
JMIR Infodemiology

Focusing on determinants and distribution of health information and misinformation on the internet, and its effect on public and individual health.

Volume 4 (2024) ISSN 2564-1891 Editor-in-Chief: Tim Ken Mackey, MAS, PhD

Contents

Original Papers

Dynamic Associations Between Centers for Disease Control and Prevention Social Media Contents and Epidemic Measures During COVID-19: Inveillance Study (e49756) Shuhua Yin, Shi Chen, Yaorong Ge.	3
Verification in the Early Stages of the COVID-19 Pandemic: Sentiment Analysis of Japanese Twitter Users (e37881) Ryuichiro Ueda, Feng Han, Hongjian Zhang, Tomohiro Aoki, Katsuhiko Ogasawara.	20
Evaluating the Disease-Related Experiences of TikTok Users With Lupus Erythematosus: Qualitative and Content Analysis (e51211) Lindsey Wanberg, David Pearson.	33
Experiences of Women With Medical Abortion Care Reflected in Social Media (VEILLE Study): Noninterventional Retrospective Exploratory Infodemiology Study (e49335) Giulia Gouy, Luisa Attali, Paméla Voillot, Patrick Fournet, Aubert Agostini.	45
The Role of Scientific Research in Human Papillomavirus Vaccine Discussions on Twitter: Social Network Analysis (e50551) Geneviève Jessiman-Perreault, Jean-Christophe Boucher, So Kim, Nicole Frenette, Abbas Badami, Henry Smith, Lisa Allen Scott.	57
Analyzing Questions About Alcohol in Pregnancy Using Web-Based Forum Topics: Qualitative Content Analysis (e58056) Nessie Frennesson, Julie Barnett, Youssouf Merouani, Angela Attwood, Luisa Zuccolo, Cheryl McQuire.	74
The Journey of Engaging With Web-Based Self-Harm and Suicide Content: Longitudinal Qualitative Study (e47699) Zoë Haime, Laura Kennedy, Lydia Grace, Rachel Cohen, Jane Derges, Lucy Biddle.	85
Exploring the Impact of the COVID-19 Pandemic on Twitter in Japan: Qualitative Analysis of Disrupted Plans and Consequences (e49699) Masaru Kamba, Wan She, Kiki Ferawati, Shoko Wakamiya, Eiji Aramaki.	105
Perceptions of Health Misinformation on Social Media: Cross-Sectional Survey Study (e51127) Anna Gaysynsky, Nicole Senft Everson, Kathryn Heley, Wen-Ying Chou.	120

TikTok as a Source of Health Information and Misinformation for Young Women in the United States: Survey Study (e54663)	
Ciera Kirkpatrick, LaRissa Lawrie.	133
The Use of Social Media to Express and Manage Medical Uncertainty in Dyskeratosis Congenita: Content Analysis (e46693)	
Emily Pearce, Hannah Raj, Ngozika Emezienna, Melissa Gilkey, Allison Lazard, Kurt Ribisl, Sharon Savage, Paul Han.	149
The Role of Social Media in Knowledge, Perceptions, and Self-Reported Adherence Toward COVID-19 Prevention Guidelines: Cross-Sectional Study (e44395)	
Camryn Garrett, Shan Qiao, Xiaoming Li.	166
Government-Nongovernmental Organization (NGO) Collaboration in Macao's COVID-19 Vaccine Promotion: Social Media Case Study (e51113)	
Xuechang Xian, Rostam Neuwirth, Angela Chang.	179
Development of a Medical Social Media Ethics Scale and Assessment of #IRad, #CardioTwitter, and #MedTwitter Posts: Mixed Methods Study (e47770)	
Vongai Mlambo, Eric Keller, Caroline Mussatto, Gloria Hwang.	195

Corrigenda and Addenda

Correction: Verification in the Early Stages of the COVID-19 Pandemic: Sentiment Analysis of Japanese Twitter Users (e57880)	
Ryuichiro Ueda, Feng Han, Hongjian Zhang, Tomohiro Aoki, Katsuhiko Ogasawara.	204

Research Letter

Using Social Listening for Digital Public Health Surveillance of Human Papillomavirus Vaccine Misinformation Online: Exploratory Study (e54000)	
Dannell Boatman, Abby Starkey, Lori Acciavatti, Zachary Jarrett, Amy Allen, Stephenie Kennedy-Rea.	206

Original Paper

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

Shuhua Yin¹, PhD; Shi Chen¹, PhD; Yaorong Ge¹, PhD

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

Corresponding Author:

Shi Chen, PhD

University of North Carolina at Charlotte

9201 University City Blvd

Charlotte, NC, 28223

United States

Phone: 1 8148800738

Email: schen56@charlotte.edu

Abstract

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

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

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

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

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

(*JMIR Infodemiology* 2024;4:e49756) doi:[10.2196/49756](https://doi.org/10.2196/49756)

KEYWORDS

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

Introduction

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

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

Infodemiology and infoveillance studies tackle health challenges, generate insights, and predict patterns and trends of diseases using previously neglected online data. Infodemiology, which is the conjunction of "information" and "epidemiology," defined by Gunther Eysenbach, is the field of distribution and determinants of information of a population through the internet or other electronic media [4]. Infoveillance takes surveillance as the primary aim and generates automated analysis from massive online data. It employs innovative computational approaches to mine and analyze unstructured online text information, such as analyzing patterns and trends, predicting potential outbreaks, and addressing current issues of public

health. Unlike traditional epidemiological surveillance systems, which include cohort studies, disease registries, population surveys, and health care records, infoveillance studies discover a wide range of health topics, monitor health issues including outbreaks and pandemics, and forecast epidemiological trends in real time. A large amount of anonymous online data can be obtained in a more timely manner with these approaches than with traditional surveillance systems, and this will help researchers and public health agencies to prepare for and tackle public health emergencies and issues more efficiently and effectively.

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

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

associations, more specifically, the CDC's COVID-19 content topic tweeting patterns and the actual COVID-19 epidemic metrics.

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

Methods

Data Acquisition and Preprocessing

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

For consistency in subsequent analyses, all CDC tweet time series and US COVID-19 variable time series were standardized

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

Natural Language Processing

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

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

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

Table 1. Data collection and preprocessing.

Variable	Data collection	Data preprocessing
CDC ^a tweets	<ul style="list-style-type: none"> Twitter API^b using a search query 17,524 English tweets by January 15, 2022 	<ul style="list-style-type: none"> Remove @mentions, hashtags, special characters, emails, punctuations, URLs, and hyperlinks Tokenization: break down sentences into individual tokens Lemmatization: each word or token is turned to its base or dictionary form Remove a list of stop words created by research experts N-grams: form phrases from the tweets Modify the date range: January 22, 2020 (the start date of reported case data) to January 15, 2022 (the end date of CDC tweets)
COVID-19 epidemic metrics	<ul style="list-style-type: none"> Public GitHub repository of the CSSE^c at Johns Hopkins University Confirmed case count: 751 records; January 22, 2020, to February 10, 2022 Death count: 908 records; January 22, 2020, to July 17, 2022 Completed COVID-19 tests: 760 records; January 13, 2020, to February 10, 2022 Complete vaccination: 428 records; December 10, 2020, to February 10, 2022 	<ul style="list-style-type: none"> Standardize metric time series to be the same as that of CDC tweets Fill missing values in the vaccination data with zeros 725 records for each of the 4 metric series Turn cumulative records to daily records

^aCDC: Centers for Disease Control and Prevention.

^bAPI: application programming interface.

^cCSSE: Center for Systems Science and Engineering.

Table 2. Subsequent analyses.

Variable	Topic modeling	Data analysis
CDC ^a tweets and COVID-19 metrics	<ul style="list-style-type: none"> Construct an LDA^b topic model using CDC tweets assigning 4 topics Extract generated topics with their top 10 unique associated keywords Produce interactive visualizations using pyLDAvis 	<ul style="list-style-type: none"> Domain experts examine topic keywords with randomly sampled tweets in iteration Domain experts determine the theme of each topic Perform multivariate time series analyses between each topic time series and each COVID-19 metric time series: <ol style="list-style-type: none"> Visualization Cross-correlation function (CCF) Mutual information (MI) Autoregressive integrated moving average with external variable (ARIMAX) model

^aCDC: Centers for Disease Control and Prevention.

^bLDA: latent Dirichlet allocation.

Topic Modeling With Latent Dirichlet Allocation

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

of topics based on both topic model evaluation and domain expert interpretations of the identified topic clusters.

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



An ideal q should have high probabilities $q(x_i)$ for the new data. Perplexity decreases as the likelihood of the words in new data

increases. Therefore, lower perplexity indicates better predictability of an LDA model.

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

$$\frac{1}{N} \sum_{i=1}^N P(w_i)$$

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

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

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

Data Preparation

Key topic time series data were derived from the previous NLP and LDA processes. We constructed a multivariate data frame with posting dates and number of tweets for each key topic at a daily resolution. Since LDA identified 4 key topics, a total of 4 CDC key topic time series were developed. There were also

4 US COVID-19 epidemic metric time series: daily cumulative reported cases, cumulative confirmed deaths, cumulative number of completed PCR tests or other approved nucleic acid amplification tests, and cumulative number of people who received a complete primary series of vaccines. These 4 sets of COVID-19 epidemic metric time series were then converted to daily measures via first order differencing. Multivariate time series analyses were implemented to investigate the associations between time series of key CDC tweet topics and US COVID-19 epidemic metrics.

Visualizations

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

Cross-Correlation Function

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

$$\frac{1}{N} \sum_{i=1}^N x_i y_i$$

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

Mutual Information

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

$$\frac{1}{N} \sum_{i=1}^N x_i y_i$$

where x_i and y_i are the i th elements of the variables X and Y , respectively. When applied to time series data, X and Y are 2

individual time series and x_i and y_i are their respective observations at the i th time step. Note that MI is a single value instead of a function over lag l as in the CCF. A larger MI value indicates a higher shared mutual dependency between the 2 time series.

Autoregressive Integrated Moving Average With External Variable

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



or in backward shift operator form:




See [Multimedia Appendix 1](#) for details on the parameters.

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



or in backward shift operator form [26]:



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

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

In addition to reporting the values of the ARIMAX model parameter set (p, d, q), difference in Akaike information criterion (dAIC), root mean square error (RMSE), and mean absolute error (MAE) were also computed to compare different ARIMAX performances. The optimal model was the one with the lowest AIC score. dAIC was computed in between 2 models (see

[Multimedia Appendix 1](#) for detailed information on AIC). We had an ARIMA model of a single topic time series and an ARIMAX model of that topic time series with an exogenous variable. Negative dAIC values indicated that the ARIMAX model showed improvement in model performance over the ARIMA counterpart that did not include an exogenous variable.

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



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

Results

Topic Modeling and Content Results

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

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

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

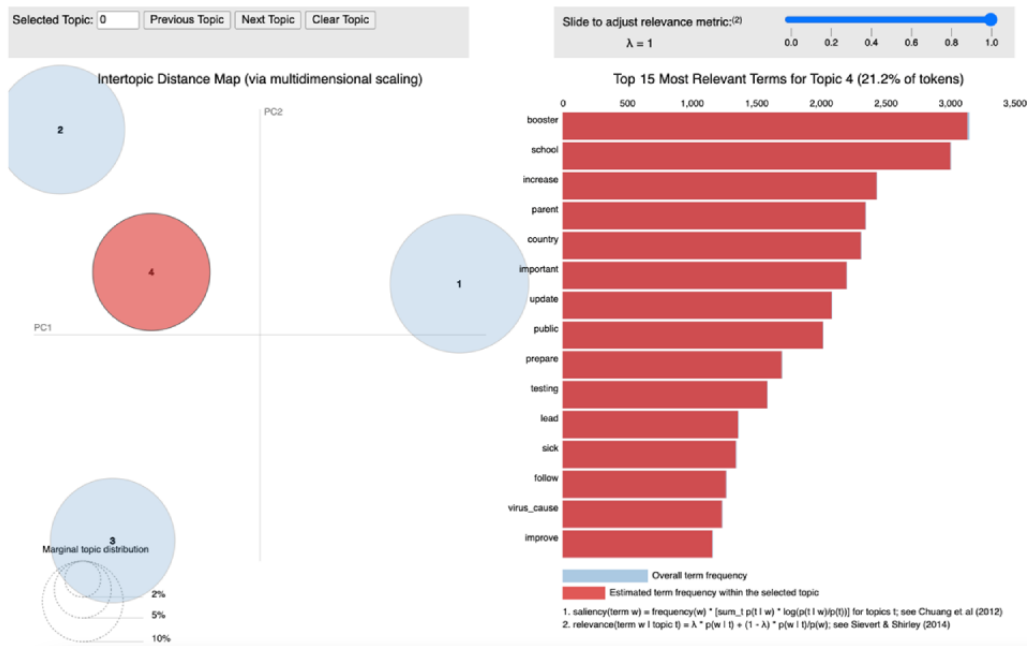
Textbox 1. Example tweets from each topic theme.

<p>Topic 1: General vaccination information and education, especially preventing adverse health outcomes of COVID-19</p> <ul style="list-style-type: none"> • “Even as the world’s attention is focused on #COVID19, this week we are taking time to highlight how #VaccinesWork and to thank the heroes who help develop and deliver lifesaving vaccines. #WorldImmunizationWeek message” • “CDC’s #COVID19 Vaccine Webinar Series is a great place to start learning about a variety of topics around COVID-19 vaccination.” • “The #DeltaVariant of the virus that causes #COVID19 is more than two times as contagious as the original strain. Wear a mask indoors in public, even if vaccinated and in an area of substantial or high transmission. Get vaccinated as soon as you can.” <p>Topic 2: Pediatric intervention, pediatric vaccination information, family safety, and school and community protection</p> <ul style="list-style-type: none"> • “Make #handwashing a family activity! Explain to children that handwashing can keep them healthy. Be a good role model—if you wash your hands often, your children are more likely to do the same. #COVID19” • “Parents: During #COVID19, well-child visits are especially important for children under 2. Schedule your child’s routine visit, so the healthcare provider can check your child’s development & provide recommended vaccines.” • “It is critically important for our public health to open schools this fall. CDC resources will help parents, teachers and administrators make practical, safety-focused decisions as this school year begins.” <p>Topic 3: Updates on COVID-19 testing, case, and death data, and relevant information of the disease</p> <ul style="list-style-type: none"> • “CDC tracks 12 different forecasting models of possible #COVID19 deaths in the US. As of May 11, all forecast an increase in deaths in the coming weeks and a cumulative total exceeding 100,000 by June 1. See national & state forecasts.” • “The latest CDC #COVIDView report shows that the percentage of #COVID19-associated deaths has been on the rise in the United States since October and has now surpassed the highest percentage seen during summer.” • “#COVID19 cases are going up dramatically. This increase is not due to more testing. As the number of cases rise, so does the percentage of tests coming back positive, which shows that COVID-19 is spreading.” <p>Topic 4: Research, study, health care, and community engagement to curb COVID-19</p> <ul style="list-style-type: none"> • “Our Nation’s medical community has been vigilant and their help in identifying confirmed cases of #COVID19 in the United States to date has been critical to containing the spread of this virus.” • “In a new report using data from Colombia, scientists found that pregnant women with symptomatic #COVID19 were at higher risk of hospitalization & death than nonpregnant women with symptomatic COVID-19. HCPs can inform pregnant women about how to stay safe.” • “A new study finds masking and fewer encounters or less time close to persons with #COVID19 can limit the spread in university settings. #MaskUp when inside indoor public places regardless of vaccination status.”

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

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

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



Multivariate Time Series Analysis Results

CCF Results

The time series of CDC tweet topics and COVID-19 metrics were plotted to visually examine patterns and potential associations. A total of 16 time series plots (4 topics × 4 COVID-19 epidemic metrics) were generated (Figures S14-S29 in Multimedia Appendix 1). CCFs were computed to quantify the dynamic association between each CDC key topic series and each of the 4 COVID-19 epidemic metrics. Quantitative results have been presented (Tables S3-S6 in Multimedia Appendix 1). Visualizations (Figures S30-S44 in Multimedia Appendix 1) illustrated CCFs between both types of time series. CCF values and plots showed that the CDC’s key COVID-19 tweet topic series was not substantially correlated with the confirmed COVID-19 case count series. As an example, there

were no specific patterns between topic 2 and daily confirmed COVID-19 cases (Figure 2A).

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

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

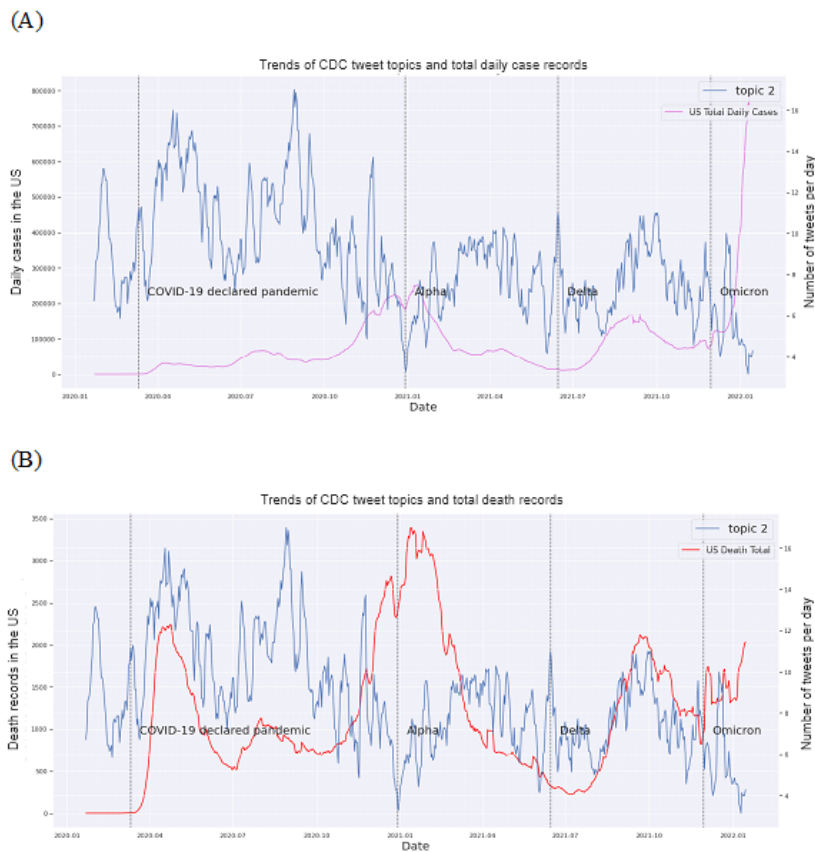


Figure 3. Cross-correlation function (CCF) between the completed COVID-19 test series and topic 2 tweets. (A) Trends of CDC tweet topics and number of completed tests; (B) CCF between COVID-19 confirmed cases and topic 2 tweets. CDC: Centers for Disease Control and Prevention.

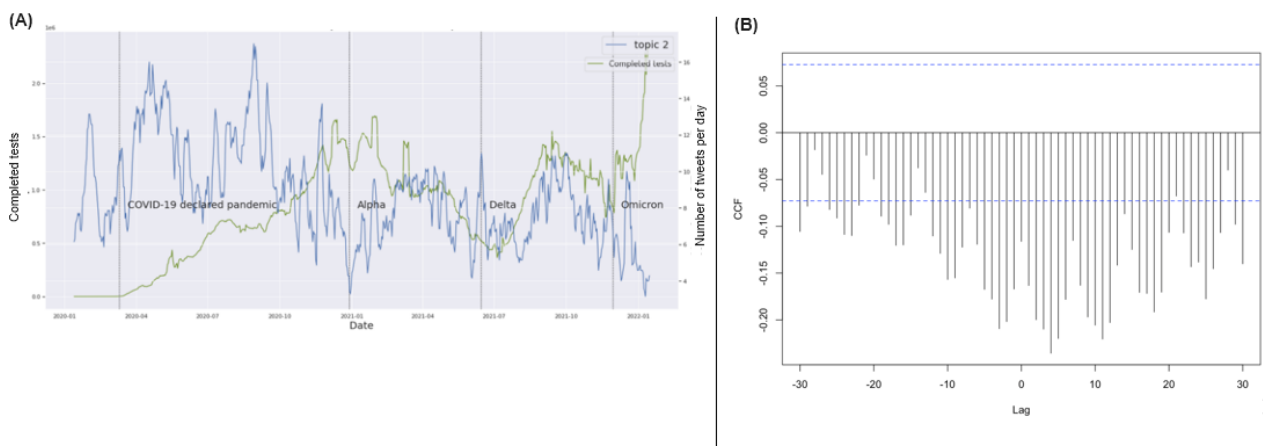
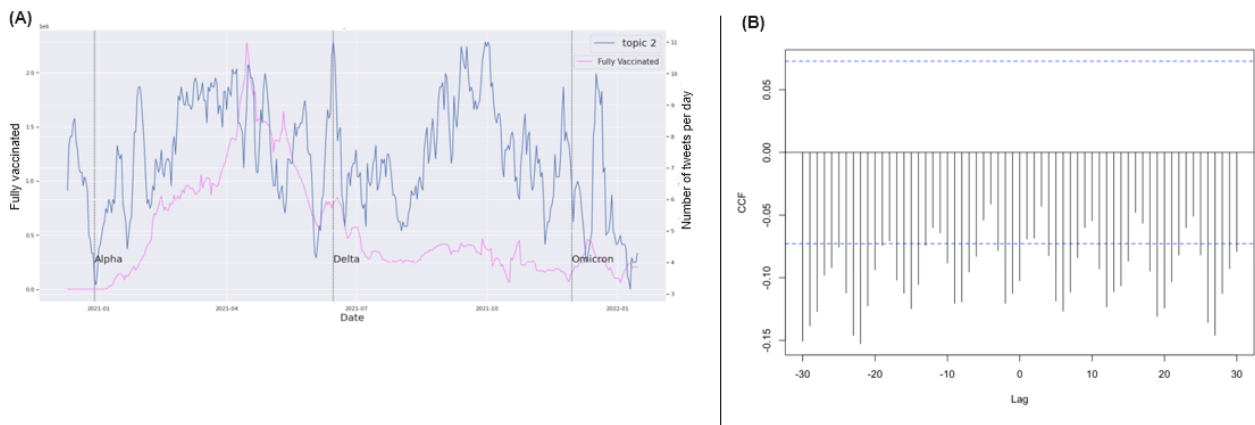


Figure 4. Cross-correlation function (CCF) between the completed COVID-19 vaccination series and topic 2 tweets. (A) Trends of CDC tweet topics and vaccination records; (B) CCF between records of fully vaccinated people and topic 2 tweets. CDC: Centers for Disease Control and Prevention.



MI Results

MI values between each CDC tweet topic and each COVID-19 metric were calculated, and they are shown in Table 4. Confirmed case counts and topic 4 (research, health care, and community engagement to restrain COVID-19) had the highest MI value (3.21), indicating that there was a strong dependency in COVID-19 cases and topic 4. On the other hand, the vaccination rate and topic 3 had the lowest MI value (0.56), indicating almost independence between the 2 series. Among all 4 key topics, topic 4 showed the highest MI values (3.21, 3.02, 3.21, and 1.65) with the 4 COVID-19 metrics. Topic 2 (pediatric intervention, family safety, and school and community protection) had consistently lower MI values with the COVID-19 metric than topic 4. The MI of topic 1 (information on COVID-19 vaccination and education on preventing its adverse health outcomes) and topic 3 (updates on COVID-19 testing, case, and death metrics, and relevant information of the disease) showed similar values with all 4 COVID-19 metrics,

although the MI values of topic 1 were slightly higher. Vaccination and educational information on the adverse health outcomes of COVID-19 appeared to not be substantially correlated with COVID-19 epidemic metrics, including testing, cases, and deaths. We speculated that the CDC considered both vaccination and preventing adverse health outcomes of COVID-19 critical to public health and disseminated these topics regardless of the current COVID-19 situation at the time of posting.

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

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

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

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

^bTopic 2: Pediatric intervention, pediatric vaccination information, family safety, and school and community protection.

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

^dTopic 4: Research, study, health care, and community engagement to curb COVID-19.

ARIMAX Results

ARIMAX performance measures, including values of ARIMAX parameters (*p*, *d*, *q*), dAIC, RMSE, and MAE, are reported in Table 5. As an external input, the vaccination rate time series significantly improved the performances of the original ARIMA models for all CDC key topics (dAIC = -108.15, -69.79, -90.54, and -91.53 for topics 1 to 4, respectively). This was the largest increase in model performance across all topics with the exogenous variable in the ARIMAX model. The COVID-19 case series improved the ARIMA model performance for CDC

topics 1 and 3 (dAIC = -104.76 and -1.53 for topics 1 and 3, respectively). Including the death or testing series did not result in substantial improvements to the ARIMA model performance for all CDC key topics.

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

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

We performed an exhaustive search to identify the optimal ARIMAX parameters (p, d, q). For example, topic 1 with death

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

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

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

^aARIMAX: autoregressive integrated moving average with external variable.

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

^cTopic 2: Pediatric intervention, pediatric vaccination information, family safety, and school and community protection.

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

^eTopic 4: Research, study, health care, and community engagement to curb COVID-19.

^fARIMAX parameters (p, d, q).

^gdAIC: delta Akaike information criterion (AIC) or difference in AIC.

^hNegative dAIC: indicates improvement of performance in the ARIMAX model compared with its autoregressive integrated moving average (ARIMA) model.

ⁱAIC values of ARIMA and its corresponding ARIMAX models.

^jRMSE: root mean square error.

^kMAE: mean absolute error.

Discussion

Principal Findings

In this study, we systematically investigated and comprehensively identified the CDC's key topics, COVID-19 epidemic metrics, and dynamic associations between the 2 types of data series based on 17,524 COVID-related English tweets from the CDC since January 2022. The LDA topic model was built to characterize and identify the dynamic shifts of topics in the CDC's COVID-19 communication over a period of more

than 2 years. For the first time, we were able to identify the following 4 key topics: (1) general vaccination information and education; (2) pediatric intervention that also involved family and school safety; (3) updates on the COVID-19 epidemic situation, such as numbers of cases, deaths, etc; and (4) research studies that were able to curb the pandemic.

Our study took a unique approach of infoveillance by identifying potential associations between COVID-19 epidemic outcome metrics in the United States and the CDC's key topic dynamics during different stages of the pandemic. This innovative

framework significantly expanded the original infoveillance approach that generally relied on the number of posts (ie, posting dynamics) without further extracting more detailed and meaningful content topics and sentiments from the textual data. Our study was able to further provide practical and useful health communication strategies for public health agencies to effectively communicate timely and accurate information to the public. It is important to investigate the dynamic associations between the CDC's tweets on COVID-19 and the progression of the pandemic for several reasons:

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

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

We then examined the dynamic associations between the 4 identified key topics and 4 COVID-19 epidemic outcome metrics. Among the 4 major topics, topic 1, which covered information on vaccination and adverse health outcomes of COVID-19, had substantially strong associations with death counts and testing records during the Alpha phase (December 29, 2020, to June 14, 2021). We found that during this phase, when the overall vaccination-related CDC tweets were decreasing, the daily vaccination rate (number of people who received a complete primary series of the COVID-19 vaccine based on the CDC Vaccine Tracker) was increasing, which

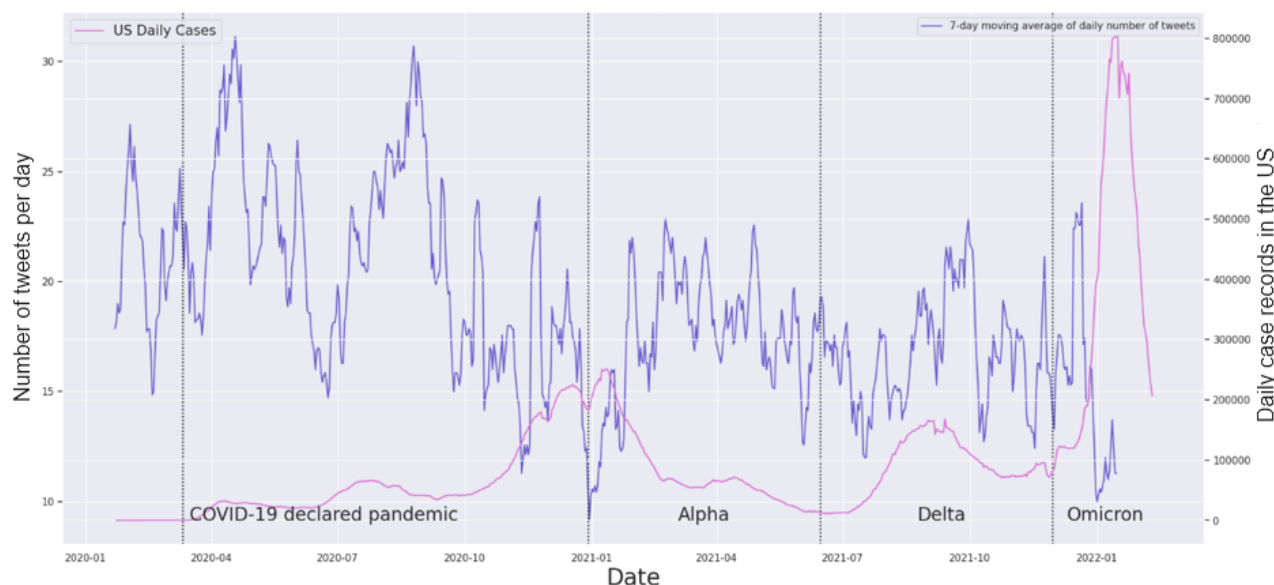
aligned with the CDC's effort in emphasizing the importance of vaccination to the public on social media. When discussions from the CDC about vaccination were increasing after the Alpha phase, the vaccination rate started to decrease. The reasons could be but are not limited to the following:

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

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

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

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



Limitations and Future Work

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

1. Time lags: What is posted might not reflect real-time situations, which can impact the association strength between the posted measures and real-world metrics; thus, we suggest aligning the content topics of health communication with up-to-date epidemic outcome metrics.
2. Discrepancies in posting methods: The CDC simplifies the data in their posts to make the information more comprehensible for the audience, which might not align with the detailed epidemic metrics posted from other sources with different interpretations of the reported metrics.
3. Variability in the data source: The data open to the public might come from sources and reporting standards that are different from the CDC's protocol, which could weaken potential associations.
4. Audience: As a government health agency, the CDC prioritizes certain data for social media to cater to the public for relevancy. For example, posting daily epidemic measures could lead to strong associations with COVID-19 metrics, but an association does not mean causality, and we assume that the CDC does not generate their tweets with

the intention to improve associations of any kind and their priority is to present a variety of reliable information to the public.

5. Fatigue from information overload: Frequent data updates on social media can lead to desensitization, making it less likely for users to pay attention and react to the information over time, for example, tweeting about daily epidemic measures decreases the public's attention over time.
6. Political and societal factors, for example, antivaccination groups and conspiracy theories about the pandemic.

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

Lastly, public engagement (ie, retweets, likes, replies, etc) of the CDC's health communication is an important indicator of the effectiveness of online health communication efforts. There have been studies that analyzed public sentiments and attitudes [31-34] toward various health-related topics. However, very few studies have investigated the associations of public sentiment shifts along disease-related metrics. In addition, public sentiments and attitudes are heavily influenced by health agencies' messages and should not be misled by misinformation. Public opinions also influence health practices and interventions, which have a significant impact on the actual epidemic outcomes (eg, case, death, vaccination, etc). Thus, it is important to further investigate the underlying association between public health communication topics and actual epidemic measures. The insights can help public health agencies develop better social media strategies to address public concerns at different stages of the emergency. Therefore, we suggest that examining the dynamics and patterns of public responses to health agencies' original communications can provide valuable insights on public perceptions and attitudes around various issues during the

pandemic, such as pharmaceutical interventions (eg, vaccination) and nonpharmaceutical interventions. Detailed content analysis can be applied to explicitly identify public concerns in response to the CDC's health communications. In addition, sentiment analysis can be applied to extract public sentiments (ie, positive, neutral, or negative) toward the CDC's health communications, and further help identify public attitudes and reactions to various content topics that the CDC has communicated. Public attitudes will ultimately determine individual health behavior and decision-making, such as vaccination acceptance and compliance with nonpharmaceutical interventions, which in turn drive the overall epidemic dynamics. Therefore, it is critical to investigate real-time public engagement, such as retweeting or replying on social media, toward public health agencies' communications to better inform health agencies about prioritizing their communications and addressing public concerns about specific content topics.

Conclusions

This study investigated the dynamic associations between the CDC's detailed COVID-19 communication topics on Twitter and epidemic metrics in the United States for almost 2 years during the pandemic. Using LDA topic modeling, we were the first to comprehensively identify and explore various COVID-related topics tweeted by the federal public health agency during the pandemic. We also collected daily COVID-19 epidemic metrics (confirmed case counts, death counts,

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

Acknowledgments

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

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary information.

[PDF File (Adobe PDF File), 3131 KB - [infodemiology_v4i1e49756_app1.pdf](#)]

References

1. WHO Coronavirus (COVID-19) Dashboard. World Health Organization. URL: <https://covid19.who.int/> [accessed 2023-11-11]
2. The Health Communicator's Social Media Toolkit. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/healthcommunication/toolstemplates/socialmediatoolkit_bm.pdf [accessed 2023-11-11]
3. Yun GW, Morin D, Park S, Joa CY, Labbe B, Lim J, et al. Social media and flu: Media Twitter accounts as agenda setters. *Int J Med Inform* 2016 Jul;91:67-73. [doi: [10.1016/j.ijmedinf.2016.04.009](https://doi.org/10.1016/j.ijmedinf.2016.04.009)] [Medline: [27185510](https://pubmed.ncbi.nlm.nih.gov/27185510/)]
4. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. *J Med Internet Res* 2009 Mar 27;11(1):e11 [FREE Full text] [doi: [10.2196/jmir.1157](https://doi.org/10.2196/jmir.1157)] [Medline: [19329408](https://pubmed.ncbi.nlm.nih.gov/19329408/)]
5. Liu Q, Zheng Z, Zheng J, Chen Q, Liu G, Chen S, et al. Health Communication Through News Media During the Early Stage of the COVID-19 Outbreak in China: Digital Topic Modeling Approach. *J Med Internet Res* 2020 Apr 28;22(4):e19118 [FREE Full text] [doi: [10.2196/19118](https://doi.org/10.2196/19118)] [Medline: [32302966](https://pubmed.ncbi.nlm.nih.gov/32302966/)]
6. Al-Ramahi M, Elnoshokaty A, El-Gayar O, Nasralah T, Wahbeh A. Public Discourse Against Masks in the COVID-19 Era: Infodemiology Study of Twitter Data. *JMIR Public Health Surveill* 2021 Apr 05;7(4):e26780 [FREE Full text] [doi: [10.2196/26780](https://doi.org/10.2196/26780)] [Medline: [33720841](https://pubmed.ncbi.nlm.nih.gov/33720841/)]

7. Abd-Alrazaq A, Alhuwail D, Househ M, Hamdi M, Shah Z. Top Concerns of Tweeters During the COVID-19 Pandemic: Infoveillance Study. *J Med Internet Res* 2020 Apr 21;22(4):e19016 [FREE Full text] [doi: [10.2196/19016](https://doi.org/10.2196/19016)] [Medline: [32287039](https://pubmed.ncbi.nlm.nih.gov/32287039/)]
8. Chen S, Xu Q, Buchenberger J, Bagavathi A, Fair G, Shaikh S, et al. Dynamics of Health Agency Response and Public Engagement in Public Health Emergency: A Case Study of CDC Tweeting Patterns During the 2016 Zika Epidemic. *JMIR Public Health Surveill* 2018 Nov 22;4(4):e10827 [FREE Full text] [doi: [10.2196/10827](https://doi.org/10.2196/10827)] [Medline: [30467106](https://pubmed.ncbi.nlm.nih.gov/30467106/)]
9. COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. GitHub. URL: <https://github.com/CSSEGISandData/COVID-19/blob/master/README.md> [accessed 2021-12-31]
10. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases* 2020 May;20(5):533-534. [doi: [10.1016/s1473-3099\(20\)30120-1](https://doi.org/10.1016/s1473-3099(20)30120-1)]
11. Dong E, Ratcliff J, Goyea TD, Katz A, Lau R, Ng TK, et al. The Johns Hopkins University Center for Systems Science and Engineering COVID-19 Dashboard: data collection process, challenges faced, and lessons learned. *The Lancet Infectious Diseases* 2022 Dec;22(12):e370-e376. [doi: [10.1016/s1473-3099\(22\)00434-0](https://doi.org/10.1016/s1473-3099(22)00434-0)]
12. A Note on Topical N-grams. University of Massachusetts Amherst. URL: <https://people.cs.umass.edu/~mccallum/papers/tng-tr05.pdf> [accessed 2023-11-11]
13. Rajaraman A, Ullman JD. Data Mining. In: *Mining of Massive Datasets*. Cambridge, United Kingdom: Cambridge University Press; 2011.
14. Bafna P, Pramod D, Vaidya A. Document clustering: TF-IDF approach. 2016 Presented at: International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT); March 03-05, 2016; Chennai, India. [doi: [10.1109/ICEEOT.2016.7754750](https://doi.org/10.1109/ICEEOT.2016.7754750)]
15. Perplexity. Wikipedia. URL: <https://en.wikipedia.org/wiki/Perplexity> [accessed 2023-11-11]
16. Stevens K, Kegelmeyer P, Andrzejewski D, Buttler D. Exploring topic coherence over many models and many topics. In: *EMNLP-CoNLL '12: Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*. 2012 Presented at: 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning; July 12-14, 2012; Jeju Island, Korea. [doi: [10.5555/2390948.2391052](https://doi.org/10.5555/2390948.2391052)]
17. Topic coherence pipeline. Gensim: topic modeling for humans. Radim Řehůřek. URL: <https://radimrehurek.com/gensim/models/coherencemodel.html#gensim.models.coherencemodel.CoherenceModel> [accessed 2021-12-31]
18. Tijare P, Jhansi Rani P. Exploring popular topic models. *J. Phys.: Conf. Ser* 2020 Dec 01;1706(1):012171. [doi: [10.1088/1742-6596/1706/1/012171](https://doi.org/10.1088/1742-6596/1706/1/012171)]
19. Derrick TR, Thomas JM. Time-Series Analysis: The cross-correlation function. In: Stergiou N, editor. *Innovative Analyses of Human Movement*. Champaign, Illinois: Human Kinetics Publishers; 2004:189-205.
20. statsmodels.tsa.stattools.adfuller. statsmodels. URL: <https://www.statsmodels.org/dev/generated/statsmodels.tsa.stattools.adfuller.html> [accessed 2021-12-31]
21. Safarnejad L, Xu Q, Ge Y, Bagavathi A, Krishnan S, Chen S. Identifying Influential Factors in the Discussion Dynamics of Emerging Health Issues on Social Media: Computational Study. *JMIR Public Health Surveill* 2020 Jul 28;6(3):e17175 [FREE Full text] [doi: [10.2196/17175](https://doi.org/10.2196/17175)] [Medline: [32348275](https://pubmed.ncbi.nlm.nih.gov/32348275/)]
22. Time Series Measures. GitHub. URL: <https://elife-asu.github.io/PyInform/timeseries.html> [accessed 2021-12-31]
23. Non-seasonal ARIMA models. OTexts. URL: <https://otexts.com/fpp2/non-seasonal-arima.html> [accessed 2023-04-09]
24. ARIMA modelling in R. OTexts. URL: <https://otexts.com/fpp2/arima-r.html> [accessed 2023-04-11]
25. Rabbi F, Tareq SU, Islam MM, Chowdhury MA, Kashem MA. A Multivariate Time Series Approach for Forecasting of Electricity Demand in Bangladesh Using ARIMAX Model. 2020 Presented at: 2nd International Conference on Sustainable Technologies for Industry 4.0 (STI); December 19-20, 2020; Dhaka, Bangladesh. [doi: [10.1109/STI50764.2020.9350326](https://doi.org/10.1109/STI50764.2020.9350326)]
26. Yang M, Xie J, Mao P, Wang C, Ye Z. Application of the ARIMAX Model on Forecasting Freeway Traffic Flow. 2017 Presented at: 17th COTA International Conference of Transportation Professionals; July 7-9, 2017; Shanghai, China. [doi: [10.1061/9780784480915.061](https://doi.org/10.1061/9780784480915.061)]
27. Hodson TO. Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. *Geosci. Model Dev* 2022 Jul 19;15(14):5481-5487. [doi: [10.5194/gmd-15-5481-2022](https://doi.org/10.5194/gmd-15-5481-2022)]
28. TMO Guide (3) – pyLDavis. WE1S project. URL: <https://we1s.ucsb.edu/research/we1s-tools-and-software/topic-model-observatory/tmo-guide/tmo-guide-pyldavis/> [accessed 2023-04-09]
29. Chuang J, Manning CD, Heer J. Termite: Visualization Techniques for Assessing Textual Topic Models. 2012 Presented at: International Working Conference on Advanced Visual Interfaces; May 21-25, 2012; Capri Island, Italy.
30. Sievert C, Shirley K. LDavis: A method for visualizing and interpreting topics. 2014 Presented at: Workshop on Interactive Language Learning, Visualization, and Interfaces; June 2014; Baltimore, MA, USA. [doi: [10.3115/v1/W14-3110](https://doi.org/10.3115/v1/W14-3110)]
31. Mohamed Ridhwan K, Hargreaves CA. Leveraging Twitter data to understand public sentiment for the COVID - 19 outbreak in Singapore. *International Journal of Information Management Data Insights* 2021 Nov;1(2):100021 [FREE Full text] [doi: [10.1016/j.jjime.2021.100021](https://doi.org/10.1016/j.jjime.2021.100021)]

32. Hu T, Wang S, Luo W, Zhang M, Huang X, Yan Y, et al. Revealing Public Opinion Towards COVID-19 Vaccines With Twitter Data in the United States: Spatiotemporal Perspective. *J Med Internet Res* 2021 Sep 10;23(9):e30854 [FREE Full text] [doi: [10.2196/30854](https://doi.org/10.2196/30854)] [Medline: [34346888](https://pubmed.ncbi.nlm.nih.gov/34346888/)]
33. Chandrasekaran R, Desai R, Shah H, Kumar V, Moustakas E. Examining Public Sentiments and Attitudes Toward COVID-19 Vaccination: Inveillance Study Using Twitter Posts. *JMIR Infodemiology* 2022 Apr 15;2(1):e33909 [FREE Full text] [doi: [10.2196/33909](https://doi.org/10.2196/33909)] [Medline: [35462735](https://pubmed.ncbi.nlm.nih.gov/35462735/)]
34. Shen L, Yao R, Zhang W, Evans R, Cao G, Zhang Z. Emotional Attitudes of Chinese Citizens on Social Distancing During the COVID-19 Outbreak: Analysis of Social Media Data. *JMIR Med Inform* 2021 Mar 16;9(3):e27079 [FREE Full text] [doi: [10.2196/27079](https://doi.org/10.2196/27079)] [Medline: [33724200](https://pubmed.ncbi.nlm.nih.gov/33724200/)]

Abbreviations

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

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

Please cite as:

Yin S, Chen S, Ge Y

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

JMIR Infodemiology 2024;4:e49756

URL: <https://infodemiology.jmir.org/2024/1/e49756>

doi: [10.2196/49756](https://doi.org/10.2196/49756)

PMID: [38261367](https://pubmed.ncbi.nlm.nih.gov/38261367/)

©Shuhua Yin, Shi Chen, Yaorong Ge. Originally published in *JMIR Infodemiology* (<https://infodemiology.jmir.org>), 23.01.2024. 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.

Original Paper

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

Ryuichiro Ueda¹, MHA; Feng Han², MHA; Hongjian Zhang¹, PhD; Tomohiro Aoki¹, MHA; Katsuhiko Ogasawara¹, MBA, PhD

¹Faculty of Health Sciences, Hokkaido University, Sapporo, Japan

²Graduate School of Medicine, Hokkaido University, Sapporo, Japan

Corresponding Author:

Katsuhiko Ogasawara, MBA, PhD

Faculty of Health Sciences

Hokkaido University

N12-W5, Kita-ku

Sapporo, 060-0812

Japan

Phone: 81 11 706 3409

Email: oga@hs.hokudai.ac.jp

Related Article:

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

Abstract

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

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

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

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

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

(*JMIR Infodemiology* 2024;4:e37881) doi:[10.2196/37881](https://doi.org/10.2196/37881)

KEYWORDS

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

Introduction

Background

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

COVID-19 spread rapidly in Japan after the first infection was confirmed on January 16, 2020, with incidents such as the mass infection on the Diamond Princess cruise ship in early February

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

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

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

Prior Work in Infodemiology

Following the spread of COVID-19, social networking services (SNSs) were used to transmit information about the virus, accelerating activity in the field of infodemiology, which utilizes this data. Infodemiology is a relatively new research field that combines health informatics and public health with data analysis. It is a scientific discipline that studies the distribution of information and its determinants in information media, particularly the internet, to provide reliable information on public health [11]. Infodemiology became widely known after the World Health Organization (WHO) used the term at the first WHO Infodemiology Conference in response to the spread of COVID-19 and stated the need to promote research activities in this field worldwide [12]. In a previous study, Su et al [13] used sentimental analysis of text information from SNS data to reflect public concerns and psychological changes in individuals, providing information to promote public health. In particular, a sentiment analysis of the Italian region of Lombardy, where the lockdown was enforced, indicated that the number of SNS

users with feelings of “anxiety” decreased after the lockdown. In addition, Heras-Pedrosa et al [14] observed through sentiment analysis that “anxiety” and “anger” toward government policies were the top feelings in Spain in the early stages of the infection. Furthermore, in Japan, Niu et al [15] conducted a sentiment analysis from SNS text data on the reasons for the delay in COVID-19 vaccine uptake compared to other countries, suggesting that concerns about side effects may have outweighed the fear of infection in the initial vaccination process. Thus, social media-based analysis reflects the psychological changes in individuals and enables the provision of real-time information to the government enacting public health policies and infection prevention measures.

SNS Usage in Japan

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

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

Study Purpose

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

Methods

Research Data

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

Data Preprocessing

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

Morphological Analysis

In contrast to structured and quantitative data, which can be easily analyzed by a computer, qualitative text data, which are

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

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

Japanese Sentiment Dictionary

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

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

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

^aJIWC: Japanese Linguistic Inquiry and Word Count.

Data Clustering

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

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

cost for predictions would become immense. As a result, we employed the CBOW model for word embedding processing. After the data collected from Twitter and the terms registered from each Japanese sentiment dictionary were vectorized, Fuzzy-C-Means (FCM) was used to cluster each of the 7 sentiments.

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



In this case, the following conditions are satisfied:



The matrix U , denoted as $[u_{it}]$, represents an $n \times c$ matrix with the membership value u_{it} as an element. Meanwhile, the matrix V , represented as $[v_{jt}]$, is an $n \times c$ matrix with cluster center v_{jt} as an element.

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



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

$$v_t^{(p)} = \frac{\sum_{i=1}^n u_{it}^{(p)m} x_i}{\sum_{i=1}^n u_{it}^{(p)m}$$

Third, we update the membership values from $u_{it}^{(p)}$ to $u_{it}^{(p+1)}$ using the following formula:

$$u_{it}^{(p+1)} = \frac{1}{\sum_{t=1}^c \left(\frac{d_{it}^{(p)}}{d_{it}^{(p)}} \right)^{\frac{1}{m-1}}$$

Finally, if $\|u_{it}^{(p+1)} - u_{it}^{(p)}\| < \epsilon$ holds for all i and t , we terminate the loop. Otherwise, we increment p by 1 and return

to the second step. Once the loop terminates, we obtain the center points for each cluster and the membership values for each sample data, completing the clustering process. In this study, FCM was used on text data to reduce the number of words included in an emotion dictionary and construct the emotion dictionary, allowing for more accurate sentiment analysis of the text data due to the influence of a single word on multiple emotions. Both tweets and retweets of Twitter data were used, and quoted retweets, which are retweets of others' posts with additional comments, were also included.

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

Figure 1. Sentiment analysis flowchart.

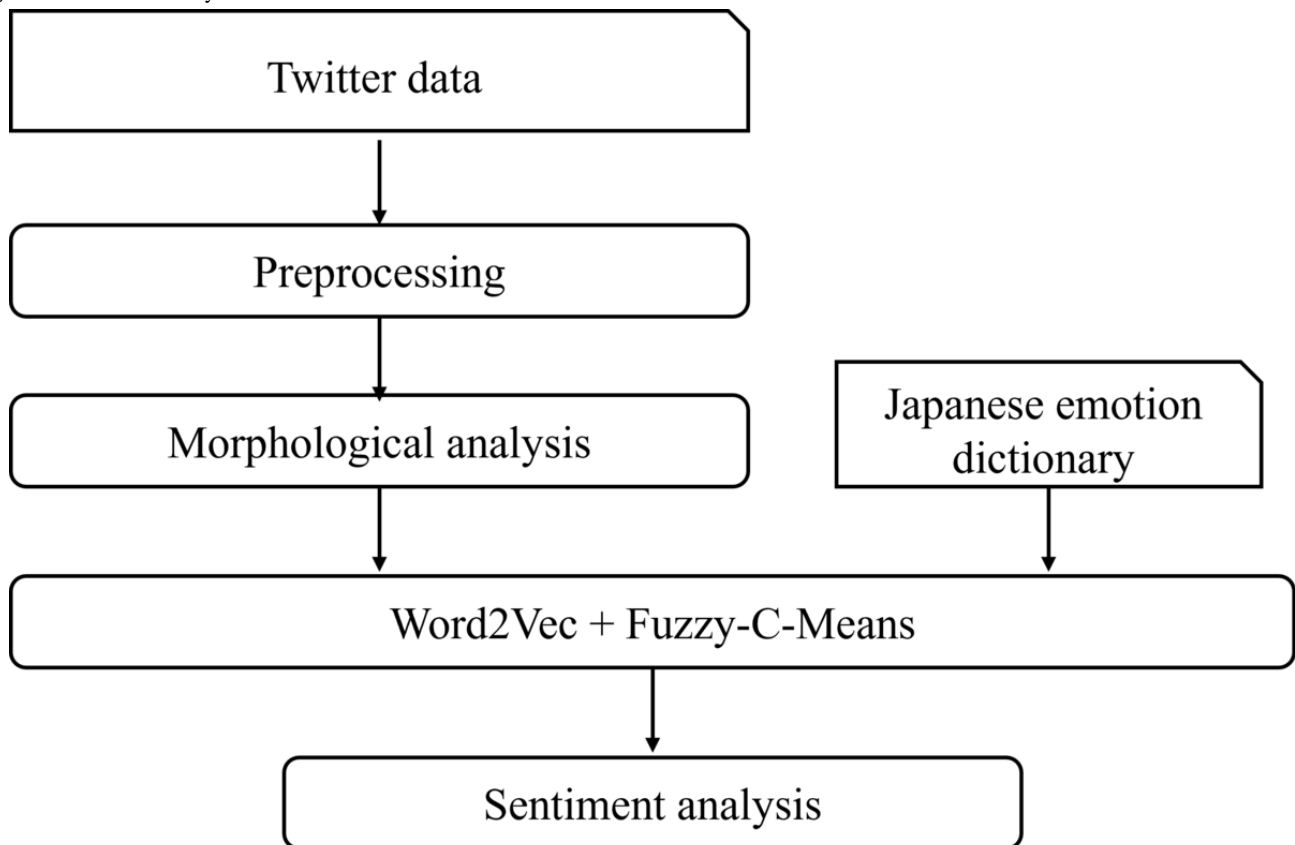
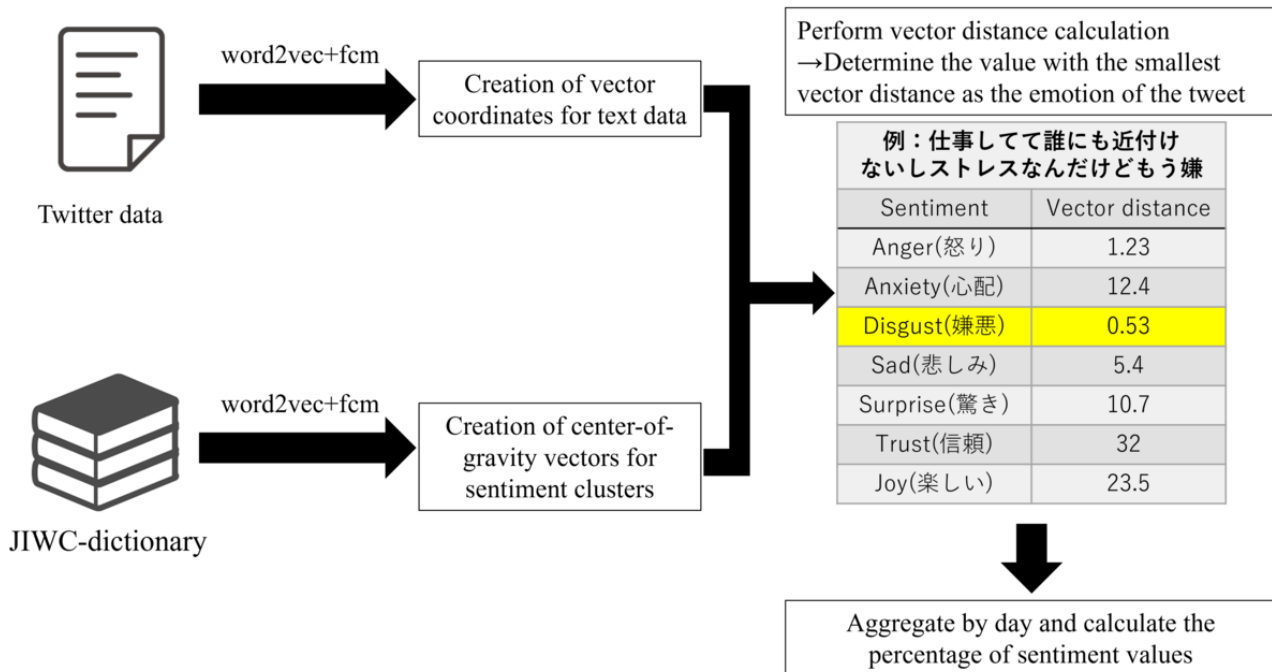


Figure 2. Diagram of the sentiment determination method. FCM: Fuzzy-C-Means; JIWC: Japanese Linguistic Inquiry and Word Count.



Examining Sentiment Changes Before and After the State of Emergency Declaration

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

conducted a median difference examination using the Mann-Whitney U test, with statistical significance set at $P < .05$, utilizing the statistical software JMP (version 16.0; SAS).

Ethical Considerations

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

Results

Research Data

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

Table 3. Daily tweet and retweet counts.

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

Percentage of Emotions in the Sentiment Analysis

The results of the sentiment analysis on the tweet and retweet groups for the period between midnight on March 23, 2020, to 23:59:59 on April 6, 2020 (before the declaration of the state of emergency) are shown in [Figure 3](#). The results for the period between midnight on April 7, 2020, and 23:59:59 on April 21, 2020 (after the declaration of the state of emergency) are shown

in [Figure 4](#). In the tweet group, the positive emotion “joy” was highest both before and after the state of emergency declaration at 40.5% (n=272,879) and 31% (n=217,074), respectively, while in the retweet group, the negative sentiment of “worry” was 34% (n=587,540), and “disgust” was 18.6% (n=322,462) during the period before the state of emergency declaration. These percentages were higher than those for the other emotions.

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

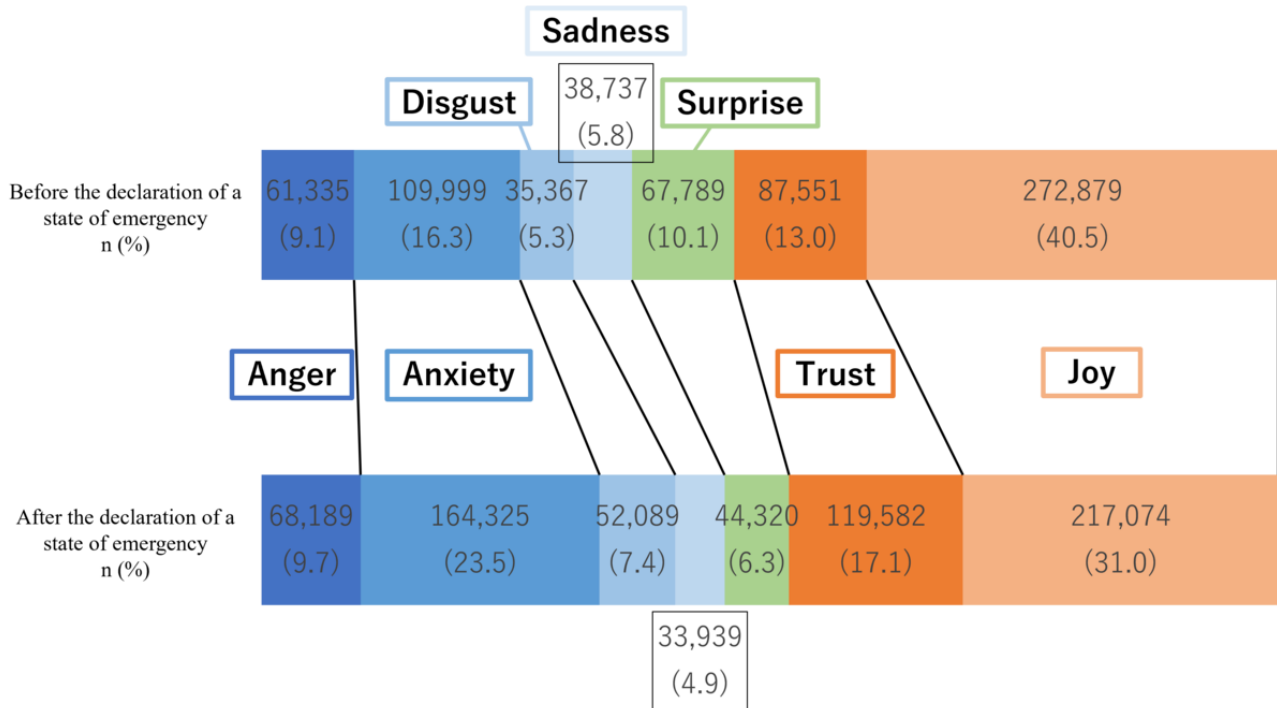
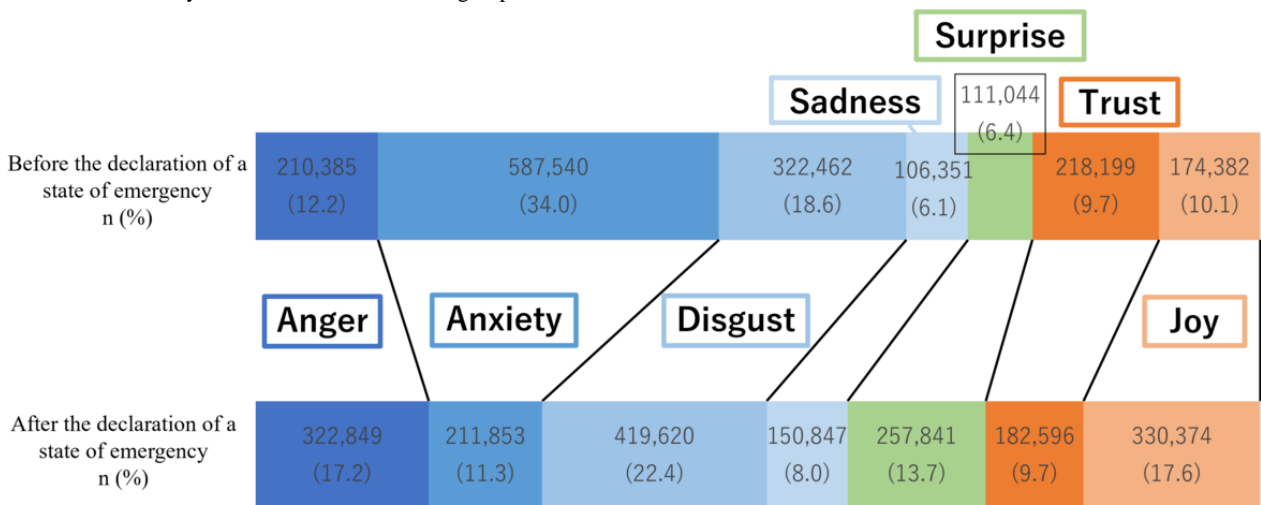


Figure 4. Sentiment analysis value ratio in the retweet group.



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

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

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

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

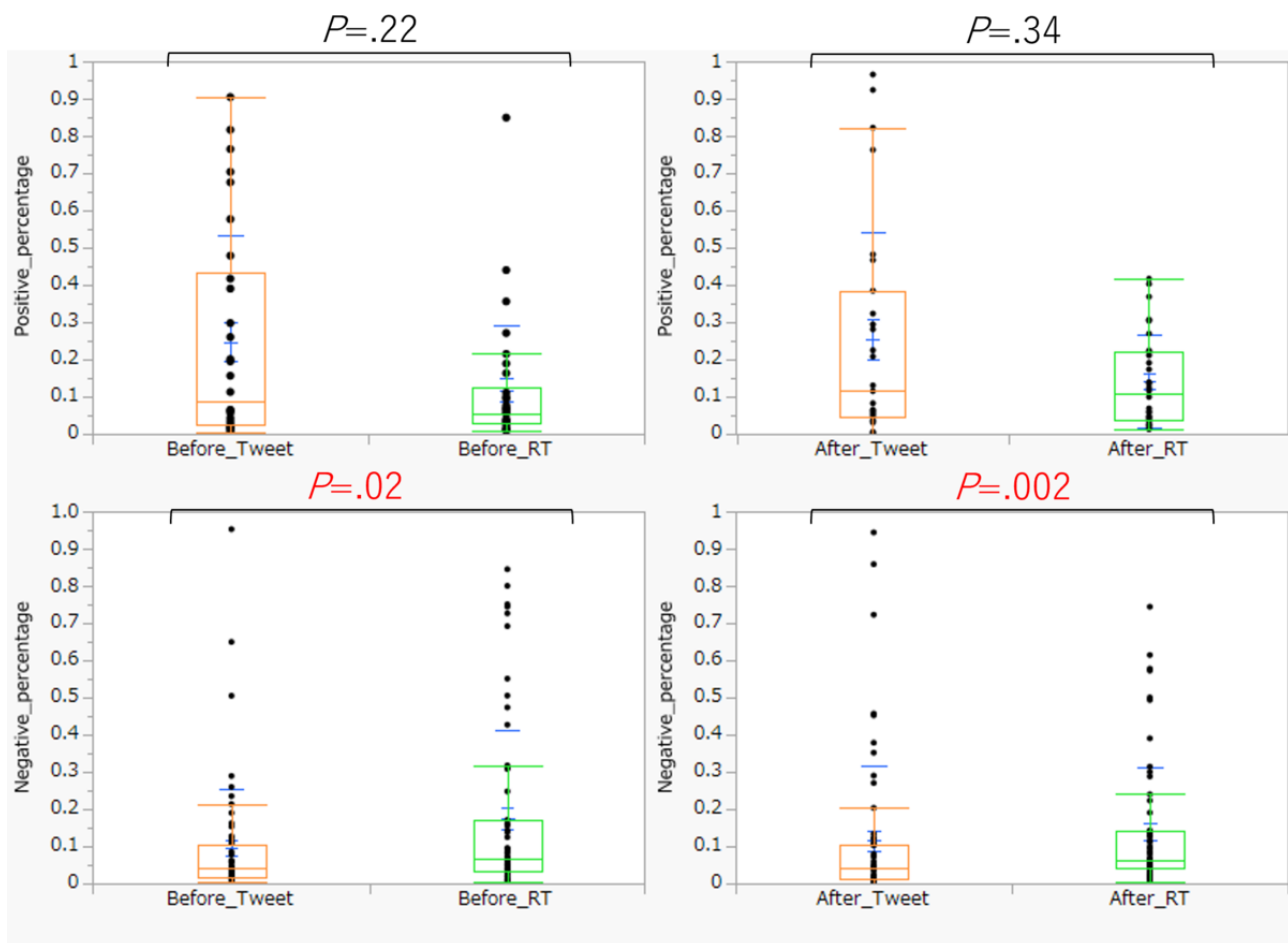
Sentiments	Before (n=15)		After (n=15)		P value
	Median	SD	Median	SD	
Anger tweet	0.042	0.061	0.024	0.082	.80
Anger retweet	0.042	0.051	0.063	0.035	.84
Anxiety tweet	0.063	0.078	0.050	0.135	.90
Anxiety retweet	0.210	0.307	0.054	0.052	.25
Disgust tweet	0.023	0.293	0.021	0.020	.59
Disgust retweet	0.073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

^aBefore refers to the period from midnight on March 23, 2020, until 11:59:59 PM on April 6, 2020, while after refers to the period from midnight on April 7, 2020, until 11:59:59 PM on April 21, 2020.

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

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

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



Discussion

Principal Findings

The Japanese language sentiment analysis conducted during this study's target period, both before and after the declaration of the state of emergency, revealed that "joy," associated with a "positive" sentiment, accounted for high proportions within the tweet group at 40.5% ($n=272,879$) before and 31% ($n=217,074$) after. On the other hand, "anxiety" and "disgust," which express "negative" feelings, accounted for high percentages in both the tweet and retweet groups, especially in the retweet group, where "anxiety" accounted for 34% ($n=587,540$) and "disgust" accounted for 18.6% ($n=322,464$) of the total retweets before the state of emergency was declared. The self-restraint approach regulating behavior during the declaration of a state of emergency in Japan allowed movement across prefectures. This may have been a contributing factor to the widespread negative posts related to movements from the target area. This surge in negative sentiment was countered by a simultaneous rise in positive emotions, attributed to the anticipation of infection prevention following the state of emergency declaration. During the early stages of the COVID-19 spread in other countries, a previous study on English-speaking users indicated elevated levels of positive emotions linked to anticipations for potential policies [22]. A generally similar emotional response was apparent among the public in other

countries. In the early stages of the spread of infection, when no vaccine or other countermeasures had been implemented, feelings of anxiety may have been expressed on social media, as well as expectations for strong countermeasures, such as behavioral restrictions. In contrast, the results of the sentiment analysis of English-language tweets corresponding to the same period showed that negative and positive emotions accounted for approximately the same proportions by late March, the end of the period covered in this study. Notably, the negative emotion "fear" occupied a higher percentage than other emotions around January and February [27]. In China and European countries, the first cases of infection were confirmed earlier than in Japan (where the initial expansion of the outbreak occurred in late March). Thus, the earlier spread of infection in those nations may have a significant impact on the sentiment analysis.

Comparative Study Between the Tweet and Retweet Groups

When comparing the tweet and retweet groups, the retweet group tended to post more negative sentiments. In this regard, a previous study revealed that in the early stage of the COVID-19 outbreak among English-speaking users, many tweets had a positive sentiment, while many retweets had a negative sentiment [28]. It is clear that much of the information users wished to disseminate was negative in nature. As for the difference between groups in this study, there is a research report

that focuses on virality, one of the characteristics of sentiment analysis using social media [29]. Virality is an explosive spread of attention and information through social media and word-of-mouth on the internet. Virality is derived from “viral”—as in a virus. Previous research indicates that negative posts increase virality, while positive posts decrease virality. Therefore, for topics that attract substantial public attention, such as COVID-19, the topic of this study, there is a tendency to spread negative content in retweets, consequently increasing virality. This suggests a noteworthy contrast between the tweet and retweet groups.

Limitations

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

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

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

The Twitter data utilized in this study underwent random sampling for both tweets and retweets. Twitter incorporates a function known as “bot,” which automatically generates tweets in response to specified times and keywords. Numerous accounts, commonly referred to as “bot accounts,” are responsible for automatic posting. Shi et al [31] conducted a sentiment analysis on Twitter focusing on the #coronavirus

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

Conclusions

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

Acknowledgments

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

Conflicts of Interest

None declared.

Multimedia Appendix 1

Total number of tweets and retweets per day.

[PNG File, 180 KB - [infodemiology_v4i1e37881_app1.png](#)]

References

1. Yamazaki T. Epidemiology of the early stages of the case outbreak in the Diamond Princess novel coronavirus infection case. IASR. 2020 Jul 31. URL: <https://www.niid.go.jp/niid/images/idsc/iasr/41/485.pdf> [accessed 2024-01-05]
2. Declaration of a state of emergency in response to the novel coronavirus disease. Prime Minister's Office of Japan. 2020 Apr 16. URL: https://japan.kantei.go.jp/ongoingtopics/_00020.html [accessed 2023-11-09]
3. Soubhik B, Stefano C, Antonio D, Paolo F, Thiemo F, Stefano F. Evaluating COVID-19 public health messaging in Italy: Self-reported compliance and growing mental health concerns. medRxiv. Preprint posted online on April 5, 2020 [FREE Full text] [doi: [10.1101/2020.03.27.20042820](https://doi.org/10.1101/2020.03.27.20042820)]
4. Outbreak of a case of pneumonia associated with novel coronavirus (case 1). Ministry of Health, Labour and Welfare. 2020 Jan 16. URL: https://www.mhlw.go.jp/stf/newpage_08906.html [accessed 2023-11-09]
5. Task force on new coronavirus infections (15th meeting). Prime Minister's Office of Japan. 2020 Feb 27. URL: https://www.kantei.go.jp/jp/98_abe/actions/202002/27corona.html [accessed 2023-11-09]
6. Regarding the "historical emergency". Cabinet Office Minister's Secretariat Document Management Division. 2020 Mar 10. URL: <https://www8.cao.go.jp/chosei/koubun/rekiren/rekishitekikinkyujitai.pdf> [accessed 2024-01-05]
7. Influenza countermeasures special measures law. Ministry of Health, Labour and Welfare. URL: https://www.mhlw.go.jp/web/t_doc?dataId=78ab2871&dataType=0 [accessed 2024-01-05]
8. Task force on new coronavirus infections (23th meeting). Prime Minister's Office of Japan. URL: https://www.kantei.go.jp/jp/98_abe/actions/202003/26corona.html [accessed 2023-11-09]
9. Facts about novel coronavirus infection. Cabinet Office. URL: https://www8.cao.go.jp/shougai/whitepaper/r03hakusho/zenbun/h2_01_02.html [accessed 2023-11-09]
10. Task force on new coronavirus infections (29th meeting). Prime Minister's Office of Japan. URL: https://www.kantei.go.jp/jp/98_abe/actions/202004/16corona.html [accessed 2023-11-09]
11. Zielinski C. Infodemics and infodemiology: a short history, a long future. Rev Panam Salud Publica 2021 Feb 23;45:e40. [doi: [10.26633/rpsp.2021.40](https://doi.org/10.26633/rpsp.2021.40)]
12. 1st WHO Infodemiology Conference. World Health Organization. URL: <https://www.who.int/news-room/events/detail/2020/06/30/default-calendar/1st-who-infodemiology-conference> [accessed 2023-11-09]
13. Su Y, Xue J, Liu X, Wu P, Chen J, Chen C, et al. Examining the impact of COVID-19 lockdown in Wuhan and Lombardy: a psycholinguistic analysis on Weibo and Twitter. Int J Environ Res Public Health 2020 Jun 24;17(12) [FREE Full text] [doi: [10.3390/ijerph17124552](https://doi.org/10.3390/ijerph17124552)] [Medline: [32599811](https://pubmed.ncbi.nlm.nih.gov/32599811/)]
14. de Las Heras-Pedrosa C, Sánchez-Núñez P, Peláez JI. Sentiment analysis and emotion understanding during the COVID-19 pandemic in Spain and its impact on digital ecosystems. Int J Environ Res Public Health 2020 Jul 31;17(15):5542 [FREE Full text] [doi: [10.3390/ijerph17155542](https://doi.org/10.3390/ijerph17155542)] [Medline: [32751866](https://pubmed.ncbi.nlm.nih.gov/32751866/)]
15. Niu Q, Liu J, Kato M, Shinohara Y, Matsumura N, Aoyama T, et al. Public opinion and sentiment before and at the beginning of COVID-19 vaccinations in Japan: Twitter analysis. JMIR Infodemiology 2022 May 9;2(1):e32335 [FREE Full text] [doi: [10.2196/32335](https://doi.org/10.2196/32335)] [Medline: [35578643](https://pubmed.ncbi.nlm.nih.gov/35578643/)]
16. White paper 2021: Information and Communications in Japan. Ministry of Internal Affairs and Communications. URL: <https://www.soumu.go.jp/johotsusintokei/whitepaper/eng/WP2021/2021-index.html> [accessed 2023-11-09]
17. Nazuki no Oto. NTT Data Corporation. URL: <https://nazuki-oto.com/main/> [accessed 2023-11-09]
18. Tomoya M, Mamoru K. Why is the Japanese language so difficult? Analysis of a realistic corpus of learners of Japanese and issues in language processing. IPSJ 2012 May:217-223.
19. Takuya K, Kaoru Y, Yuji M. Applying conditional random fields to Japanese morphological analysis. In: Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing. 2004 Presented at: 2004 Conference on Empirical Methods in Natural Language Processing; July 25-26; Barcelona, Spain p. 230-237 URL: <https://aclanthology.org/W04-3230>
20. Sato T. Neologism dictionary based on language resources on the web for Mecab. GitHub. 2015. URL: <https://github.com/neologd/mecab-ipadic-neologd> [accessed 2023-11-09]
21. JIWC dictionary. GitHub. URL: <https://github.com/sociocom/JIWC-Dictionary> [accessed 2023-11-09]
22. Boon-Itt S, Skunkan Y. Public perception of the COVID-19 pandemic on Twitter: sentiment analysis and topic modeling study. JMIR Public Health Surveill 2020 Nov 11;6(4):e21978 [FREE Full text] [doi: [10.2196/21978](https://doi.org/10.2196/21978)] [Medline: [33108310](https://pubmed.ncbi.nlm.nih.gov/33108310/)]
23. Tomas M, Kai C, Greg C, Jeffrey D. Efficient estimation of word representations in vector space. arXiv. Preprint posted online on January 16, 2013 [FREE Full text] [doi: [10.48550/arXiv.1301.3781](https://doi.org/10.48550/arXiv.1301.3781)]

24. Tomas M, Ilya S, Kai C, Greg C, Jeffrey D. Distributed representations of words and phrases and their compositionality. arXiv. Preprint posted online on October 16, 2013 [[FREE Full text](#)]
25. Zadeh L. Fuzzy sets. *Information and Control* 1965 Jun;8(3):338-353. [doi: [10.1016/s0019-9958\(65\)90241-x](https://doi.org/10.1016/s0019-9958(65)90241-x)]
26. Bezdek JC. Pattern recognition with fuzzy objective function algorithms. In: *Advanced Applications in Pattern Recognition*. NY: Springer New York; 1981:43.
27. Lwin MO, Lu J, Sheldenkar A, Schulz PJ, Shin W, Gupta R, et al. Global sentiments surrounding the COVID-19 pandemic on Twitter: analysis of Twitter trends. *JMIR Public Health Surveill* 2020 May 22;6(2):e19447 [[FREE Full text](#)] [doi: [10.2196/19447](https://doi.org/10.2196/19447)] [Medline: [32412418](https://pubmed.ncbi.nlm.nih.gov/32412418/)]
28. Xue J, Chen J, Chen C, Zheng C, Li S, Zhu T. Public discourse and sentiment during the COVID 19 pandemic: Using Latent Dirichlet Allocation for topic modeling on Twitter. *PLoS One* 2020 Sep 25;15(9):e0239441 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0239441](https://doi.org/10.1371/journal.pone.0239441)] [Medline: [32976519](https://pubmed.ncbi.nlm.nih.gov/32976519/)]
29. Jiménez-Zafra SM, Sáez-Castillo AJ, Conde-Sánchez A, Martín-Valdivia MT. How do sentiments affect virality on Twitter? *R Soc Open Sci* 2021 Apr 14;8(4):201756 [[FREE Full text](#)] [doi: [10.1098/rsos.201756](https://doi.org/10.1098/rsos.201756)] [Medline: [33996120](https://pubmed.ncbi.nlm.nih.gov/33996120/)]
30. 2020 Survey on information and communication media usage time and information behavior. Ministry of Internal Affairs and Communications. URL: https://www.soumu.go.jp/main_content/000765258.pdf [accessed 2023-11-10]
31. Shi W, Liu D, Yang J, Zhang J, Wen S, Su J. Social bots' sentiment engagement in health emergencies: a topic-based analysis of the COVID-19 pandemic discussions on Twitter. *Int J Environ Res Public Health* 2020 Nov 23;17(22):8701 [[FREE Full text](#)] [doi: [10.3390/ijerph17228701](https://doi.org/10.3390/ijerph17228701)] [Medline: [33238567](https://pubmed.ncbi.nlm.nih.gov/33238567/)]
32. Yum S. Social network analysis for coronavirus (COVID-19) in the United States. *Soc Sci Q* 2020 Jul 28;101(4):1642-1647 [[FREE Full text](#)] [doi: [10.1111/ssqu.12808](https://doi.org/10.1111/ssqu.12808)] [Medline: [32836475](https://pubmed.ncbi.nlm.nih.gov/32836475/)]

Abbreviations

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

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

Please cite as:

Ueda R, Han F, Zhang H, Aoki T, Ogasawara K
Verification in the Early Stages of the COVID-19 Pandemic: Sentiment Analysis of Japanese Twitter Users
JMIR Infodemiology 2024;4:e37881
URL: <https://infodemiology.jmir.org/2024/1/e37881>
doi: [10.2196/37881](https://doi.org/10.2196/37881)
PMID: [38127840](https://pubmed.ncbi.nlm.nih.gov/38127840/)

©Ryuichiro Ueda, Feng Han, Hongjian Zhang, Tomohiro Aoki, Katsuhiko Ogasawara. Originally published in *JMIR Infodemiology* (<https://infodemiology.jmir.org>), 06.02.2024. 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.

Original Paper

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

Lindsey J Wanberg¹, BA; David R Pearson^{1,2}, MD

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

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

Corresponding Author:

David R Pearson, MD

University of Minnesota Medical School

516 Delaware St SE

Minneapolis, MN, MMC 98

United States

Phone: 1 612 626 8625

Email: pearsond@umn.edu

Abstract

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

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

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

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

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

(*JMIR Infodemiology* 2024;4:e51211) doi:[10.2196/51211](https://doi.org/10.2196/51211)

KEYWORDS

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

Introduction

Background

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

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

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

Objectives

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

Methods

Data Collection

A new TikTok account was created for data collection to avoid TikTok algorithms that prioritize videos based on prior user activity [13,14]. The account was used to search for #lupus on the TikTok app on August 21, 2022; the app then displayed the

most popular videos tagged with #lupus [13]. The links, captions, usernames, likes, comments, views, and shares were extracted from each video identified through this search. Videos were then evaluated for inclusion and exclusion criteria. Videos were included if they were downloadable, were in English, and involved the personal opinions or experiences of an individual with LE. They were excluded if they did not relate to LE or were primarily about a condition other than LE.

Codebook Development

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

Data Analysis

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

Ethical Considerations

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

Results

Overview

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

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

experiences with treatment (58/153, 37.9%). *Experiences with symptoms* and *mucocutaneous symptoms* had the greatest number of cumulative views (24,426,874 and 14,082,409 views, respectively; [Table 1](#)).

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

Table 1. Top codes and their popularity metrics.

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

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

Themes and TikTok user quotes

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

In another video, a user made the following announcement:

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

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

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

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

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

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

Textbox 2. Common hashtags explained.**Hashtag and description**

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

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

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

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

Natural medicine saved my life...[I was] told that I would be on pharmaceuticals for life and that I would never be able to exercise again, I could barely walk, I wouldn't be able to have kids, I wouldn't be able to have a job...But luckily, I didn't listen. Because if I did, I don't know where I would be today. Instead, I

run three different businesses, I found movement that works for me, I've completely reversed all fertility issues, I'm in remission from lupus, and most of my markers are completely normal. The secret to my healing, you may be asking? Well, it's the food that you eat, the herbs that you put into your body, and the habits that you practice on a daily basis that set the course for your entire life. [User 26]

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

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

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

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

Another user stated as follows:

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

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

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

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

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

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

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

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

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

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

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

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

Theme 5: Although TikTok Users Frequently Had Strong Support Systems, Feelings of Isolation or

Misunderstanding Were Often Attributed to Battling an “Invisible Illness”

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

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

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

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

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

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

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

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

Discussion

Principal Findings

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

information for clinicians that may ultimately improve patient care.

In this study, we found that TikTok videos on LE experiences have extensive audiences, garnering millions of views and high user engagement. Videos that mentioned mucocutaneous symptoms of LE, such as hair loss and rash, were pervasive among this sample, with *mucocutaneous symptoms* being the second most frequent code used. Consistent with findings of

previous qualitative and survey studies, mucocutaneous symptoms seemed to be major drivers of poor mental health and negative body image among TikTok users with LE [27-33]. Our study suggests the need to regularly assess for mental health and body image concerns in patients with LE, especially among those with active dermatologic symptoms. It also underscores the importance of treating hair loss and rash to mitigate mental health burdens in this population (Textbox 3).

Textbox 3. Clinical applications of themes.

Themes and clinical applications

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

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

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

Our sample had high symptom burden and frequently described how musculoskeletal symptoms and fatigue were interfering with daily functioning. TikTok users did not always feel that others understood these symptoms, leading to feelings of isolation and the thought that they have an "invisible illness." These ideas have been described in numerous qualitative

analyses on LE [3,7,24,31,36,41-43]. Even so, overall, TikTok users demonstrated robust social support systems made up of friends, family, and partners. This is important because social support has been associated with improved mental health in patients with SLE, whereas a lack of substantial social support has been associated with increased disease activity [24,44-47]. These combined findings suggest the importance of educating patients' support people on LE morbidity and disability to facilitate successful support relationships (Textbox 3).

Strengths

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

Limitations

Although qualitative studies are inherently not designed to be generalizable because they provide rich, narrative data from the group being studied, it is important to note that this study only examined TikTok content and did not examine content from

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

Conclusions

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

Authors' Contributions

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

Conflicts of Interest

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

Multimedia Appendix 1

Final codebook.

[DOCX File, 34 KB - [infodemiology_v4i1e51211_app1.docx](#)]

References

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

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

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

Abbreviations

LE: lupus erythematosus

SLE: systemic lupus erythematosus

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

Please cite as:

Wanberg LJ, Pearson DR

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

JMIR Infodemiology 2024;4:e51211

URL: <https://infodemiology.jmir.org/2024/1/e51211>

doi: [10.2196/51211](https://doi.org/10.2196/51211)

PMID: [38631030](https://pubmed.ncbi.nlm.nih.gov/38631030/)

©Lindsey J Wanberg, David R Pearson. Originally published in JMIR Infodemiology (<https://infodemiology.jmir.org>), 17.04.2024. 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.

Original Paper

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

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

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

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

³Kap Code, Paris, France

⁴Centre de Santé SoMeD, Rouen, France

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

* all authors contributed equally

Corresponding Author:

Giulia Gouy, MD

Service de Gynécologie-Obstétrique

Hôpital de la Croix-Rousse

103 Gd Rue de la Croix-Rousse

Lyon, 69004

France

Phone: 33 472 071 936

Email: giulia.gouy@chu-lyon.fr

Abstract

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

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

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

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

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

(*JMIR Infodemiology* 2024;4:e49335) doi:[10.2196/49335](https://doi.org/10.2196/49335)

KEYWORDS

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

Introduction

Background

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

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

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

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

Analyzing online information represents a developing alternative means to understand patients' health compared with self-administered questionnaires. These patient-generated health data are produced spontaneously (and are thus not limited to

medical consultations, for instance), mostly anonymously. Therefore, these data may better correspond to patients' feelings compared with closed-ended questions. Moreover, text-mining techniques applied to analyze social media data can be used with relative ease [12], providing new opportunities to bridge the gap between qualitative and quantitative data analyses [13]. A new research discipline and methodology has thus emerged. This scientific discipline called infodemiology focuses on health-related content analysis published online [14].

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

Objective

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

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

Methods

Study Design

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

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

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

Data Extraction

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

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

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

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

Data Preprocessing and Modeling

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

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

Data Analysis

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

A *topic model* was applied to identify the topics addressed in the posts constituting the study data set [24]. Topic models consist of text-mining approaches that aim to automatically identify the abstract topics addressed in a collection of

documents. Such models are based on the hypothesis that each document corresponds to a distribution of several topics.

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

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

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

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

Ethical Considerations

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

adopted as all usernames, web forum names, geographic locations, URLs, or any other sensitive information was substituted by identifiers before being stored in the analysis corpus.

Results

Population and Posts

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

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

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

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

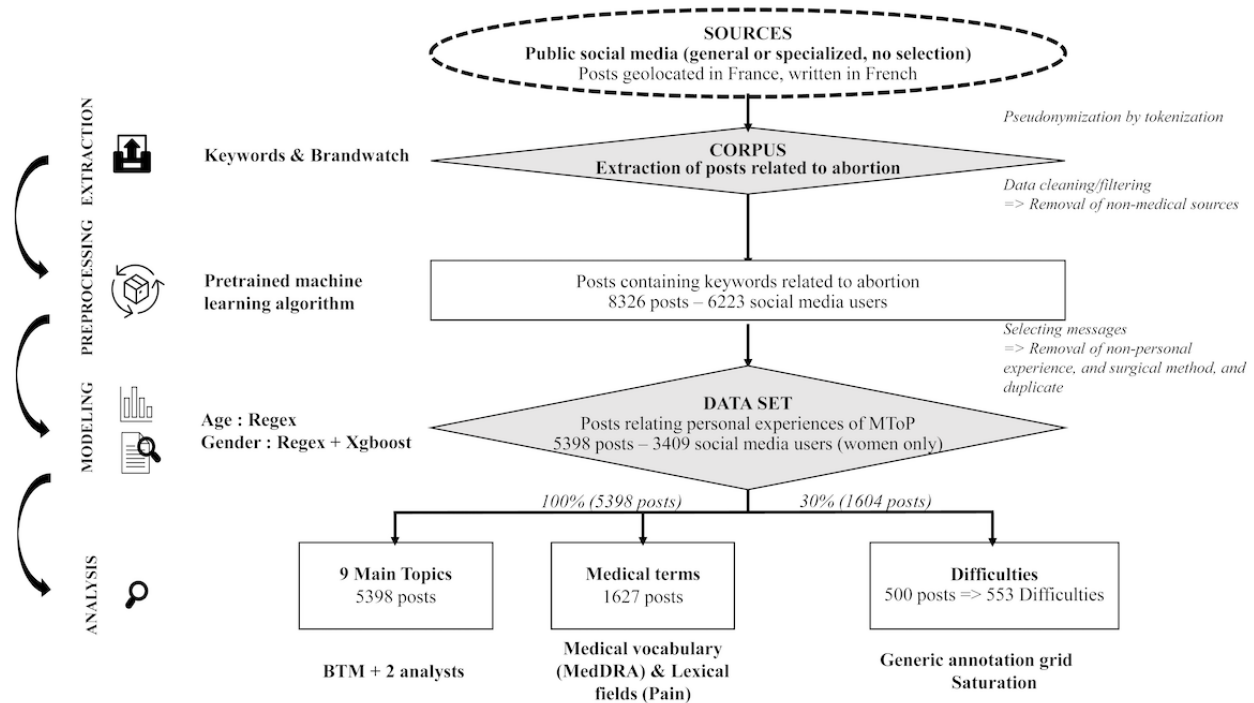


Table 1. List of forums reporting women's medical termination of pregnancy experiences and the top 6 extraction keywords.

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

^aIVG: French abbreviation for voluntary termination of pregnancy.

^bmedicament: French word for medicine.

^cN/A: not applicable.

^dAvort: first letters of the French word *avortement*, which means abortion.

Discussion Topics

From the 5398 posts, 9 topics of interest were identified (Table 2). Personal experience and community support were the most

prominent topics. The 7 other topics were as follows (in decreasing order): pain and bleeding, psychological experience, questioned efficacy, social pressure, positive experiences, menstrual cycle disorders, and reported inefficacy.

Table 2. Topics and topic description ranked by frequency.^a

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

^aA single post may contain several topics.

Medical Terms

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

emotional distress was associated with the feeling of regret and grief.

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

Table 3. Most frequently reported medical terms related to difficulties after medical termination of pregnancy.^a

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

^aA single post may contain several medical terms.

Figure 2. Focus on pain medical terms.

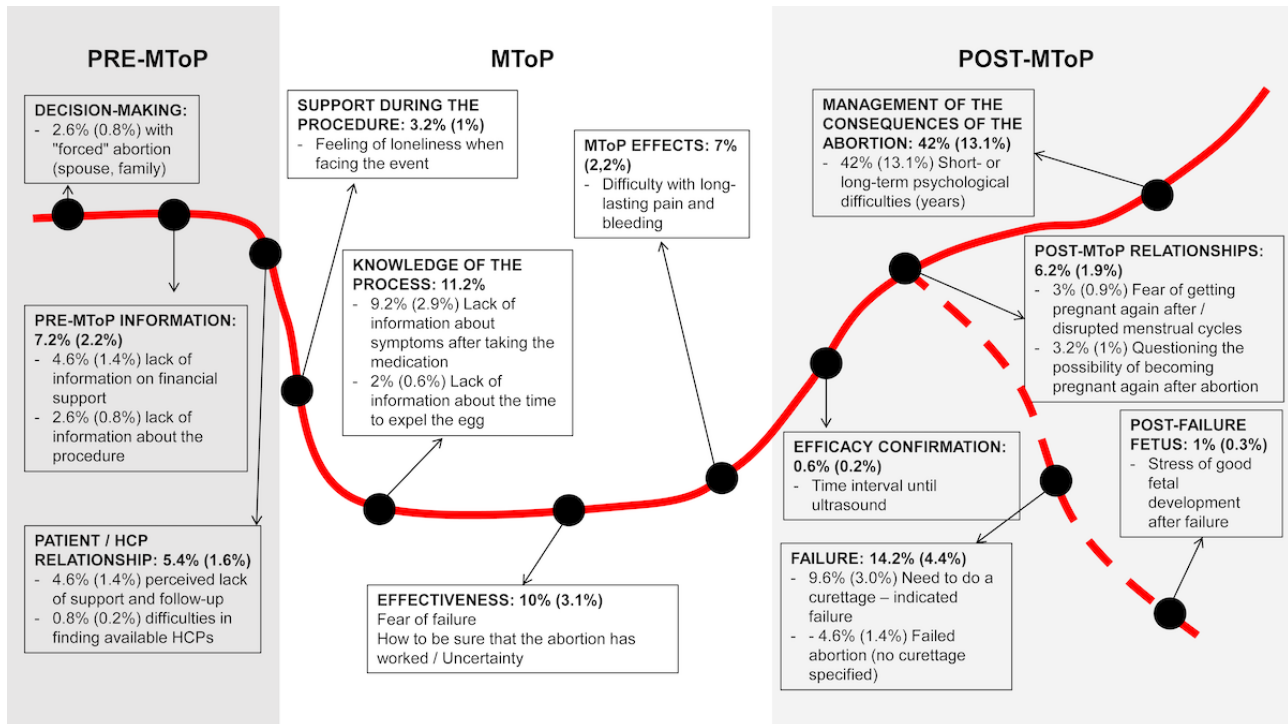


Encountered Difficulties

A total of 553 difficulties were identified from 500 posts derived from a randomized sample containing 1604 posts (30%).

Reported difficulties were encountered along the MToP process (Figure 3).

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



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

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

Overall, 5.4% of the 553 encountered difficulties (n=30) highlighted a lack of health care support during and after the procedure. Users reported that they were looking for health care

providers who did not judge them and who could provide them with an environment of mutual trust.

Discussion

Principal Findings

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

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

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

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

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

Study Strengths

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

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

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

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

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

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

Study Limitations

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

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

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

Moreover, our analysis was based on the spontaneous testimonies of social media users on a single topic of interest and according to their feelings (subjectivity). The media or influential people could direct the discussions and encourage a

peak of comments at a given moment (eg, the change of legislation around abortion in the United States).

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

Conclusion

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

Acknowledgments

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

Data Availability

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

Conflicts of Interest

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

Multimedia Appendix 1

Keywords for data extraction.

[PNG File , 301 KB - [infodemiology_v4i1e49335_app1.png](#)]

Multimedia Appendix 2

Data extraction filters.

[PNG File , 33 KB - [infodemiology_v4i1e49335_app2.png](#)]

References

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

6. Rapport public d'évaluation Mifégyne 200 mg, comprimé (mifépristone). Agence Française de Sécurité Sanitaire des Produits de Santé. 2007. URL: https://archive.ansm.sante.fr/var/ansm_site/storage/original/application/776b24988710e52725e3584edb4cdc2a.pdf [accessed 2023-04-21]
7. IVG médicamenteuse jusqu'à 9 semaines: pérenniser ce droit pour les femmes. Haute Autorité de Santé (HAS). 2021 Apr 12. URL: <https://tinyurl.com/mr58pvnz> [accessed 2024-01-30]
8. Divay S. L'avortement: une déviance légale. *Deviance Société* 2004;28(2):195-209. [doi: [10.3917/ds.282.0195](https://doi.org/10.3917/ds.282.0195)]
9. Donegan K, Ovelgonne H, Flores G, Fuglerud P, Georgescu A. Social media and m-Health data. Subgroup report. European Medicines Agency. URL: https://www.ema.europa.eu/en/documents/report/social-media-m-health-data-subgroup-report_en.pdf [accessed 2023-04-21]
10. Audience Internet Global en Février 2024. Médiamétrie. 2024 Mar 26. URL: <https://www.mediametrie.fr/fr/audience-internet-global-en-fevrier-2024> [accessed 2024-04-05]
11. Cordonier L. Information et santé. Analyse des croyances et comportements d'information des Français liés à leur niveau de connaissances en santé, au refus vaccinal et au renoncement médical. Fondation Descartes. 2023. URL: <https://www.fondationdescartes.org/2023/10/information-et-sante/> [accessed 2024-01-30]
12. McDonald L, Malcolm B, Ramagopalan S, Syrad H. Real-world data and the patient perspective: the PROMISE of social media? *BMC Med* 2019 Jan 16;17(1):11 [FREE Full text] [doi: [10.1186/s12916-018-1247-8](https://doi.org/10.1186/s12916-018-1247-8)] [Medline: [30646913](https://pubmed.ncbi.nlm.nih.gov/30646913/)]
13. Wiedemann G. Opening up to big data: computer-assisted analysis of textual data in social sciences. *Forum Qualitative Sozialforschung* 2013 May 25;14(2):23. [doi: [10.17169/fqs-14.2.1949](https://doi.org/10.17169/fqs-14.2.1949)]
14. Eysenbach G. Infodemiology: the epidemiology of (mis)information. *Am J Med* 2002 Dec;113(9):763-765. [doi: [10.1016/S0002-9343\(02\)01473-0](https://doi.org/10.1016/S0002-9343(02)01473-0)]
15. Perone H, Herweck A, Stump H, Levine H, Wong A, Carugno J. The virtual infertility community: a qualitative analysis of patient experiences shared on Instagram. *J Assist Reprod Genet* 2021 Mar;38(3):613-620 [FREE Full text] [doi: [10.1007/s10815-020-02028-6](https://doi.org/10.1007/s10815-020-02028-6)] [Medline: [33411326](https://pubmed.ncbi.nlm.nih.gov/33411326/)]
16. Mercier R, Senter K, Webster R, Henderson Riley A. Instagram users' experiences of miscarriage. *Obstet Gynecol* 2020 Jan;135(1):166-173. [doi: [10.1097/AOG.0000000000003621](https://doi.org/10.1097/AOG.0000000000003621)] [Medline: [31809440](https://pubmed.ncbi.nlm.nih.gov/31809440/)]
17. Vazquez Corona M, Betrán AP, Bohren M. The portrayal and perceptions of cesarean section in Mexican media Facebook pages: a mixed-methods study. *Reprod Health* 2022 Feb 22;19(1):49 [FREE Full text] [doi: [10.1186/s12978-022-01351-8](https://doi.org/10.1186/s12978-022-01351-8)] [Medline: [35193590](https://pubmed.ncbi.nlm.nih.gov/35193590/)]
18. Tapi Nzali MD, Bringay S, Lavergne C, Mollevi C, Opitz T. What patients can tell us: topic analysis for social media on breast cancer. *JMIR Med Inform* 2017 Jul 31;5(3):e23 [FREE Full text] [doi: [10.2196/medinform.7779](https://doi.org/10.2196/medinform.7779)] [Medline: [28760725](https://pubmed.ncbi.nlm.nih.gov/28760725/)]
19. Jacques L, Carpenter E, Valley T, Alvarez B, Higgins J. Medication or surgical abortion? An exploratory study of patient decision making on a popular social media platform. *Am J Obstet Gynecol* 2021 Sep;225(3):344-347 [FREE Full text] [doi: [10.1016/j.ajog.2021.05.011](https://doi.org/10.1016/j.ajog.2021.05.011)] [Medline: [34022196](https://pubmed.ncbi.nlm.nih.gov/34022196/)]
20. Kapp N, Lohr PA. Modern methods to induce abortion: safety, efficacy and choice. *Best Pract Res Clin Obstet Gynaecol* 2020 Feb;63:37-44. [doi: [10.1016/j.bpobgyn.2019.11.008](https://doi.org/10.1016/j.bpobgyn.2019.11.008)] [Medline: [32029379](https://pubmed.ncbi.nlm.nih.gov/32029379/)]
21. Industry leading social media management solution. Brandwatch. URL: <https://www.brandwatch.com/suite/social-media-management/> [accessed 2023-04-21]
22. Talmatkadi M, Foulquié P, Déguilhem A, Renner S, Châteauneuf L, Mebarki A, et al. Création d'un algorithme d'identification d'expériences vécues par des patients ou leurs proches à partir de messages issus des réseaux sociaux : un cas d'usage sur le COVID long. *Rev Épidémiol Santé Publique* 2022 Nov;70:S265. [doi: [10.1016/j.respe.2022.09.045](https://doi.org/10.1016/j.respe.2022.09.045)]
23. Schäfer F, Faviez C, Voillot P, Foulquié P, Najm M, Jeanne J, et al. Mapping and modeling of discussions related to gastrointestinal discomfort in French-speaking online forums: results of a 15-year retrospective infodemiology study. *J Med Internet Res* 2020 Nov 03;22(11):e17247 [FREE Full text] [doi: [10.2196/17247](https://doi.org/10.2196/17247)] [Medline: [33141087](https://pubmed.ncbi.nlm.nih.gov/33141087/)]
24. Voillot P, Riche B, Portafax M, Foulquié P, Gedik A, Barbarot S, et al. Social media platforms listening study on atopic dermatitis: quantitative and qualitative findings. *J Med Internet Res* 2022 Jan 28;24(1):e31140 [FREE Full text] [doi: [10.2196/31140](https://doi.org/10.2196/31140)] [Medline: [35089160](https://pubmed.ncbi.nlm.nih.gov/35089160/)]
25. Renner S, Loussikian P, Foulquié P, Arnould B, Marrel A, Barbier V, et al. Perceived unmet needs in patients living with advanced bladder cancer and their caregivers: infodemiology study using data from social media in the United States. *JMIR Cancer* 2022 Sep 20;8(3):e37518 [FREE Full text] [doi: [10.2196/37518](https://doi.org/10.2196/37518)] [Medline: [36125861](https://pubmed.ncbi.nlm.nih.gov/36125861/)]
26. Article R1121-1. French Public Health Code. 2021. URL: https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000034696952/
27. Baum SE, Ramirez AM, Larrea S, Filippa S, Egwuatu I, Wydrzynska J, et al. "It's not a seven-headed beast": abortion experience among women that received support from helplines for medication abortion in restrictive settings. *Health Care Women Int* 2020 Oct;41(10):1128-1146. [doi: [10.1080/07399332.2020.1823981](https://doi.org/10.1080/07399332.2020.1823981)] [Medline: [33156737](https://pubmed.ncbi.nlm.nih.gov/33156737/)]
28. Attali L. Psychological aspects of abortion. *J Gynecol Obstet Biol Reprod* 2016 Dec;45(10):1552-1567. [doi: [10.1016/j.jgyn.2016.09.030](https://doi.org/10.1016/j.jgyn.2016.09.030)] [Medline: [28029465](https://pubmed.ncbi.nlm.nih.gov/28029465/)]
29. Attali L, Guceve E, Schoch F, Bettahar K, Warynski F. Histoires d'IVG, Histoires de femmes. In: *Recherches Familiales*. Paris: Vuibert; Jan 2021:144.

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

Abbreviations

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

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

Please cite as:

Gouy G, Attali L, Voillot P, Fournet P, Agostini A

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

JMIR Infodemiology 2024;4:e49335

URL: <https://infodemiology.jmir.org/2024/1/e49335>

doi: [10.2196/49335](https://doi.org/10.2196/49335)

PMID: [38696232](https://pubmed.ncbi.nlm.nih.gov/38696232/)

©Giulia Gouy, Luisa Attali, Paméla Voillot, Patrick Fournet, Aubert Agostini. Originally published in *JMIR Infodemiology* (<https://infodemiology.jmir.org>), 02.05.2024. 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.

Original Paper

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

Geneviève Jessiman-Perreault¹, BA, MSc; Jean-Christophe Boucher², PhD; So Youn Kim², MDSA; Nicole Frenette¹, MSW; Abbas Badami², MEng; Henry M Smith², MA; Lisa K Allen Scott^{1,3,4}, PhD

¹Alberta Health Services, Calgary, AB, Canada

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

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

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

Corresponding Author:

Geneviève Jessiman-Perreault, BA, MSc

Alberta Health Services

2210 2 St SW

Calgary, AB, T2S 3C3

Canada

Phone: 1 587 794 4232

Email: genevieve.jessiman-perreault@albertahealthservices.ca

Abstract

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

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

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

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

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

(*JMIR Infodemiology* 2024;4:e50551) doi:[10.2196/50551](https://doi.org/10.2196/50551)

KEYWORDS

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

Introduction

Background

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

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

Despite a large body of evidence demonstrating the safety and efficacy of the HPV vaccine [10,11], attitudes toward the vaccine and the accuracy of information shared about this topic in web-based settings vary markedly from extremely negative and erroneous to supportive and factually accurate [12]. In addition, in recent years, there has been a rapid increase in the accessibility of scientific journals and subsequent dissemination of scientific findings through social media [13]. Simultaneously, there has been a decline in the role of unbiased science journalists and other communication experts as mediators between scientists and the public [14]. While these changes have had a democratizing effect on scientific knowledge and allowed for better communication between scientific communities and the public, this unfiltered access to scientific research also creates an environment where individuals may have difficulty in differentiating valid and credible information from biased and unreliable information or may misinterpret legitimate findings [15]. In contrast, researchers have also noted that the growth of open science can create opportunities for people to discuss novel research across polarized boundaries [16], but the type and quality of scientific research about HPV vaccination that is being shared in web-based discussions is

unknown. Finally, with a wealth of open-access scientific research available, there are concerns about how ideologically motivated communities, such as vaccine-hesitant groups, integrate scientific knowledge into their social media communication strategies to amplify uncertainty around vaccines [17]. It is prudent to investigate how scientific research is integrated into web-based HPV vaccine discussions, given that web-based information is typically considered to be more credible, reliable, and authoritative if supported by scientific citation, notwithstanding the source of journal, authorship, or other features [18].

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

Objectives

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

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

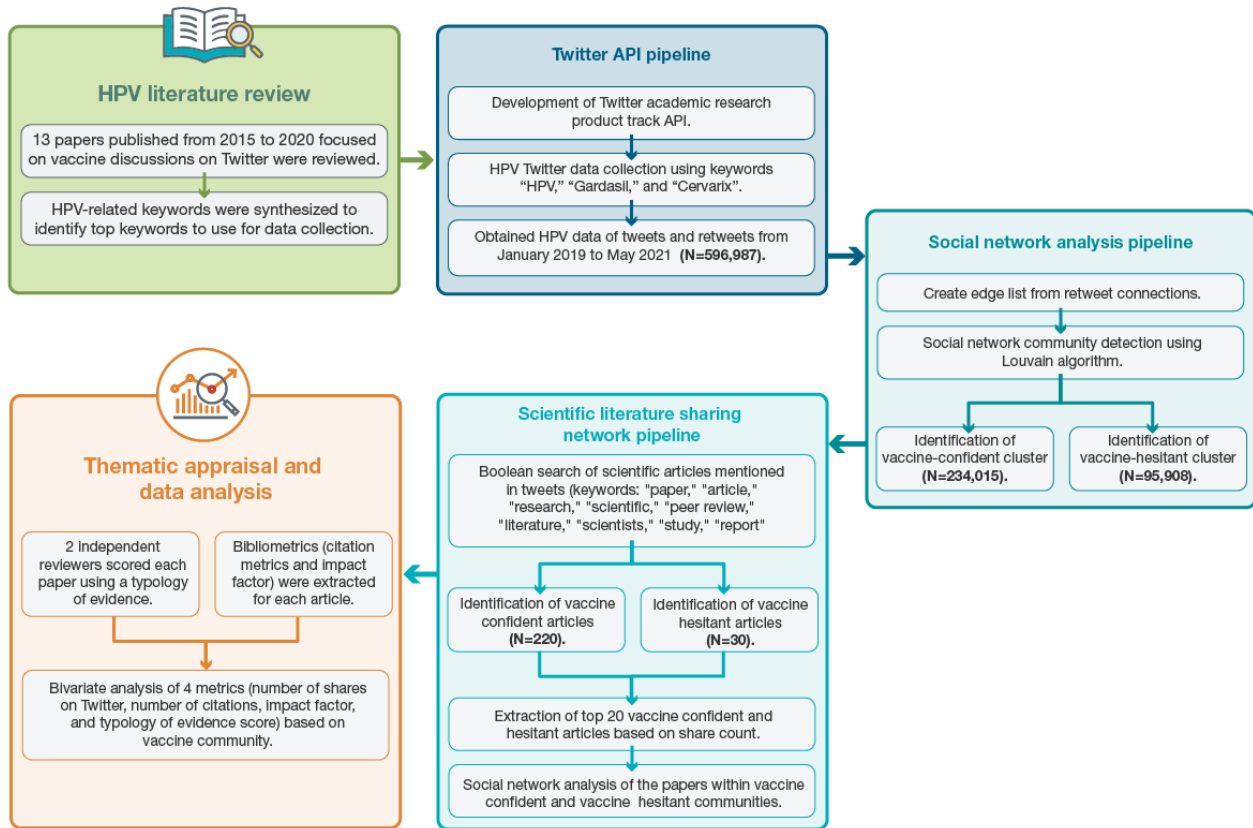
Methods

Overview

Our methods followed a multistep process. First, we conducted a rapid review to inform HPV and HPV vaccine keywords. Second, we used these keywords to filter tweets and create a data set. Third, we detected vaccine-confident and vaccine-hesitant communities and generated social network maps of each community based on tweets and retweet. Fourth, we detected the mentions of scientific literature in each

community and extracted those papers for future statistical and social network analysis. A summary of these methods is presented in [Figure 1](#) (adapted from the paper by Elyashar et al [29]), and further details are presented in the following sections.

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



Literature Review to Inform Data Collection

To determine the most applicable keywords to guide this study, a rapid review was first conducted to determine the most frequently used keywords in literature focused on HPV and HPV immunization discourse on Twitter. The rapid review methodology was selected due to its efficiency in synthesizing a large volume of information in a timely yet systematic manner [30]. This review yielded 13 papers published between 2015

and 2020 about the topic of HPV immunization discussions on social media, with 11 (85%) focusing on HPV immunization discussions on Twitter specifically. We extracted the keywords used in each paper to filter content on social media (Textbox 1). Then, we synthesized these keywords to compile a list of the most used keywords to represent HPV and HPV vaccine discussions on social media, and the top 3 keywords were used to generate the data set.

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

Papers and keywords

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

Data Collection

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

Ethical Considerations

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

Social Network Analysis

First, we created a social network of accounts by creating an edge list using retweets. The retweet edge list consisted of nodes

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

Scientific Literature Sharing Network Analysis

From the vaccine-confident and vaccine-hesitant data sets, we identified tweets that either mentioned or shared scientific literature through a Boolean search for tweets with an embedded http secure link or any of the select list of words (“paper,” “article,” “research,” “scientific,” “peer review,” “literature,” “scientists,” “study,” and “report”) [48]. This filter identified 220 papers from the vaccine-confident community and 30 papers from the vaccine-hesitant community. The titles of or links to these papers were extracted from the data set along with associated metrics such as number of shares for further analysis (as described in the *Data Analysis* section). We identified the top 20 most shared scientific publications in these respective communities. We chose to identify the top 20 most shared scientific publications due to the proportion of shares that these papers had—accounting for >97% of shares in the

vaccine-hesitant community and approximately 61% in the vaccine-confident community. Then, we repeated the social network analysis steps by creating a retweet network of accounts sharing the top 20 prominent scientific publications within the vaccine-confident and vaccine-hesitant communities. The edge list for the vaccine-confident community comprised 989 nodes and 1013 edges, whereas the vaccine-hesitant group had 355 nodes and 422 edges. The primary influencers in this network were again identified using degree centrality measures, and we qualitatively analyzed these accounts on Twitter through their Twitter bio descriptions. The social network analysis of the scientific papers was conducted using Gephi (version 0.9.2) [47].

Typology of Evidence for Thematic and Critical Appraisal

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

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

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

^aRCT: randomized controlled trial.

Bibliometric Indicators

Traditionally, the prestige and quality of a journal was evaluated using citation metrics such as impact factor [91]. In the past few years, as assessment of scientific information has grown exponentially, new tools have been developed to capture the visibility and reach of web-based scientific information. Examples of these alternative metrics or altmetrics include likes, shared tweets, and retweets [92]. To compare traditional scholarly measures of quality to altmetrics, we collected data about the number of times the paper was shared by each vaccine community and the impact factor of the journal the paper was published in. We also collected data about the number of citations each shared paper had received through Google Scholar. Given that citations are impacted by the length of time since publication, we used the SCImago Journal Ranking (SJR) indicator, which provides a weighted average score that remains consistent each year and accounts for the prestige of the citing journal and the differences across subject fields, allowing for more equal comparisons across subject fields [93]. Each paper was assigned an SJR indicator, where a lower score indicates

lower-ranking journals and higher scores indicate higher-ranking journals [94]. Journals that were not indexed in the Scopus database were not assigned an SJR score and were marked as missing in our database. These metrics were used to assess the influence of the shared papers in scientific research and the prestige of the journal the shared papers were published in.

Data Analysis

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

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

Results

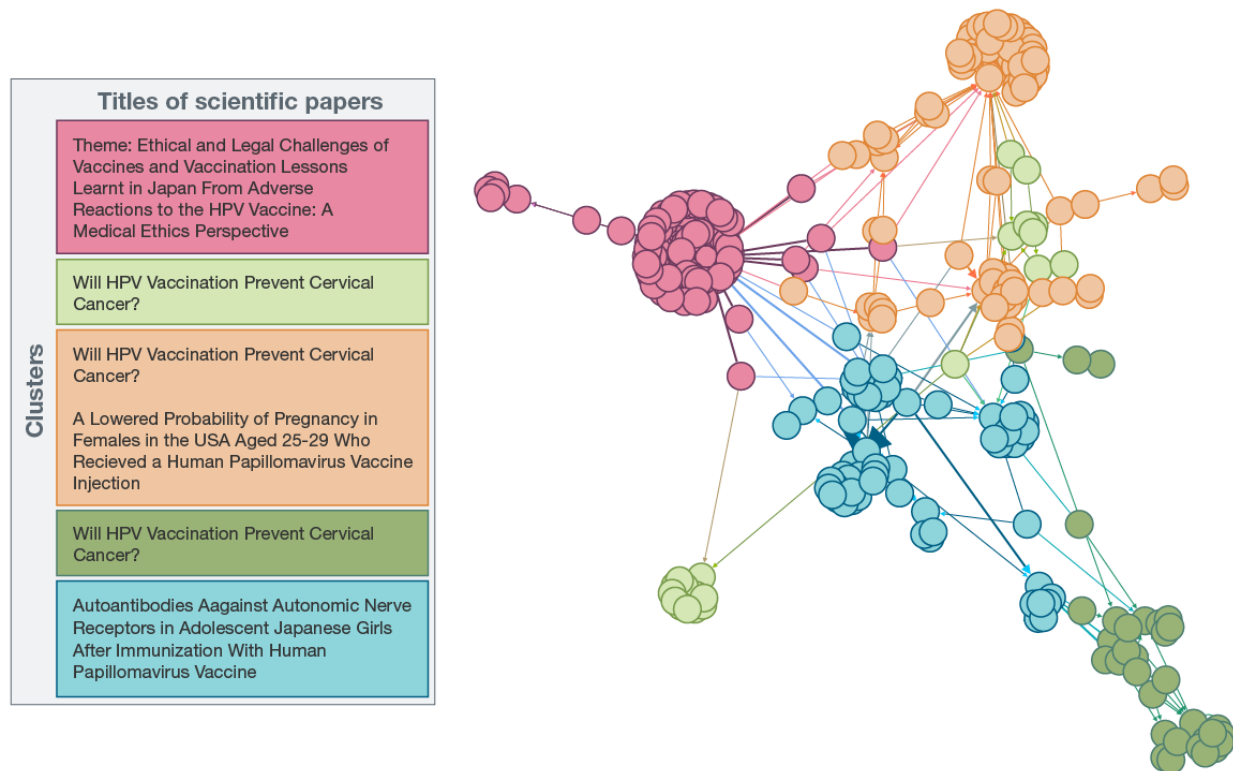
Overview

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

Vaccine-Hesitant Social Network

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

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



The orange cluster of the vaccine-hesitant community circulated a retracted paper, which alleged that HPV vaccines affected the vaccine recipients' fertility and focused on safety [52]. Furthermore, the orange cluster's location in the network (ie, adjacent to the light green cluster) suggests social influence and connection between the 2 clusters. In contrast, there was little interaction between the accounts in the light green cluster and the dark green cluster, suggesting that the influential accounts

in these clusters independently found the same scientific literature and circulated it among a relatively isolated cohort of users. Finally, in the blue cluster, a European support group for those who had experienced vaccine injuries was the leading influential account, whereas a medical society's account that published a widely shared paper in this cluster [51] was an account of secondary influence. Again, the influential accounts in this cluster shared scientific papers, which were retweeted

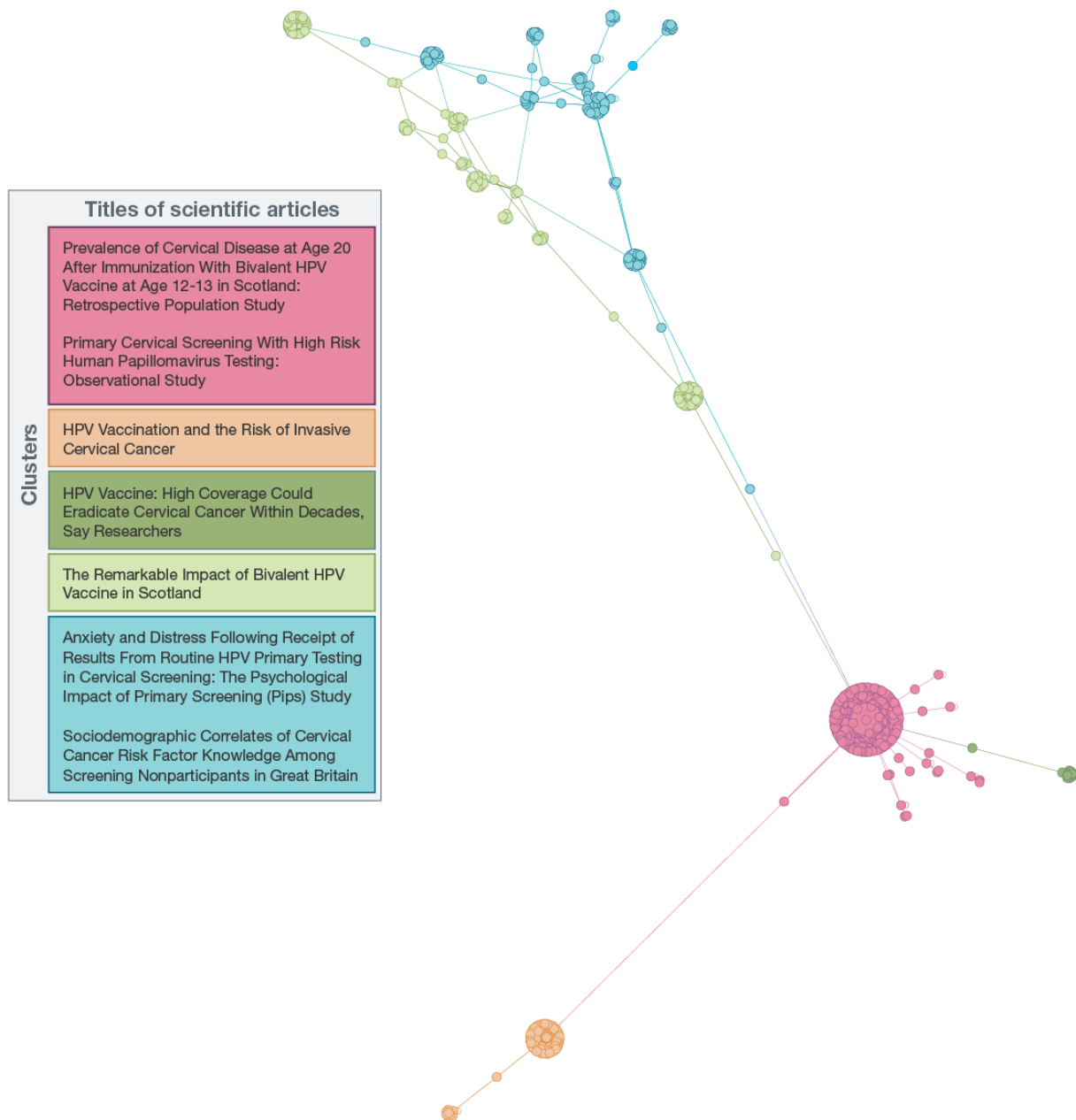
by accounts that are more peripheral to the central clusters of influential accounts. The primary scientific paper circulated among users in this cluster focused on the theme of safety of the HPV vaccine by measuring the serum levels of autoantibodies in a cohort of girls who had possible adverse reactions following the receipt of the HPV vaccine [53].

Vaccine-Confident Social Network

The retweet network of scientific research shared among the web-based vaccine-confident community can similarly be divided into 5 distinct subclusters, as shown in Figure 3. The red cluster primarily included users retweeting literature from

the *British Medical Journal*. There were 2 main papers circulated in this cluster, both of which focused on the effectiveness of the HPV vaccine. The first was a retrospective population study about the efficacy of the HPV vaccine in the prevention of cervical cancer in Scotland, focusing on the theme of satisfaction with service [71], whereas the second was an observational study about the outcomes of HPV screening in high-risk populations in England [72]. In the orange cluster, we observed a similar influence exerted by a government-funded public health agency, which shared a popular paper about effectiveness, focusing on the potential of the HPV vaccine to lower the risk of cervical cancer in a cohort population [73].

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



In the red, orange, and dark green clusters, there were physicians and health care workers among the users who retweeted influential tweets. For example, in the orange cluster, 1 particularly influential physician circulated an editorial paper

about the effectiveness of the HPV vaccine, which indicated that high HPV vaccine coverage could eradicate cervical cancer within a few decades [74]. A science correspondent for a pre-eminent American newspaper was the leading influencer

in the light green cluster wherein the primary paper circulated was an editorial, also focused on effectiveness, related to the positive impacts of HPV vaccination in Scotland [75]. Finally, in the blue cluster, a leading cancer prevention researcher from a British research institute was the leading influencer and author of the scientific papers circulated. In this cluster, papers about the psychological impacts of HPV screening [76] and the sociodemographic correlates of cervical cancer risk among those who did not participate in cervical screening programs in the United Kingdom [77] were recirculated by the accounts influenced by the leading researcher. Unlike the other clusters, health care workers were not overrepresented in the light green and blue clusters.

Overall, results from the vaccine-confident community suggest that health care, scientific, and news media communities are operating in closed systems. As we can see in Figure 3, there are relatively few bridging connections among the different communities discussing influential HPV vaccination literature in the vaccine-confident space. In contrast, the vaccine-hesitant space (Figure 2) is a more cohesive and tightly connected community, suggesting that there are stronger knowledge flows between subclusters in this group. Twitter accounts in the vaccine-hesitant community appear to be more efficient in sharing information than the more fragmented vaccine-confident community (Multimedia Appendix 3). Furthermore, the

vaccine-hesitant Twitter accounts are more effective in communicating the results and research of interest to one another, whereas those in the vaccine-confident space appear to struggle to disseminate the research of interest beyond their personal and professional communities. These findings are supported by the descriptive statistics presented later in the paper, which indicated that while the vaccine-confident community shares far more scientific papers than the vaccine-hesitant community, the scientific literature shared by the vaccine-confident community received far fewer shares per paper despite being published in higher-ranked journals.

Typology of Evidence and Bibliometric Analysis

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

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

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

Table 3 presents the descriptive statistics about the 4 metrics for all the scientific papers shared by vaccine-confident (220/250, 88% papers) and vaccine-hesitant (30/250, 12% papers) communities. The 4 metrics described in Table 3 are the median shares per paper, the median number of citations each shared paper received, the median SJR score of the journal that published each shared paper, and the median typology of evidence score. Table 3 also presents the results from the Mann-Whitney *U* test. Tweets containing scientific papers shared by the vaccine-confident community received a median of 3 shares, compared to a median of 4 shares by the vaccine-hesitant community. Results from the Mann-Whitney *U* test indicate that there are statistically significant differences ($P=.01$) in shares of tweets containing papers about HPV vaccination between the vaccine-hesitant and vaccine-confident communities and that this difference is small (Cohen $d=0.37$). Scientific papers shared by the vaccine-confident community received a median of 13 citations compared to a median of 17

citations for the scientific papers shared by the vaccine-hesitant community. We did not find evidence of statistically significant differences in the number of citations received by papers shared between the vaccine-confident and vaccine-hesitant communities. Scientific papers shared by the vaccine-confident community received a median SJR score of 1.83 compared to a median score of 0.84 for the papers shared by the vaccine-hesitant community. Results from the effect size calculation found this to be a medium standardized difference (Cohen $d=0.61$). The Mann-Whitney *U* test also found evidence of statistically significant ($P<.001$) differences in SJR scores of the HPV-related papers shared between the vaccine-confident and vaccine-hesitant communities. Finally, scientific papers shared by both the vaccine-confident and the vaccine-hesitant communities received a median typology of evidence score of 1, and results from the Mann-Whitney *U* test did not find evidence of a statistically significant difference.

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

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

Discussion

Principal Findings

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

Furthermore, the quality of journals that published the papers shared in these communities varied markedly. The scientific publications shared by the vaccine-confident community were significantly more likely to be published in higher-ranked

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

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

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

in health authorities and reducing the proliferation of conspiratorial beliefs [105].

Limitations

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

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

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

Twitter toward HPV vaccination amidst the context of the pandemic [107].

Strengths

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

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

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

Conclusions

Many of the communication strategies initially used by health promotion communities, including the use of the logical fallacy such as appealing to scientific authority and scientific knowledge, appear to have been co-opted by the vaccine-hesitant community and are being used to create controversy by focusing

on questions about the effectiveness and safety of the HPV vaccine. While the scientific literature shared within these vaccine-hesitant communities is often published in lower-ranked journals, they deliver a substantially more successful, coordinated strategy when it comes to communicating about HPV vaccine on Twitter, compared to the vaccine-confident communities. By widely sharing a curated selection of scientific publications among like-minded individuals, the vaccine-hesitant community members' communication around the HPV vaccine yields much more interaction (ie, shares and retweets) than is observed in the vaccine-confident community's efforts to disseminate research findings. While the scientific literature shared by members of the vaccine-confident community is

published in higher-ranked journals, these papers receive far fewer interactions and have lesser reach on Twitter.

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

Acknowledgments

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

Conflicts of Interest

None declared.

Multimedia Appendix 1

Summary of the top 20 most shared scientific papers on Twitter by the vaccine-hesitant community.

[[DOCX File, 21 KB](#) - [infodemiology_v4i1e50551_app1.docx](#)]

Multimedia Appendix 2

Summary of the top 20 most shared scientific papers on Twitter by the vaccine-confident community.

[[DOCX File, 23 KB](#) - [infodemiology_v4i1e50551_app2.docx](#)]

Multimedia Appendix 3

Retweet network map of human papillomavirus immunization conversations (N=596,987).

[[DOCX File, 457 KB](#) - [infodemiology_v4i1e50551_app3.docx](#)]

References

1. Volesky KD, El-Zein M, Franco EL, Brenner DR, Friedenreich CM, Ruan Y. Corrigendum to "Cancers attributable to infections in Canada" [Prev. Med. 122 (2019) 109-117]. *Prev Med* 2019 Dec;129:105730. [doi: [10.1016/j.ypmed.2019.05.018](#)] [Medline: [31151672](#)]
2. Global strategy to accelerate the elimination of cervical cancer as a public health problem. World Health Organization. 2020 Nov 17. URL: <https://www.who.int/publications/i/item/9789240014107> [accessed 2024-02-05]
3. Hopkins TG, Wood N. Female human papillomavirus (HPV) vaccination: global uptake and the impact of attitudes. *Vaccine* 2013 Mar 25;31(13):1673-1679. [doi: [10.1016/j.vaccine.2013.01.028](#)] [Medline: [23375978](#)]
4. Dubé E, Gagnon D, MacDonald N, Bocquier A, Peretti-Watel P, Verger P. Underlying factors impacting vaccine hesitancy in high income countries: a review of qualitative studies. *Expert Rev Vaccines* 2018 Nov;17(11):989-1004. [doi: [10.1080/14760584.2018.1541406](#)] [Medline: [30359151](#)]
5. Tonsaker T, Bartlett G, Trpkov C. Health information on the internet: gold mine or minefield? *Can Fam Physician* 2014 May;60(5):407-408 [FREE Full text] [Medline: [24828994](#)]
6. Lutkenhaus RO, Jansz J, Bouman MP. Mapping the Dutch vaccination debate on Twitter: identifying communities, narratives, and interactions. *Vaccine X* 2019 Mar 21;1:100019 [FREE Full text] [doi: [10.1016/j.jvacx.2019.100019](#)] [Medline: [31485580](#)]
7. Castells M. *Networks of Outrage and Hope: Social Movements in the Internet Age*, 2nd Edition. Hoboken, NJ: Wiley; Apr 2015.
8. Baldwin AS, Tiro JA, Zimet GD. Broad perspectives in understanding vaccine hesitancy and vaccine confidence: an introduction to the special issue. *J Behav Med* 2023 Apr;46(1-2):1-8 [FREE Full text] [doi: [10.1007/s10865-023-00397-8](#)] [Medline: [36802315](#)]
9. Larson HJ, Gakidou E, Murray CJ. The vaccine-hesitant moment. *N Engl J Med* 2022 Jul 07;387(1):58-65 [FREE Full text] [doi: [10.1056/NEJMr2106441](#)] [Medline: [35767527](#)]

10. Huh WK, Joura EA, Giuliano AR, Iversen OE, de Andrade RP, Ault KA, et al. Final efficacy, immunogenicity, and safety analyses of a nine-valent human papillomavirus vaccine in women aged 16-26 years: a randomised, double-blind trial. *Lancet* 2017 Nov 11;390(10108):2143-2159. [doi: [10.1016/S0140-6736\(17\)31821-4](https://doi.org/10.1016/S0140-6736(17)31821-4)] [Medline: [28886907](https://pubmed.ncbi.nlm.nih.gov/28886907/)]
11. Arbyn M, Xu L. Efficacy and safety of prophylactic HPV vaccines. A Cochrane review of randomized trials. *Expert Rev Vaccines* 2018 Dec;17(12):1085-1091. [doi: [10.1080/14760584.2018.1548282](https://doi.org/10.1080/14760584.2018.1548282)] [Medline: [30495978](https://pubmed.ncbi.nlm.nih.gov/30495978/)]
12. Madden K, Nan X, Briones R, Waks L. Sorting through search results: a content analysis of HPV vaccine information online. *Vaccine* 2012 May 28;30(25):3741-3746. [doi: [10.1016/j.vaccine.2011.10.025](https://doi.org/10.1016/j.vaccine.2011.10.025)] [Medline: [22019758](https://pubmed.ncbi.nlm.nih.gov/22019758/)]
13. Piwowar H, Priem J, Larivière V, Alperin JP, Matthias L, Norlander B, et al. The state of OA: a large-scale analysis of the prevalence and impact of open access articles. *PeerJ* 2018 Feb 13;6:e4375 [FREE Full text] [doi: [10.7717/peerj.4375](https://doi.org/10.7717/peerj.4375)] [Medline: [29456894](https://pubmed.ncbi.nlm.nih.gov/29456894/)]
14. Bucchi M. Credibility, expertise and the challenges of science communication 2.0. *Public Underst Sci* 2017 Nov;26(8):890-893. [doi: [10.1177/0963662517733368](https://doi.org/10.1177/0963662517733368)] [Medline: [29025369](https://pubmed.ncbi.nlm.nih.gov/29025369/)]
15. Steffens MS, Dunn AG, Leask J. Meeting the challenges of reporting on public health in the new media landscape. *Aust J Rev* 2017 Dec;39(2):119-132.
16. Barberá P. How social media reduces mass political polarization. Evidence from Germany, Spain, and the U.S. In: Proceedings of the American Political Science Association Annual Meeting 2015. 2015 Presented at: APSA 2015; September 3-6, 2015; San Francisco, CA URL: http://pablobarbera.com/static/barbera_polarization_APSA.pdf
17. van Schalkwyk F, Dudek J, Costas R. Communities of shared interests and cognitive bridges: the case of the anti-vaccination movement on Twitter. *Scientometrics* 2020 Jun 13;125:1499-1416. [doi: [10.1007/s11192-020-03551-0](https://doi.org/10.1007/s11192-020-03551-0)]
18. Boothby C, Murray D, Waggy AP, Tsou A, Sugimoto CR. Credibility of scientific information on social media: variation by platform, genre and presence of formal credibility cues. *Quant Sci Stud* 2021 Nov 05;2(3):845-863. [doi: [10.1162/qss_a_00151](https://doi.org/10.1162/qss_a_00151)]
19. Bello-Orgaz G, Hernandez-Castro J, Camacho D. Detecting discussion communities on vaccination in Twitter. *Future Gener Comput Syst* 2017 Jan;66:125-136. [doi: [10.1016/j.future.2016.06.032](https://doi.org/10.1016/j.future.2016.06.032)]
20. Chakraborty P, Colditz JB, Silvestre AJ, Friedman MR, Bogen KW, Primack BA. Observation of public sentiment toward human papillomavirus vaccination on Twitter. *Cogent Med* 2017 Dec 12;4(1). [doi: [10.1080/2331205x.2017.1390853](https://doi.org/10.1080/2331205x.2017.1390853)]
21. Du J, Xu J, Song H, Liu X, Tao C. Optimization on machine learning based approaches for sentiment analysis on HPV vaccines related tweets. *J Biomed Semantics* 2017 Mar 03;8(1):9 [FREE Full text] [doi: [10.1186/s13326-017-0120-6](https://doi.org/10.1186/s13326-017-0120-6)] [Medline: [28253919](https://pubmed.ncbi.nlm.nih.gov/28253919/)]
22. Du J, Xu J, Song HY, Tao C. Leveraging machine learning-based approaches to assess human papillomavirus vaccination sentiment trends with Twitter data. *BMC Med Inform Decis Mak* 2017 Jul 05;17(Suppl 2):69 [FREE Full text] [doi: [10.1186/s12911-017-0469-6](https://doi.org/10.1186/s12911-017-0469-6)] [Medline: [28699569](https://pubmed.ncbi.nlm.nih.gov/28699569/)]
23. Kunneman F, Lambooi M, Wong A, van den Bosch A, Mollema L. Monitoring stance towards vaccination in Twitter messages. *BMC Med Inform Decis Mak* 2020 Feb 18;20(1):33 [FREE Full text] [doi: [10.1186/s12911-020-1046-y](https://doi.org/10.1186/s12911-020-1046-y)] [Medline: [32070334](https://pubmed.ncbi.nlm.nih.gov/32070334/)]
24. Tavošchi 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](https://doi.org/10.1080/21645515.2020.1714311)] [Medline: [32118519](https://pubmed.ncbi.nlm.nih.gov/32118519/)]
25. Du J, Luo C, Shegog R, Bian J, Chen Y, Tao C. Deep learning and behavioral theory: an improved analytic method to understand HPV vaccination intentions from Twitter discussion. *SSRN J* 2019 Oct 03;1-30 [FREE Full text] [doi: [10.2139/ssrn.3460688](https://doi.org/10.2139/ssrn.3460688)]
26. Shah Z, Surian D, Dydya 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](https://doi.org/10.2196/14007)] [Medline: [31682571](https://pubmed.ncbi.nlm.nih.gov/31682571/)]
27. Sanawi JB, Samani MC, Taibi M. #vaccination: identifying influencers in the vaccination discussion on Twitter through social network visualisation. *Int J Bus Soc* 2017 Oct;18(S4):718-726.
28. Caulfield T, Marcon AR, Murdoch B, Brown JM, Perrault ST, Jarry J, et al. Health misinformation and the power of narrative messaging in the public sphere. *Can J Bioeth* 2019 Mar 20;2(2):52-60. [doi: [10.7202/1060911ar](https://doi.org/10.7202/1060911ar)]
29. Elyashar A, Plohotnikov I, Cohen IC, Puzis R, Cohen O. The state of mind of health care professionals in light of the COVID-19 pandemic: text analysis study of Twitter discourses. *J Med Internet Res* 2021 Oct 22;23(10):e30217 [FREE Full text] [doi: [10.2196/30217](https://doi.org/10.2196/30217)] [Medline: [34550899](https://pubmed.ncbi.nlm.nih.gov/34550899/)]
30. Dobbins M. Rapid Review Guidebook. Hamilton, ON: National Collaborating Centre for Methods and Tools; 2017.
31. Shapiro GK, Surian D, Dunn AG, Perry R, Kelaher M. Comparing human papillomavirus vaccine concerns on Twitter: a cross-sectional study of users in Australia, Canada and the UK. *BMJ Open* 2017 Oct 05;7(10):e016869 [FREE Full text] [doi: [10.1136/bmjopen-2017-016869](https://doi.org/10.1136/bmjopen-2017-016869)] [Medline: [28982821](https://pubmed.ncbi.nlm.nih.gov/28982821/)]
32. Massey PM, Leader A, Yom-Tov E, Budenz A, Fisher K, Klassen AC. Applying multiple data collection tools to quantify human papillomavirus vaccine communication on Twitter. *J Med Internet Res* 2016 Dec 05;18(12):e318 [FREE Full text] [doi: [10.2196/jmir.6670](https://doi.org/10.2196/jmir.6670)] [Medline: [27919863](https://pubmed.ncbi.nlm.nih.gov/27919863/)]

33. Keim-Malpass J, Mitchell EM, Sun E, Kennedy C. Using Twitter to understand public perceptions regarding the #HPV vaccine: opportunities for public health nurses to engage in social marketing. *Public Health Nurs* 2017 Jul;34(4):316-323. [doi: [10.1111/phn.12318](https://doi.org/10.1111/phn.12318)] [Medline: [28261846](https://pubmed.ncbi.nlm.nih.gov/28261846/)]
34. Nelon JL, Moscarelli M, Stupka P, Sumners C, Uselton T, Patterson MS. Does scientific publication inform public discourse? A case study observing social media engagement around vaccinations. *Health Promot Pract* 2021 May;22(3):377-384. [doi: [10.1177/1524839919899925](https://doi.org/10.1177/1524839919899925)] [Medline: [32009444](https://pubmed.ncbi.nlm.nih.gov/32009444/)]
35. Surian D, Nguyen DQ, Kennedy G, Johnson M, Coiera E, Dunn AG. Characterizing Twitter discussions about HPV vaccines using topic modeling and community detection. *J Med Internet Res* 2016 Aug 29;18(8):e232 [FREE Full text] [doi: [10.2196/jmir.6045](https://doi.org/10.2196/jmir.6045)] [Medline: [27573910](https://pubmed.ncbi.nlm.nih.gov/27573910/)]
36. Zhou X, Coiera E, Tsafnat G, Arachi D, Ong MS, Dunn AG. Using social connection information to improve opinion mining: identifying negative sentiment about HPV vaccines on Twitter. *Stud Health Technol Inform* 2015;216:761-765. [Medline: [26262154](https://pubmed.ncbi.nlm.nih.gov/26262154/)]
37. Becker BF, Larson HJ, Bonhoeffer J, van Mulligen EM, Kors JA, Sturkenboom MC. Evaluation of a multinational, multilingual vaccine debate on Twitter. *Vaccine* 2016 Dec 07;34(50):6166-6171. [doi: [10.1016/j.vaccine.2016.11.007](https://doi.org/10.1016/j.vaccine.2016.11.007)] [Medline: [27840012](https://pubmed.ncbi.nlm.nih.gov/27840012/)]
38. Dyda A, Shah Z, Surian D, Martin P, Coiera E, Dey A, et al. HPV vaccine coverage in Australia and associations with HPV vaccine information exposure among Australian Twitter users. *Hum Vaccin Immunother* 2019;15(7-8):1488-1495. [doi: [10.1080/21645515.2019.1596712](https://doi.org/10.1080/21645515.2019.1596712)] [Medline: [30978147](https://pubmed.ncbi.nlm.nih.gov/30978147/)]
39. Dunn AG, Surian D, Leask J, Dey A, Mandl KD, Coiera E. Mapping information exposure on social media to explain differences in HPV vaccine coverage in the United States. *Vaccine* 2017 May 25;35(23):3033-3040 [FREE Full text] [doi: [10.1016/j.vaccine.2017.04.060](https://doi.org/10.1016/j.vaccine.2017.04.060)] [Medline: [28461067](https://pubmed.ncbi.nlm.nih.gov/28461067/)]
40. Budenz A, Massey P, Leader A, Klassen A, Fisher K, Yom-Tov E. Let's talk about HPV: a focus on gender- and MSM-specific Twitter communication. In: *Proceedings of the American Public Health Association Annual Meeting 2017*. 2017 Presented at: APHA 2017; November 4-8, 2017; Atlanta, GA.
41. Zhang H, Wheldon C, Dunn AG, Tao C, Huo J, Zhang R, et al. Mining Twitter to assess the determinants of health behavior toward human papillomavirus vaccination in the United States. *J Am Med Inform Assoc* 2020 Feb 01;27(2):225-235 [FREE Full text] [doi: [10.1093/jamia/ocz191](https://doi.org/10.1093/jamia/ocz191)] [Medline: [31711186](https://pubmed.ncbi.nlm.nih.gov/31711186/)]
42. X data for academic research. X Developer Platform. URL: <https://developer.twitter.com/en/use-cases/do-research/academic-research> [accessed 2024-02-05]
43. X API documentation. X Developer Platform. URL: <https://developer.twitter.com/en/docs/twitter-api> [accessed 2024-02-05]
44. Van Rossum G, Drake FL. *Python 3 Reference Manual*. North Charleston, SC: CreateSpace; Mar 2009.
45. Yan Y, Toriumi F, Sugawara T. Understanding how retweets influence the behaviors of social networking service users via agent-based simulation. *Comput Soc Netw* 2021 Sep 13;8:18. [doi: [10.1186/S40649-021-00099-8](https://doi.org/10.1186/S40649-021-00099-8)]
46. Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. *J Stat Mech* 2008 Oct 09;2008:P10008. [doi: [10.1088/1742-5468/2008/10/P10008](https://doi.org/10.1088/1742-5468/2008/10/P10008)]
47. Bastian M, Heymann S, Jacomy M. Gephi: an open source software for exploring and manipulating networks. *Proc Int AAAI Conf Web Soc Media* 2009 Mar 19;3(1):361-362. [doi: [10.1609/icwsm.v3i1.13937](https://doi.org/10.1609/icwsm.v3i1.13937)]
48. pandas.Series.str.contains. Pandas. URL: <https://pandas.pydata.org/docs/reference/api/pandas.Series.str.contains.html> [accessed 2024-02-05]
49. Gray JA. *Evidence-Based Healthcare*. London, UK: Churchill Livingstone; 1997.
50. Beppu H, Minaguchi M, Uchide K, Kumamoto K, Sekiguchi M, Yaju Y. Lessons learnt in Japan from adverse reactions to the HPV vaccine: a medical ethics perspective. *Indian J Med Ethics* 2017 Apr 1;82 [FREE Full text] [doi: [10.20529/IJME.2017.021](https://doi.org/10.20529/IJME.2017.021)]
51. Rees CP, Brhlikova P, Pollock AM. Will HPV vaccination prevent cervical cancer? *J R Soc Med* 2020 Feb 21;113(2):64-78 [FREE Full text] [doi: [10.1177/0141076819899308](https://doi.org/10.1177/0141076819899308)] [Medline: [31962050](https://pubmed.ncbi.nlm.nih.gov/31962050/)]
52. DeLong G. A lowered probability of pregnancy in females in the USA aged 25-29 who received a human papillomavirus vaccine injection. Retracted in: *J Clin Hypertens (Greenwich)* 2018 Jun 11;81(14):661-674. [doi: [10.1080/15287394.2018.1477640](https://doi.org/10.1080/15287394.2018.1477640)] [Medline: [29889622](https://pubmed.ncbi.nlm.nih.gov/29889622/)]
53. Hineno A, Ikeda S, Scheibenbogen C, Heidecke H, Forster K, Junker J, et al. Autoantibodies against autonomic nerve receptors in adolescent Japanese girls after immunization with human papillomavirus vaccine. *Ann Arthritis Clin Rheumatol* 2019;2:1014 [FREE Full text]
54. Atkinson AE, Mandujano CA, Bejarano S, Kennedy LS, Tsongalis GJ. Screening for human papillomavirus in a low- and middle-income country. *J Glob Oncol* 2019 May;5:JGO1800233 [FREE Full text] [doi: [10.1200/JGO.18.00233](https://doi.org/10.1200/JGO.18.00233)] [Medline: [31050922](https://pubmed.ncbi.nlm.nih.gov/31050922/)]
55. Bizjak M, Bruck O, Praprotnik S, Dahan S, Shoenfeld Y. Pancreatitis after human papillomavirus vaccination: a matter of molecular mimicry. *Immunol Res* 2017 Mar;65(1):164-167. [doi: [10.1007/s12026-016-8823-9](https://doi.org/10.1007/s12026-016-8823-9)] [Medline: [27421720](https://pubmed.ncbi.nlm.nih.gov/27421720/)]
56. Blitshteyn S, Brinth L, Hendrickson JE, Martinez-Lavin M. Autonomic dysfunction and HPV immunization: an overview. *Immunol Res* 2018 Dec;66(6):744-754. [doi: [10.1007/s12026-018-9036-1](https://doi.org/10.1007/s12026-018-9036-1)] [Medline: [30478703](https://pubmed.ncbi.nlm.nih.gov/30478703/)]

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

78. Artemchuk H, Eriksson T, Poljak M, Surcel HM, Dillner J, Lehtinen M, et al. Long-term antibody response to human papillomavirus vaccines: up to 12 years of follow-up in the Finnish maternity cohort. *J Infect Dis* 2019 Jan 29;219(4):582-589. [doi: [10.1093/infdis/jiy545](https://doi.org/10.1093/infdis/jiy545)] [Medline: [30239832](https://pubmed.ncbi.nlm.nih.gov/30239832/)]
79. Baker P, Kelly D, Medeiros R, Morrissey M, Price R. Eliminating HPV-caused cancers in Europe: achieving the possible. *J Cancer Policy* 2021 Jun;28:100280. [doi: [10.1016/j.jcpo.2021.100280](https://doi.org/10.1016/j.jcpo.2021.100280)] [Medline: [35559909](https://pubmed.ncbi.nlm.nih.gov/35559909/)]
80. Colafrancesco S, Perricone C, Tomljenovic L, Shoenfeld Y. Human papilloma virus vaccine and primary ovarian failure: another facet of the autoimmune/inflammatory syndrome induced by adjuvants. *Am J Reprod Immunol* 2013 Oct;70(4):309-316. [doi: [10.1111/aji.12151](https://doi.org/10.1111/aji.12151)] [Medline: [23902317](https://pubmed.ncbi.nlm.nih.gov/23902317/)]
81. Cromwell I, Smith LW, van der Hoek K, Hedden L, Coldman AJ, Cook D, et al. Cost-effectiveness analysis of primary human papillomavirus testing in cervical cancer screening: results from the HPV FOCAL Trial. *Cancer Med* 2021 May;10(9):2996-3003 [FREE Full text] [doi: [10.1002/cam4.3864](https://doi.org/10.1002/cam4.3864)] [Medline: [33811457](https://pubmed.ncbi.nlm.nih.gov/33811457/)]
82. Drolet M, Bénard É, Pérez N, Brisson M, HPV Vaccination Impact Study Group. Population-level impact and herd effects following the introduction of human papillomavirus vaccination programmes: updated systematic review and meta-analysis. *Lancet* 2019 Aug 10;394(10197):497-509 [FREE Full text] [doi: [10.1016/S0140-6736\(19\)30298-3](https://doi.org/10.1016/S0140-6736(19)30298-3)] [Medline: [31255301](https://pubmed.ncbi.nlm.nih.gov/31255301/)]
83. Gee J, Naleway A, Shui I, Baggs J, Yin R, Li R, et al. Monitoring the safety of quadrivalent human papillomavirus vaccine: findings from the Vaccine Safety Datalink. *Vaccine* 2011 Oct 26;29(46):8279-8284. [doi: [10.1016/j.vaccine.2011.08.106](https://doi.org/10.1016/j.vaccine.2011.08.106)] [Medline: [21907257](https://pubmed.ncbi.nlm.nih.gov/21907257/)]
84. Gleber-Netto FO, Rao X, Guo T, Xi Y, Gao M, Shen L, et al. Variations in HPV function are associated with survival in squamous cell carcinoma. *JCI Insight* 2019 Jan 10;4(1):e124762 [FREE Full text] [doi: [10.1172/jci.insight.124762](https://doi.org/10.1172/jci.insight.124762)] [Medline: [30626753](https://pubmed.ncbi.nlm.nih.gov/30626753/)]
85. Kardas-Nelson M. Vaccine uptake and prevalence of HPV related cancers in US men. *BMJ* 2019 Mar 18;364:11210. [doi: [10.1136/bmj.11210](https://doi.org/10.1136/bmj.11210)] [Medline: [30885904](https://pubmed.ncbi.nlm.nih.gov/30885904/)]
86. Klein NP, Hansen J, Chao C, Velicer C, Emery M, Slezak J, et al. Safety of quadrivalent human papillomavirus vaccine administered routinely to females. *Arch Pediatr Adolesc Med* 2012 Dec;166(12):1140-1148. [doi: [10.1001/archpediatrics.2012.1451](https://doi.org/10.1001/archpediatrics.2012.1451)] [Medline: [23027469](https://pubmed.ncbi.nlm.nih.gov/23027469/)]
87. Nagarsheth NB, Norberg SM, Sinkoe AL, Adhikary S, Meyer TJ, Lack JB, et al. TCR-engineered T cells targeting E7 for patients with metastatic HPV-associated epithelial cancers. *Nat Med* 2021 Mar;27(3):419-425 [FREE Full text] [doi: [10.1038/s41591-020-01225-1](https://doi.org/10.1038/s41591-020-01225-1)] [Medline: [33558725](https://pubmed.ncbi.nlm.nih.gov/33558725/)]
88. Simms KT, Hanley SJ, Smith MA, Keane A, Canfell K. Impact of HPV vaccine hesitancy on cervical cancer in Japan: a modelling study. *Lancet Public Health* 2020 Apr;5(4):e223-e234 [FREE Full text] [doi: [10.1016/S2468-2667\(20\)30010-4](https://doi.org/10.1016/S2468-2667(20)30010-4)] [Medline: [32057317](https://pubmed.ncbi.nlm.nih.gov/32057317/)]
89. Hoss A, Meyerson BE, Zimet GD. State statutes and regulations related to human papillomavirus vaccination. *Hum Vaccin Immunother* 2019;15(7-8):1519-1526 [FREE Full text] [doi: [10.1080/21645515.2019.1627817](https://doi.org/10.1080/21645515.2019.1627817)] [Medline: [31241406](https://pubmed.ncbi.nlm.nih.gov/31241406/)]
90. Petticrew M, Roberts H. Evidence, hierarchies, and typologies: horses for courses. *J Epidemiol Community Health* 2003 Jul;57(7):527-529 [FREE Full text] [doi: [10.1136/jech.57.7.527](https://doi.org/10.1136/jech.57.7.527)] [Medline: [12821702](https://pubmed.ncbi.nlm.nih.gov/12821702/)]
91. Callaham M, Wears RL, Weber E. Journal prestige, publication bias, and other characteristics associated with citation of published studies in peer-reviewed journals. *JAMA* 2002 Jun 05;287(21):2847-2850. [doi: [10.1001/jama.287.21.2847](https://doi.org/10.1001/jama.287.21.2847)] [Medline: [12038930](https://pubmed.ncbi.nlm.nih.gov/12038930/)]
92. Scarlat MM, Mavrogenis AF, Pečina M, Niculescu M. Impact and alternative metrics for medical publishing: our experience with International Orthopaedics. *Int Orthop* 2015 Aug;39(8):1459-1464. [doi: [10.1007/s00264-015-2766-y](https://doi.org/10.1007/s00264-015-2766-y)] [Medline: [25947897](https://pubmed.ncbi.nlm.nih.gov/25947897/)]
93. Guerrero-Bote VP, Moya-Anegón F. A further step forward in measuring journals' scientific prestige: the SJR2 indicator. *J Informetr* 2012 Oct;6(4):674-688. [doi: [10.1016/j.joi.2012.07.001](https://doi.org/10.1016/j.joi.2012.07.001)]
94. Scimago journal and country rank homepage. Scimago Journal & Country Rank. URL: <https://www.scimagojr.com/> [accessed 2023-07-04]
95. McKnight PE, Najab J. Mann-Whitney U Test. In: The Corsini Encyclopedia of Psychology. Hoboken, NJ: John Wiley & Sons; 2010.
96. Cohen J. Statistical Power Analysis for the Behavioral Sciences Second Edition. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
97. Kata A. Anti-vaccine activists, Web 2.0, and the postmodern paradigm--an overview of tactics and tropes used online by the anti-vaccination movement. *Vaccine* 2012 May 28;30(25):3778-3789. [doi: [10.1016/j.vaccine.2011.11.112](https://doi.org/10.1016/j.vaccine.2011.11.112)] [Medline: [22172504](https://pubmed.ncbi.nlm.nih.gov/22172504/)]
98. Martini C. Ad hominem arguments, rhetoric, and science communication. *Stud Log Gramm Rhetor* 2018;55(1):66. [doi: [10.2478/slgr-2018-0033](https://doi.org/10.2478/slgr-2018-0033)]
99. Vivion M, Anassour Laouan Sidi E, Betsch C, Dionne M, Dubé E, Driedger SM, et al. Prebunking messaging to inoculate against COVID-19 vaccine misinformation: an effective strategy for public health. *J Commun Healthc* 2022 Mar 04;15(3):232-242. [doi: [10.1080/17538068.2022.2044606](https://doi.org/10.1080/17538068.2022.2044606)]
100. van der Linden S, Roozenbeek J, Maertens R, Basol M, Kácha O, Rathje S, et al. How can psychological science help counter the spread of fake news? *Span J Psychol* 2021 Apr 12;24:e25. [doi: [10.1017/sjp.2021.23](https://doi.org/10.1017/sjp.2021.23)]

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

Abbreviations

API: application programming interface

HPV: human papillomavirus

SJR: SCImago Journal Ranking

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

Please cite as:

Jessiman-Perreault G, Boucher JC, Kim SY, Frenette N, Badami A, Smith HM, Allen Scott LK

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

JMIR Infodemiology 2024;4:e50551

URL: <https://infodemiology.jmir.org/2024/1/e50551>

doi: [10.2196/50551](https://doi.org/10.2196/50551)

PMID: [38722678](https://pubmed.ncbi.nlm.nih.gov/38722678/)

©Geneviève Jessiman-Perreault, Jean-Christophe Boucher, So Youn Kim, Nicole Frenette, Abbas Badami, Henry M Smith, Lisa K Allen Scott. Originally published in *JMIR Infodemiology* (<https://infodemiology.jmir.org>), 09.05.2024. 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.

Original Paper

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

Nessie Felicia Frennesson¹, MSc; Julie Barnett², PhD; Youssouf Merouani³, MSc; Angela Attwood¹, PhD; Luisa Zuccolo⁴, PhD; Cheryl McQuire⁵, PhD

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

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

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

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

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

Corresponding Author:

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

Abstract

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

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

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

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

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

(*JMIR Infodemiology* 2024;4:e58056) doi:[10.2196/58056](https://doi.org/10.2196/58056)

KEYWORDS

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

Introduction

Background

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

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

Kingdom, evaluated both knowledge and attitudes toward alcohol in pregnancy before and after their guidelines changed to advising women not to drink during pregnancy and showed no changes in either knowledge or attitudes [23].

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

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

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

thread starts, this study will explore what topics related to alcohol during pregnancy are discussed and if their nature has changed since the start of Mumsnet in 2002, with the change of CMO guidelines in 2016 and the COVID-19 pandemic as key time points for comparison.

Aims and Objectives

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

The specific research questions in this study were as follows:

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

Methods

Data Source

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

Ethical Considerations

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

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

Search Strategy

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

Data Analysis

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

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

Results

Overview

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

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

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

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

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

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

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

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

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

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

Consumed Alcohol Before Knowing About Pregnancy

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

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

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

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

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

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

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

Research, Guidelines, and Official Information About Alcohol in Pregnancy

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

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

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

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

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

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

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

NoLo Products

This category covered conversations about the consumption of NoLo products during pregnancy. Reasons for seeking NoLo alternatives included not feeling left out in a group that is consuming alcohol and also because some missed the taste of

alcohol. The need for this type of product appeared greater during celebrations such as Christmas and weddings:

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

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

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

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

How to Hide Not Consuming Alcohol to Conceal Pregnancy?

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

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

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

Have Been Consuming Alcohol During Pregnancy But Now Worried

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

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

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

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

Overall, 4.3% (17/395) of all the included thread starts appeared in this category. For this category, the proportion of thread starts for the first period was 6.6% (14/213), the second period had a

percentage of 1.9% (2/103), and the last period had a proportion of 1.3% (1/79), falling in this category.

Are Consuming Alcohol During Pregnancy, Not Worried About PAE

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

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

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

Consumed Alcohol by Mistake

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

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

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

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

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

It Is Hard Not to Consume Alcohol During Pregnancy

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

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

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

Discussion

Principal Findings

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

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

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

It was evident that there was a concern among the thread starters about having consumed alcohol before they found out that they were pregnant and that this caused stress and anxiety for some thread starters. There is no known safe limit for alcohol consumption during pregnancy. Some studies have shown that there was no relationship between consuming alcohol during the early days of pregnancy and outcomes such as low birth weight and spontaneous preterm birth [46], while others have suggested that alcohol consumption during the first trimester of pregnancy can increase the risk of spontaneous abortion [47,48]. Many of the thread starters had discussed the consumption of alcohol with their midwives or other professionals, but the concern remained. Previous research has shown that there is a lack of a standardized approach to how midwives approach the topic of alcohol consumption during pregnancy [16]. This could indicate that there is a need for professionals to give accurate information and at the same time being able to reduce any anxiety and stress that alcohol consumption could have caused. It has been suggested that midwives should be offered training in communication skills and in delivering alcohol interventions [16]. Reducing stress during pregnancy is especially important as it has been reported that stress can lead to outcomes such as low birth weight [49] and obesity in the offspring [50]. Research has shown that there

exists a social pressure to consume alcohol, which causes a challenge when someone wishes to conceal their pregnancy [51]. This was also prevalent in the discussions on Mumsnet, where thread starters wanted advice on hiding that they were not consuming alcohol due to their pregnancy.

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

Moreover, many of the thread starts in the category "Asking for advice on whether it is safe to consume alcohol or on safe limits" during the first period mentioned the timing of the pregnancy, which could have been a result of the change in guidelines. The change in guidelines was that the previous National Institute for Health and Care Excellence guidelines suggested that pregnant women should avoid alcohol in the first 3 months of pregnancy, and if they chose to drink, they should not consume more than 1 to 2 units twice per week [16]. However, a 2020 survey by the National Organisation for FASD showed that awareness of the current CMOs guidance that the safest approach is not to consume alcohol at all if you are pregnant or if you could become pregnant remains low among

some population subgroups, particularly young people (aged 18-25 years) [53]. This shows how important it is to communicate research in a way that is acceptable, understandable, and accessible for all. Our study showed that some Mumsnet users missed drinking alcohol while they were pregnant and wanted to know if others agreed that a glass of alcohol was acceptable, particularly on special occasions such as weddings or birthday celebrations. It was also evident that some of the forum users were trying to minimize the potential risks of consuming alcohol by referring to how previous generations had been consuming alcohol without clear adverse outcomes or how other countries have less strict guidelines. This was most apparent in content posted before the introduction of updated CMO low-risk drinking guidance in 2016.

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

Strengths and Limitations

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

CMO guidance and can offer valuable priorities to inform improved communication, reach, and preference for prenatal health information.

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

Conclusions

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

Acknowledgments

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

Data Availability

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

Authors' Contributions

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

Conflicts of Interest

None declared.

Multimedia Appendix 1

Included categories.

[[DOCX File, 24 KB - infodemiology_v4i1e58056_app1.docx](#)]

References

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

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

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

Abbreviations

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

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

Please cite as:

Frennesson NF, Barnett J, Merouani Y, Attwood A, Zuccolo L, McQuire C

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

JMIR Infodemiology 2024;4:e58056

URL: <https://infodemiology.jmir.org/2024/1/e58056>

doi: [10.2196/58056](https://doi.org/10.2196/58056)

PMID:

©Nessie Felicia Frennesson, Julie Barnett, Youssef Merouani, Angela Attwood, Luisa Zuccolo, Cheryl McQuire. Originally published in JMIR Infodemiology (<https://infodemiology.jmir.org>), 20.06.2024. 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.

Original Paper

The Journey of Engaging With Web-Based Self-Harm and Suicide Content: Longitudinal Qualitative Study

Zoë Haime¹, BSc, MSc, PhD; Laura Kennedy^{2,3,4}, BA, MPhil, PhD; Lydia Grace³, BSc, MSc, PhD; Rachel Cohen^{1,5}, BA, MA, PhD; Jane Derges¹, MSc, PhD; Lucy Biddle¹, BA, MA, PhD

¹Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom

²Centre for Society and Mental Health, Kings College, London, United Kingdom

³Samaritans, Surrey, United Kingdom

⁴Department of Health Service and Population Research, King's College London, London, United Kingdom

⁵Health and Social Services Group, Welsh Government, Cardiff, United Kingdom

Corresponding Author:

Zoë Haime, BSc, MSc, PhD

Population Health Sciences

Bristol Medical School

University of Bristol

Canyng Hall

39 Whatley Road

Bristol, BS8 2PS

United Kingdom

Phone: 44 01179289000

Email: zoe.haime@bristol.ac.uk

Abstract

Background: Self-harm and suicide are major public health concerns worldwide, with attention focused on the web environment as a helpful or harmful influence. Longitudinal research on self-harm and suicide-related internet use is limited, highlighting a paucity of evidence on long-term patterns and effects of engaging with such content.

Objective: This study explores the experiences of people engaging with self-harm or suicide content over a 6-month period.

Methods: This study used qualitative and digital ethnographic methods longitudinally, including one-to-one interviews at 3 time points to explore individual narratives. A trajectory analysis approach involving 4 steps was used to interpret the data.

Results: The findings from 14 participants established the web-based journey of people who engage with self-harm or suicide content. In total, 5 themes were identified: initial interactions with self-harm or suicide content, changes in what self-harm or suicide content people engage with and where, changes in experiences of self-harm or suicide behaviors associated with web-based self-harm or suicide content engagement, the disengagement-reengagement cycle, and future perspectives on web-based self-harm or suicide content engagement. Initial engagements were driven by participants seeking help, often when offline support had been unavailable. Some participants' exposure to self-harm and suicide content led to their own self-harm and suicide behaviors, with varying patterns of change over time. Notably, disengagement from web-based self-harm and suicide spaces served as a protective measure for all participants, but the pull of familiar content resulted in only brief periods of disconnection. Participants also expressed future intentions to continue returning to these self-harm and suicide web-based spaces, acknowledging the nonlinear nature of their own recovery journey and aiming to support others in the community. Within the themes identified in this study, narratives revealed that participants' behavior was shaped by cognitive flexibility and rigidity, metacognitive abilities, and digital expertise. Opportunities for behavior change arose during periods of cognitive flexibility prompted by life events, stressors, and shifts in mental health. Participants sought diverse and potentially harmful content during challenging times but moved toward recovery-oriented engagements in positive circumstances. Metacognitive and digital efficacy skills also played a pivotal role in participants' control of web-based interactions, enabling more effective management of content or platforms or sites that posed potential harms.

Conclusions: This study demonstrated the complexity of web-based interactions, with beneficial and harmful content intertwined. Participants who demonstrated metacognition and digital efficacy had better control over web-based engagements. Some attributed these skills to study processes, including taking part in reflective diaries, showing the potential of upskilling users. This study

also highlighted how participants remained vulnerable by engaging with familiar web-based spaces, emphasizing the responsibility of web-based industry leaders to develop tools that empower users to enhance their web-based safety.

(*JMIR Infodemiology* 2024;4:e47699) doi:[10.2196/47699](https://doi.org/10.2196/47699)

KEYWORDS

suicide; self-harm; online; longitudinal; qualitative

Introduction

Background

Self-harm and suicide are major global public health concerns, with >700,000 people worldwide dying by suicide each year [1]. Attention has increasingly focused on the role of the web environment in triggering, exacerbating, or normalizing self-harm and suicide [2-4]. The amount of suicide-related information accessible on the web has grown [5], and graphic content depicting self-harm is increasingly available on social networking platforms [6]. Research shows that self-harm and suicide-related internet use is common among young people [7], particularly those who are under psychiatric care [8] and who go on to die by suicide [9].

There is a range of self-focused and social motivations for engaging with web-based self-harm and suicide content. These include accessing ongoing peer support or immediate help during a crisis [2,10,11], documenting recovery from self-harm [1,8], and researching suicide methods [12]. Moreover, research has shown that the ways in which people interact with web-based self-harm and suicide content vary depending on their level of distress [11,13].

The diversity in self-harm and suicide material complicates the experiences of content engagement. Research has identified these content interactions as being both a public health concern and a possible preventative measure [3,14], and studies have recognized the potential for engagement to have both benefits and costs [15]. Content with the potential to harm includes information on high-lethality suicide methods [16], pro-suicide websites that may encourage suicide [13], and content describing novel methods of self-harm [17]. Benefits associated with accessing content include the role of the online community in peer support, validation and acceptance of one's own self-harm or suicide feelings, and the opportunity for altruism when helping others [2,10,18-20]. These benefits may be particularly valuable given existing gaps in mental health care services and the widespread stigma that people who self-harm or experience suicidal thoughts encounter offline [2]. However, a recent review suggested that the impact of engaging with particular types of web-based self-harm or suicide-related content varies both between and within individuals, with content that benefits some having negative consequences for others and vice versa [15]. The review also identified only 4 longitudinal studies on the impact of self-harm and suicide-related internet use. Of these studies, 2 identified preventative effects of suicide prevention websites and web-based health forums on suicidal ideation [21,22]. One study showed minimal effects of search engine helpline notices on future suicide queries [23], and another study found that exposure to self-harm on Instagram predicted suicidal ideation and self-harm-related outcomes [17]. However, none

of these studies used qualitative methods with their participants, emphasizing the current paucity of evidence on how self-harm and suicide-related web-based behavior evolves and the long-term effects and experiences of engaging with such content from the user's perspective, including whether these are brief or permanent.

Objectives

The aim of this study was to explore the motivations for and consequences of viewing, searching for, and posting web-based self-harm or suicide content over a longitudinal period. Specifically, this study builds on existing knowledge by using qualitative and digital ethnographic methods to explore individual narratives of web-based engagement. Exploration of "significant moments" and points of transition within the web journey could also have substantial implications for the prevention of suicide and reduction of self-harm [24].

Methods

Design

This was a 6-month qualitative ethnographic study that investigated the stability and change in engagement with web-based self-harm and suicide content. This involved 3 one-to-one interviews and daily diary completion by participants over the study duration. We selected a 6-month time frame to ensure that we could observe changes over time in web-based engagement and associated behaviors [25,26] while also remaining mindful of the considerable commitment required for this ethnographic approach to maintain retention of participants.

Ethical Considerations

Ethics approval was obtained from the University of Bristol Faculty of Health Sciences ethics committee (reference: 117491). All participants provided written informed consent before participation, and were informed that they could withdraw from the study (including data withdrawal up to the time of analysis), without giving a reason. During consent, participants were assigned a participant ID used to identify their data and ensure anonymity. They were also informed that their data would be held confidentially and securely by the University of Bristol according to its duties and obligations under GDPR and the Data Protection Act. All participants were also compensated for their time, receiving a total of up to £75 (US \$94.79) for full study completion.

Sampling and Recruitment

UK residents aged ≥16 years who were able to communicate in English and had experience engaging with web-based self-harm or suicide content were eligible. This included posting

images, videos, memes, forum posts, blog posts, recovery posts, or comments related to self-harm or suicide or engaging with others' self-harm or suicide-related content through reposting and reblogging, quoting, liking, sharing, saving, subscribing to, or commenting. They did not need to have previous experience with self-harm or suicidal thoughts or behaviors.

Potential participants responded to advertisements posted between November 2021 and April 2022 on social media platforms (Facebook, Twitter [subsequently rebranded as X], and Reddit subreddits ["r/AdultSelfHarm," "r/StopSelfHarm," "r/BPD," "r/MentalHealthUK," and "r/maementalhealth"]), via Tellmi—a UK-based young person mental health app), and through charity websites and newsletters (Samaritans, SMARtEN, The McPin Foundation, and MQ Mental Health Research). Advertisements were posted once to platforms or sites until the end of recruitment in April 2022; however, due to web-based posting and reposting, it is possible that they were also shared elsewhere by others. Permission was sought from moderators or administrators before posting. Advertisements included a link to an expression of interest form in which participants consented via completion to the collection of brief demographic information, if and when the person last self-harmed, the way they were engaging with web-based self-harm or suicide content, and what platforms they used. All respondents had engaged with web-based self-harm or suicide content in some way.

This information was used to sample a diverse range of participants from those who expressed interest and target recruitment advertisements. Potential participants were sent the study information sheet via email, and those who were still interested in participating completed a consent form. Interviews were then arranged via email. The demographic data of those who did not participate were deleted. Once 14 baseline interviews had been conducted, the study team considered that there was good participant diversity in ethnicity and sufficient gender diversity. In addition, we had a broad range of platforms and apps represented in participant use. The authors also identified high-quality dialogue data sufficient for analysis and consistent themes to address the research aims. This resulted in the data achieving good information power [27], and therefore, recruitment was terminated. Information power was used as an alternative to data saturation in this study as the diverse nature of participant narratives meant that we were unlikely to reach a point of saturation.

Data Collection

Written consent to participate was provided by participants before entering the study. Participants were also required to complete a mandatory safety plan, including contact details for someone who could support them, their general practitioner's details (in case serious safety concerns arose), and a self-care plan that was individually designed by each participant to suit their needs (Multimedia Appendix 1). Study information was sent to the parents or guardians of those aged 16 to 18 years as a transparency measure. However, formal parental or guardian consent was not deemed a requirement by the ethics committee given the ages of the participants involved. As part of the study, a distress protocol was developed with a clinician to manage the risk of worsening mental health or increased self-harm or

suicidal thoughts as a result of participation in the study. According to the protocol, participants would first be referred to their own safety plan if their mental health declined as a result of the study. A hierarchy of responses was specified in cases of more serious distress, including the options of offering follow-up support from UK suicide charity "Samaritans" or calling upon the advice of a named senior clinician. However, study-induced distress was not reported by participants during the study, and therefore, such responses were not actioned by researchers.

One-to-one interviews were conducted at baseline and the 3- and 6-month time points via Zoom (Zoom Video Communications) with just the researcher and participant present. The interviews were open-ended and flexible, using probing techniques where appropriate, and structured loosely using a topic guide. The main topics explored were "history of self-harm and suicide feelings"; "current and historic web-based activity related to self-harm and suicide content"; "patterns, motivations, and impact of web-based content engagement"; "critical moments in the web-based content engagement journey"; "keeping safe on the web"; and "experiences of web-based moderation and blocking." The topic guides were originally refined using feedback from 2 lived-experience experts. Throughout the study, the topic guides continued to be iteratively adapted between interviews, grounding question modifications in the study data. The interviews were conducted by ZH, LK, or LB and lasted between 35 and 80 minutes (with baseline interviews averaging 65 [SD 8.55] min and follow-up interviews averaging 45 [SD 2.87] min). They were audio recorded using an encrypted device and then transcribed.

Diaries

Participants completed daily diaries independently between interviews. These diaries served as an ethnographic tool and were introduced at the end of the baseline interview. Blank digital templates were then provided periodically via email. Each covered a 4-week period and had 3 main components (daily recording of content engagement, mood ratings, and a weekly reflection of content impact). Each participant was asked to complete 5 diaries in total. Entries were used to formulate personalized follow-up interview schedules in which further information or clarifications could be sought from participants.

Measures

Self-reported mental well-being data were collected from participants at baseline and monthly intervals to coincide with diary data collection. This was done via surveys on SurveyMonkey and included validated measures for assessing anxiety, depression, and psychological well-being (Multimedia Appendix 2 [28-31]). These data were used to characterize the sample and identify whether changes in mental health and mood reported by participants during the study interviews and in the diary data were reflected in outcome measure scores.

Data Analysis

Descriptive Analysis

Participant baseline demographic characteristics were reported as proportions or frequencies, as appropriate. Individual

trajectories for well-being measures were represented visually using line graphs.

Qualitative Analysis

A trajectory analysis approach [32] was undertaken to interpret interview data temporally using the following steps:

1. Baseline interviews were transcribed, and then, through coding, themes were derived deductively from topic guide questions and inductively from the data themselves. ZH, LK, and LB separately listed preliminary themes and then refined and revised them collaboratively (Table S1 in [Multimedia Appendix 2](#)).
2. Initial matrices were produced for each participant, which included data from the baseline and the 3- and 6-month interviews. These were ordered so that each row was dedicated to a theme established in the previous step. Time points were then assigned to each column. Web-based engagement time points included “initiation,” “historic,” “current,” “never,” and periods of “disengagement and reengagement.” These time points were adapted from the original trajectory approach [31] to preserve the “chronological flow” of the data collected during this study. This allowed us to acknowledge historical content engagement and the nonlinear flow of participant journeys as the levels of engagement fluctuated, ceased, and restarted. This also enabled the inclusion of participants who were only interviewed at baseline (due to dropout) as their data included information about past experiences. Data were formatted according to a “key” using text color to denote the site or platform used and highlighting whether it was related to a significant web event. An event was deemed to be “significant” if the participant recalled it as such or if the researchers found evidence within the narrative that it had a significant impact on the participant’s thinking or behavior. Matrices were developed by extracting relevant quotes or context summaries for 2 participants by ZH, LK, and LB, and once consistency in interpretation was achieved, ZH and LK separately constructed the remaining initial matrices, with ongoing discussion between the researchers to ensure that all the data were captured.

3. Second matrices were then constructed for each participant. These were ordered with the initial themes as column headings. Each row represented an web-based platform or site used by the individual and included condensed versions of the “journey” that participants had experienced for each theme. The comparison allowed us to explore possible patterns in theme content by platforms or sites used. Second matrices were created by ZH for each participant and reviewed by LK and LB.
4. With all matrices complete, ZH, LK, and LB met to discuss similarities and differences across participant trajectories, noting trends, patterns, and outliers. Member checking of transcripts did not occur in this study due to funding and time constraints. During qualitative meeting discussions, overarching longitudinal themes were finalized.

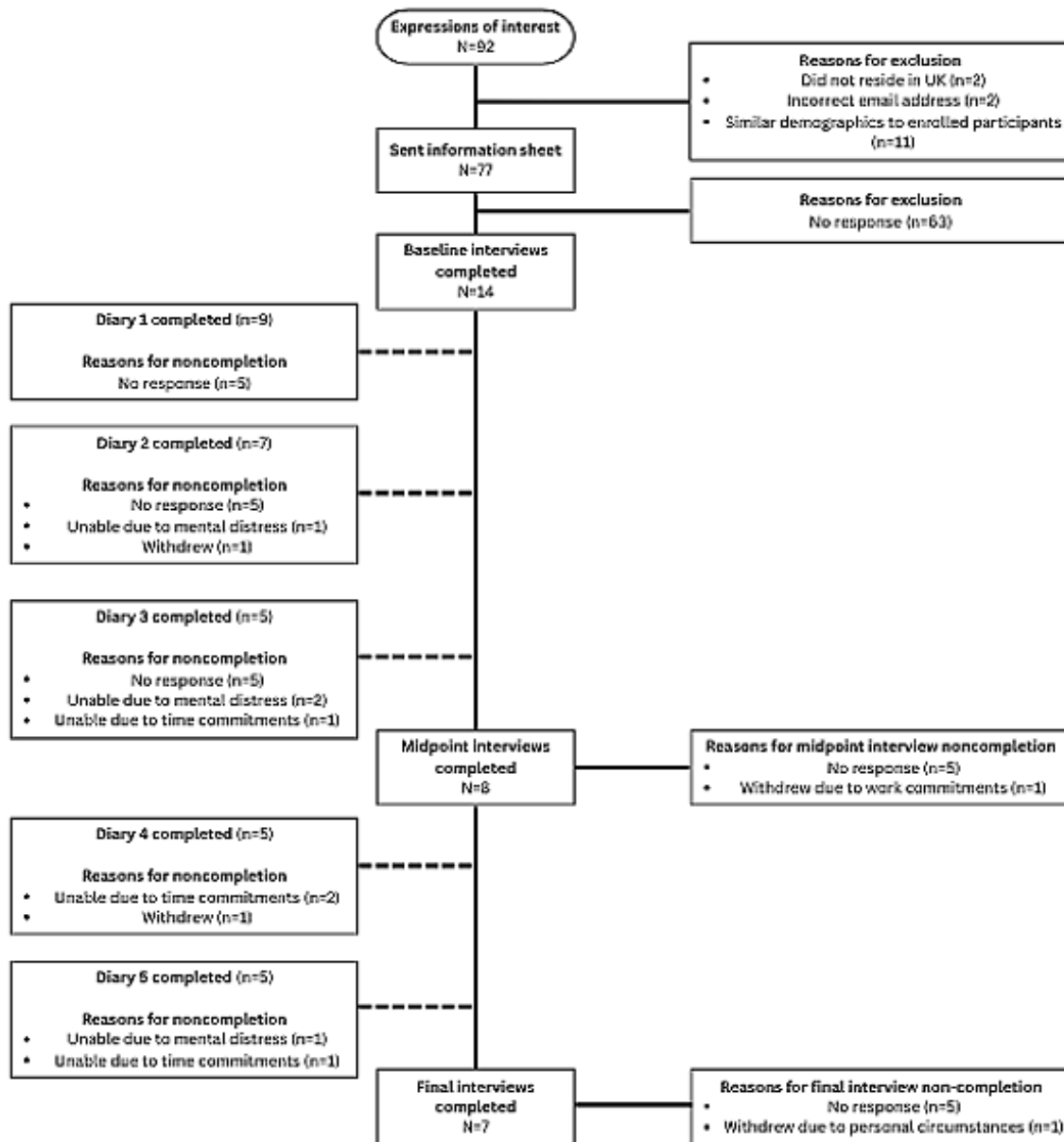
Results

Participant Flow

The participant flow through the study is shown in [Figure 1](#). There were 92 expressions of interest. Of the 77 individuals who were sampled and sent study information, 63 (82%) did not respond and 14 (18%) took part. Data from the expression of interest forms showed that participants were less likely to respond to the research invite if they were younger (aged 16-24 years), had never hurt themselves on purpose, or had self-harmed in the last week.

Of the 14 participants who completed a baseline interview, 8 (57%) completed a midpoint interview, and 7 (50%) also completed the end-point interview. On the basis of preliminary observations of demographic characteristic data from the final sample, it appears that participants of non-British ethnicity may have had a lower likelihood of completing the study compared with those of British ethnicity. However, it should be noted that this observation was not tested statistically (Table S2 in [Multimedia Appendix 2](#)). Throughout the study, participants regularly completed their diaries, with study completers returning 77% (27/35) of the distributed diaries.

Figure 1. Participant flow through the study. Note: diaries were considered completed if one or more responses were provided.



Participant Characteristics

The characteristics of those who completed the baseline interviews (N=14) are displayed in Table 1. Of the 14 participants, 4 (29%) self-identified as male and 10 (71%) self-identified as female. Their ages ranged from 16 to 52 years,

with 18 to 24 years being the most prevalent age group represented. There was a range of ethnicities, with almost half (6/14, 43%) of the participants being from global majority groups. Participants had engaged with self-harm or suicide-related content on a wide variety of sites and platforms.

Table 1. Participant characteristics at baseline (N=14).

Demographic variables	Participants, n (%)
Gender	
Man	4 (29)
Woman	10 (71)
Nonbinary	0 (0)
Ethnicity	
Asian British	2 (14)
Asian other	2 (14)
Black British	1 (7)
Black other	0 (0)
White British	7 (50)
White other	1 (7)
Mixed	1 (7)
Age (y)	
16-17	1 (7)
18-24	7 (50)
25-35	0 (0)
36-45	4 (29)
46-54	2 (14)
≥55	0 (0)
Have you ever hurt yourself on purpose?	
Yes	14 (100)
No	0 (0)
Website or platform used to access content^a	
Instagram	3 (21)
Facebook	5 (36)
TikTok	1 (7)
Twitter	6 (43)
Tumblr	3 (21)
Weibo	1 (7)
Discord	1 (7)
WhatsApp	1 (7)
YouTube	3 (21)
Suicide forums	3 (21)

^aParticipants were able to select more than one option.

Participant IDs

Participant IDs were assigned during the consent process to ensure anonymity. As participants were aware of their assigned IDs, these were changed in the manuscript (see further details in [Multimedia Appendix 2](#)).

Descriptive Analysis Results

Individual line graphs for each well-being measure demonstrated fluctuations in mental health throughout the 6-month study

period that reflected participant journeys recalled through interviews ([Multimedia Appendix 2](#)). One participant, IDB, scored poorer at 6 months on the Entrapment Scale–Short Form (which measures feelings of entrapment in a concise manner) than at baseline; however, the decline was minimal. All other study participants (13/14, 93%) improved from baseline or had no change in total score at the study end point in all quantitative measures, although no statistical analysis of change was undertaken.

Longitudinal Qualitative Analysis

Overview

The themes developed following trajectory analysis included (1) initial engagements with web-based self-harm or suicide content, (2) changes in what self-harm or suicide content people engage with and where, (3) changes in self-harm or suicide behaviors associated with web-based self-harm or suicide

content engagement, (4) the disengagement-reengagement cycle, and (5) future perspectives on self-harm and suicide content engagement. The themes and their constituent subthemes are summarized in [Textbox 1](#).

Within these themes, fluctuations in mental health and control were identified as significant factors impacting behavioral and emotional responses to web-based content and, therefore, will be further explored in the following sections.

Textbox 1. Themes and subthemes.

<p>Initial engagement with web-based self-harm or suicide content</p> <ul style="list-style-type: none"> • Motivations for initial web-based self-harm or suicide content engagement • Experience of engaging with self-harm or suicide content for the first time <p>Changes in what self-harm or suicide content people engage with and where</p> <ul style="list-style-type: none"> • Changes in types of web-based self-harm or suicide content engagement over time • Balancing curiosity and control • Changes in posting web-based self-harm or suicide content over time <p>Changes in self-harm or suicide behaviors associated with web-based self-harm or suicide content engagement</p> <ul style="list-style-type: none"> • Personal risk associated with web-based self-harm or suicide content engagement • The precipitative and protective effects of engagement with self-harm or suicide content on self-harm or suicide behavior <p>The disengagement-reengagement cycle</p> <ul style="list-style-type: none"> • Disengagement from web-based self-harm and suicide content • Reengagement with web-based self-harm and suicide content • Longer periods of disengagement • Limiting content engagement: strategies <p>Future perspectives on self-harm and suicide content engagement</p>

Initial Engagement With Web-Based Self-Harm or Suicide Content

Motivations for Initial Web-Based Self-Harm or Suicide Content Engagement

Our first theme captured historical accounts of engaging with web-based self-harm and suicide content. Participants in this study, most of whom (12/14, 86%) had already self-harmed, initially engaged with web-based self-harm or suicide content following attempts to seek help offline during mental health declines. Those who attended mental health services and received new or changes in diagnoses generally reported leaving unsatisfied, citing reasons that included lack of support, inadequate availability, or feelings of being “dismissed” (IDH; baseline interview) due to a perception of low risk. Some were unable to access services at all or felt that attending was not worthwhile. These mental health declines alongside gaps in service provision were the common catalysts for initial web-based searches for self-harm and suicide content. While some of these searches were motivated by a desire to seek help,

they varied among participants, with some also seeking information on self-harm and suicide methods:

So, I'd been to the doctors...I'd already tried looking for help, I was waiting for a referral to the CMHT [Community Mental Health Team]...And then within a couple of days I'd started lightly [cutting] on my hand then I moved up to my arm, and then I was looking for support groups online, just general support groups. [IDB; baseline interview]

Experience of Engaging With Self-Harm or Suicide Content for the First Time

The experience of initially encountering self-harm or suicide content on the web is captured through the participant responses in [Table 2](#). Only 14% (2/14) of the participants recalled first coming across content unintentionally, with most (12/14, 86%) describing purposeful searches to access material. While most of these searches were for help and support, 14% (2/14) of the participants reported seeking information about methods for self-harm or suicide, and 7% (1/14) of the participants were uncertain about what they were hoping to find but acknowledged that support-focused sites were unhelpful to them at that point.

Table 2. Quotations related to the experience of first encountering web-based self-harm or suicide content.

Reaction	Description	Quotations
Negative	First engagement with web-based self-harm or suicide content produced a negative response.	<ul style="list-style-type: none"> • “That’s not what I was looking for [support sites], I didn’t want help, at that point I was beyond help.” [IDH; baseline interview] • “...I was researching [a suicide method]...what’s required and the best way to manage [that]...It was scary. It’d have been really easy just to have thought, well, actually, I know more about it now and I can do that.” [IDC; baseline interview] • “I received a picture on WhatsApp of someone, of a friend at the time who was self-harming and she basically just sent me a picture of her scars. I think that that image has stayed with me until today, and I think it’s one of the reasons why I’m so careful because it’s not something that I want to see again.” [IDI; baseline interview]
Mixed	Participants experienced both positive and negative responses to the first engagement with web-based self-harm or suicide content.	<ul style="list-style-type: none"> • “...because people were experiencing very similar things to what you were experiencing you wanted to have more of that. It was a good environment in one respect, but it was a very toxic environment in the next because you were listening and you were going, ‘Oh, I’ve been through that.’ But it wasn’t helping. It was actually pushing you down a bit because you were getting ideas [about how to self-harm].” [IDG; baseline interview] • “I think I was just surprised that there was so much content out there. And yeah, that they haven’t been removed, and I think...I guess a sort of comfort knowing that there were others out there who were also going through tough times...And I think, I guess also shocked at how severe some [images of self-harm] are yeah.” [IDL; baseline interview]
Positive	First engagement with web-based self-harm or suicide content produced a positive response.	<ul style="list-style-type: none"> • “I applied to go onto that [Facebook] group just so that I could reach out to people and find out more from survivor-led experiences. And people offered support to each other, and I felt that was quite a good thing to do.” [IDA; midpoint interview] • “It made me feel a lot less alone just knowing, even if they were anonymous people out on the internet that could be wherever in the world, that there were other people, and I wasn’t the only person feeling like this. It was so beneficial, especially as a young teen.” [IDF; baseline interview]

Some participants sought support-related content, and others not intending to access self-harm or suicide content at all unintentionally came across graphic content (eg, images of fresh self-harm) or suicide method descriptions during their first engagements. Those whose initial interactions were with this type of self-harm or suicide content described feelings of distress even when this was the content they were seeking out. Some of these participants (2/14, 14%) recognized that this content could inadvertently validate and trigger their own self-harm and suicide feelings and behaviors, making them feel more at risk. In cases in which participants first engaged with web-based self-harm or suicide content in a discussion forum or peer support group, they were more likely to respond positively, describing how they felt less alone and were able to share experiences with others. However, some participants (2/14, 14%) had mixed emotions—it was comforting to know similar others existed, but processing extreme content was challenging and subjected them to information about novel self-harm and suicide methods, revealing their lack of control over what they were exposed to.

Changes in What Self-Harm or Suicide Content People Engage With and Where

Changes in Types of Web-Based Self-Harm or Suicide Content Engagement Over Time

All participants continued interacting with online self-harm or suicide content after their initial encounter even if it had been a negative experience. In cases in which they had positive initial engagements, participants continued to use the same platforms to access self-harm or suicide content in the long term. When those platforms or sites became obsolete, they sought out equivalent content in other web locations. Participants who had negative initial interactions accessed different platforms or sites searching for self-harm or suicide material that resonated with them.

Although participants had self-harm and suicide content that they accessed in a stable and routine manner, many also described occasions when they would change what they were accessing. Most participants (12/14, 86%) explained that different content satisfied different needs depending on their current mental state or mood. Examples of this can be found in [Textbox 2](#).

Textbox 2. Web-based self-harm and suicide content accessed during mental health changes.

Change in content accessed due to mental health declines

- “On a good day I can be in there and I can be supporting others and helping them and building them up. And then on a bad day I’ll be the one looking for support and asking for somebody to you know pick me up a little bit. So, it very much depends on what mood I’m in that day to be honest.” [IDB; baseline interview]
- “If you are depressed and you start like looking at videos that are to do with that sort of thing it’s so easy for you to be in the spiral of just like looking at more and more content about suicide and stuff like that...” [IDK; baseline interview]

Change in content accessed due to suicidal feelings or intentions

- “When I am thinking about self-harm, I will just look it up online. I go to the text service when I have suicidal feelings.” [IDF; baseline interview]
- “That’s when [‘if I’m in a really bad crisis’] I’m more seeking it out, so Tik Tok I’m not actively seeking out that content [ok] but that’s when I’m actually seeking it out, thinking I want to die, that’s when I start accessing suicide forums and stuff.” [IDK; baseline interview]

Change in content accessed due to mental health improvements

- “I’ve reached a place where I feel like I want to kind of, hear more about recovery and things like that. I think that’s why I found this sort of [‘recovery-based images’] content useful to look at. And I think that, I don’t think I’m triggered by it, but I also don’t want to interact with that kind of content where people are talking about their own [recovery] journeys because I’m not in that kind of place or not in a place where I want to hear about that kind of stuff at the moment. So, yeah, kind of like more interactions with the positive stuff, I think.” [IDI; final interview]

Dips in mental health often resulted in changes in the way participants engaged on the web, such as posting their own self-harm or suicide material rather than just interacting with others’ content. In cases in which participants experienced sustained episodes of poorer mental health, self-harm and suicide content was also seemingly accessed more frequently and sometimes uncontrollably through “habit” (IDG) or “addiction” (IDC and IDH), with 21% (3/14) of the participants describing it as falling down a “rabbit hole” (IDJ, IDL, and IDF). In total, 7% (1/14) of the participants reported how this compulsive engagement with self-harm or suicide content interrupted elements of their usual social and occupational functioning:

Even through work time I would take ten minutes and just read some of it. [IDH; baseline interview]

Directly questioning participants about web-based engagement when feeling “actively suicidal” elicited similar reported changes in behavior. A couple of participants described engaging with different content—notably turning to web-based suicide organizations and charities or friends and family members offline when they needed support for suicidal thoughts rather than their usual web-based resources for self-harm or suicide content. However, another 21% (3/14) of the participants described how prominent suicidal thoughts were more likely to result in them returning to prosuicide forums, where they would seek or check resources for their own suicide plans.

Improvements in mental health saw participants more likely to transition to web content of a recovery-based nature while often sticking to the same web-based locations. Some participants (3/14, 21%) also attempted to limit web engagement with greater use of offline resources such as community help centers or taking part in meaningful activities.

Balancing Curiosity and Control

Other participants came across content unexpectedly in their web journey or seemed to spontaneously seek out different self-harm or suicide content due to “curiosity.” Some described the ability to negotiate novel self-harm and suicide content with

a developed sense of control over time, skipping over or avoiding engaging with content that was undesirable to them:

Being able to scroll past content with trigger warnings of self-harm pics has been quite a new thing. Like in the last year-ish, before then I wouldn’t have been able to have done that. I’d have looked. [IDJ; baseline interview]

However, others described tensions between curiosity and control and how that curiosity led them to seek out different self-harm or suicide content. For example, 14% (2/14) of the participants, who read a news article on a person’s death by suicide that referenced web-based prosuicide forums, went on to search for them:

...I saw it [article on death-by-suicide of person who used pro-suicide forums] in the news. When you see something in the news, especially on the BBC website you know...it’s quite serious stuff. So, then you end up looking further. Now sometimes you have to be careful because you get drawn into it and I think you have to sort of say to yourself, “I’m only going to spend a few minutes doing this...” [IDE; midpoint interview]

The functions of social media sites (eg, hashtags or algorithms) could also enable unintentional content encounters, making control over engagements less feasible:

I guess sometimes that like tags on social media and...it’s usually by chance, I don’t actively go and seek them, but sometimes it appears and then I kind of just go down a rabbit hole of looking at more of such content. Even though I didn’t do it intentionally. [IDL; baseline interview]

Another participant explained that, in transitioning from self-harm and suicide content that no longer resonated, they had less control over what they engaged with:

I think recently, it’s like I don’t know what I’m looking for, but it’s like I know that I haven’t been able to

find it...So, I think it's normally looking through my explore page instead of searching for anything in particular... [IDI; midpoint interview]

Changes in Posting Web-Based Self-Harm or Suicide Content Over Time

For some participants (5/14, 36%), posting content seeking help and support regarding self-harm or suicide feelings or looking for ways to stay safe while self-harming was an early action in their web journey (IDA, IDB, IDC, IDD, and IDG). Others also posted detailed descriptions of suicide methods they were considering on discussion forums (IDH, IDK, and IDD), blog posts detailing their own self-harm and suicide feelings (IDN), and images of quotes on Instagram with captions about their mental health (IDI). One participant sent images of their own self-harm via direct messaging after other users requested them (IDK).

A total of 21% (3/14) of the participants in the study refrained from publishing their own content publicly (IDF, IDL, and IDJ). Of these 3 participants, 2 (67%) posted content privately (meaning that it was posted on the web but was only visible to them). Both participants described this as their way to “vent” (IDF) or “rant” (IDL) when upset and an opportunity to document their journey.

Notably, all 3 “observation-only” participants mentioned valuing their anonymity in the web space and refraining from online community interactions. They also emphasized that the potential for posts to negatively affect others deterred them from posting self-harm or suicide content publicly:

I always felt quite conflicted about reposting other people's content related to it [photos or videos of fresh self-harm]. I feel like it's one thing for me to look at it because they've posted it...versus me reblogging it to my own blog. I don't know. It's odd to explain it but it just felt weird. [IDF; baseline interview]

Another participant reported posting content in one context (asking for support on a Facebook group) but not posting “graphic images” (IDC) due to fear that it may cause harm to children. This particular concern for young people viewing content was echoed by IDF, IDH, and IDK.

IDN, who initially described making public blog posts about their own self-harm, later made these private due to a realization that the material may negatively affect others as well as an attempt to maintain anonymity. IDI also reported a change in posting behavior during and as a result of taking part in the study. After initially posting about their experiences in an attempt to raise mental health awareness, they reflected on their tendency to put a “positive spin” (IDI) on content, and by the 6-month follow-up interview, they had reduced the frequency of their posts as they began to question their own authenticity. They considered that, if they posted about their negative experiences, it would likely have a harmful effect on others, and so they refrained from posting.

Finally, one participant also noted that access to psychological therapy reduced their need to post on the web for support:

[I haven't posted] for quite some time actually. I can't remember the last time I did that. It would be over a month ago easily. Yeah, I haven't needed to really. [IDA; final interview]

Why do you think that is? [Interviewer; final interview]

Because I could handle whatever I was thinking probably on my own or bring it to the next...because I'm having weekly sessions with my psychologist... [IDA; final interview]

Changes in Self-Harm or Suicide Behaviors Associated With Web-Based Self-Harm or Suicide Content Engagement

Personal Risk Associated With Web-Based Self-Harm or Suicide Content Engagement

As described previously, some participants identified risks after their initial exposure to web-based self-harm or suicide content (Table 2). Others recognized potentially harmful consequences after longer periods of engagement. Some thought that the content they engaged with gave them implicit “permission” to carry out similar self-harm or suicide behaviors (IDG, IDH, IDJ, IDK, and IDN):

It makes it [completing suicide] feel less scary and like being able to hear people talk about what happened to them, them saying it's not that bad, like it wasn't...it just felt like nothing, it makes it feel a lot easier to do it if you know what I mean? [IDK; baseline interview]

Some found that their own self-harm or suicidal behaviors were influenced by self-harm and suicide information they had gathered on the web (IDJ, IDK, IDL, and IDB):

...there were some posts which would link to other websites where you could get resources [information on overdose statistics]. I'd say definitely at the start of my mental health journey that was quite a turning point for me. Because it was just an idea and then it became a possible thing to do. [IDJ; baseline interview]

Another participant experienced feelings of jealousy over the self-harm people had engaged in, which resulted in them feeling the need to escalate their self-harm behaviors:

I think that was that self-comparison to myself...maybe I'm being too scared or I'm not trying hard enough... [IDL; baseline interview]

The Precipitative and Protective Effects of Engagement With Self-Harm or Suicide Content on Self-Harm and Suicide Behavior

The feelings and behaviors that participants experienced following engagement with web-based self-harm or suicide content are shown in Table 3. Content could be precipitative or protective for participants depending on when they encountered it in their journey. Several participants (5/14, 36%) recalled engaging in self-harm and suicide behavior as a result of engaging with web-based content. A few of these participants

(4/14, 29%) went on to describe changes in their self-harm and suicide behaviors related to content engagement over time, implying that the effect could change from precipitative to protective. This included living vicariously through others' self-harm images (IDF and IDL), finding content engagement relaxing (IDC), and using content searches as a way to delay or stop their own suicidal behavior (IDD). One participant suggested that such changes in behavior were due to building their own mental resilience over time:

I wouldn't say the internet content changed, it would be more like I've changed to deal with what the internet's providing me. [IDJ; baseline interview]

Another participant recognized the need to consistently "check-in" with their own mental health before engaging on the web:

It's just about how I'm feeling, like do I feel like I have the capacity to deal with the internet really, do I actually want to look at what people are saying and what people are posting. [IDF; midpoint interview]

For some, there was less consistency regarding whether engagements with self-harm and suicide content would result in helpful or harmful circumstances. This was exemplified by one participant who stated that their searches were usually protective and kept them occupied when their suicidal thoughts were most intense:

I think there is a part of me that does it [conducts searches for self-harm and suicide content] to buy time. [IDB; final interview]

However, this participant also reported attempting a new form of self-harm at the midpoint interview after learning about it through a peer support group on Facebook.

Table 3. Precipitative and protective effects of web-based self-harm or suicide content engagement identified by participants.

Factor and description	Quotations
Precipitative factors	
Self-harm or suicide behavior as a consequence of engaging with web-based self-harm and suicide content	<ul style="list-style-type: none"> “It could also be really detrimental because many times, I would just come away feeling much more triggered than previously and then would engage in the behaviour [self-harm].” [IDF; baseline interview] “One of the [posts] got taken out of a group [by me] because it was talking about bloodletting and since then, I've bought syringes and needles to try and do it myself.” [IDB; midpoint interview] “How did you then cope with the fallout of what you'd seen [distressing self-harm and suicide content]?” [Interviewer; baseline interview] “I coped by self-harming. Yeah, and I write lots as well. So yeah, writing about how I feel and what I saw.” [IDC; baseline interview]
Protective factors	
Vicarious experiences through self-harm or suicide content	<ul style="list-style-type: none"> “It would mainly be trying to vicariously live out things through other people. So, I had a particular urge but wasn't in a position where I felt like I could self-harm or necessarily wanted to and almost living those experiences through somebody else's experience which was one of the ways that it [viewing self-harm material] could be really beneficial for me because it could almost meet that urge without me having to engage in the behaviour.” [IDF; baseline interview]
Delaying or stopping own self-harm or suicide behavior	<ul style="list-style-type: none"> “I don't really need to research it [suicide method] anymore. Sometimes, I do it anyway and I just re-research, re-read it and re-check my facts but it can be a way of preventing me from doing anything.” [IDB; final interview] “How do you mean?” [Interviewer; final interview] “It's like there are levels to it, aren't there? That's what I find anyway. It starts with thoughts, then it turns to urges and once you get to that urge stage, you need to feel like you're doing something, whereas, re-researching it [suicide method] is better than actually putting the tablets in your mouth. It gives you that extra step before you get to that point, if you see what I mean.” [IDB; final interview]
Calming effect	<ul style="list-style-type: none"> “How did you feel [coming across images of self-harm]?” [Interviewer; midpoint interview] “Quite relaxed because that's what I do [self-harm], so I could identify with them, those people who'd done things like that.” [IDC; midpoint interview]

Disengagement-Reengagement Cycle

Disengagement From Web-Based Self-Harm and Suicide Content

Most participants (8/14, 57%) reported entering a cycle of disengagement and reengagement during their web-based self-harm and suicide content journey. Disengagement was usually temporary, with participants choosing to have “no phone days,” deleting their accounts, finding offline activities to take

part in, or being forced to disengage due to lack of internet access.

Most often, disengagement was purposeful but impulsive. It would usually occur during periods of compulsive engagement when participants recognized a lapse in their control or as a reaction to a significant life event that resulted in mental health decline. Life events that occurred during this study included suicide bereavement, hospitalization, bullying or victimization, and experiences of exam- or work-related stress. The act of intentionally disengaging from self-harm or suicide content was

usually a conscious decision to reclaim control over their web-based actions.

A total of 14% (2/14) of the participants reflected on changes in their disengagement behavior while in the study (IDC and IDI). Previously, similar to other participants, they reported a tendency to compulsively access content during periods of poorer mental health followed by impulsive disengagement. However, at the 6-month interview, both participants described an improved ability to recognize their patterns of web-based behavior (Table 4). This understanding and insight empowered the participants to purposefully disengage during declining mental health episodes as a strategic means of regaining control over their behavior.

However, when other participants were forced to disengage from content because of intermittent internet access or physical health problems, they were often left with feelings of loss. Although one participant described this unintentional disengagement as an opportunity for brief respite from self-harm and suicide content engagement, its existence remained a reassuring presence:

I knew I could access them if I needed to, but I thought, “No, I’m having a week off and I’m going to try and distance myself from this as much as I can.”
[IDA; final interview]

Table 4. Reasons for disengagement from web-based self-harm or suicide content—from final interviews.

Participant ID and reason for disengagement	Quote
IDC	
Mental well-being	“It felt like I needed to look after myself and that I needed that break to try and keep myself safe. One of the things that this research has taught me and helped me understand, it’s helped me understand more about how social media impacts upon me. So, I think social media can be a source for good. I think you also need to recognise that sometimes you need a break.”
Regaining control	“I really crashed down, and it scared me because I’d had a lovely weekend. Things are generally a lot better, and it scared me in that I can still crash down and fluster myself. I didn’t trust myself to be researching suicide, self-harm...And there was a part of me that knew that I wanted to live, there was a part of me that knew I could spiral out of control, and I didn’t want to spiral out of control. And I’ve alluded to the fact that I’ve learnt personally a lot about myself during the six month’s research and how I use social media. And for me that Monday when I made that decision [to disengage from self-harm content] it was really positive.”
IDF	
Regaining control	“I think just to prevent myself from falling down a rabbit hole and looking at content that I know wasn’t good for me. And just feeling like so I’ve always been one of those people that I sort of like to sort of physically remove myself from things and remove things from me. So that’s one of the reasons why I do that.”
Mental well-being	“So, I think it was about a month ago now and someone who was quite active in Twitter (X) and the mental health recovery community passed away from what I feel was suicide. That’s not been confirmed but when all of that happened, I did take a couple of days off the internet just to, I guess, process things there.”

Reengagement With Web-Based Self-Harm and Suicide Content

Participants described various reasons for reengagement with web-based self-harm and suicide content, including a “fear of missing out” with the community (IDA, IDC, IDM, and IDK), wanting to use the site or platform to access other types of content (IDE, IDF, and IDI), procrastination or boredom (IDI and IDK), and the need to perform web-based responsibilities (eg, work or moderating roles within self-harm and suicide communities; IDI, IDA, and IDB). Some participants (5/14, 36%) claimed that they weighed the advantages and disadvantages of web-based content engagement before reengaging. Several participants (3/14, 21%) felt that the benefits of reengaging with self-harm and suicide content, such as feeling part of a community, were enough to justify the potential risks. As this participant noted, while the experience could be upsetting at times, it was still considered useful in light of the rewards of engagement:

With Twitter [X], I deleted that as well, but I felt like actually I missed the community and felt out of touch with people, so I actually found that useful

[reengaging], as much as sometimes it’s upsetting, it was useful. [IDK; baseline interview]

There were also differing accounts of reengagement due to mental health improvements and declines. One participant described feeling more in control once their mental health was stable:

I think I was in a better place emotionally and with my mental health...And I just felt stronger, I genuinely felt stronger and more positive. It’s a better time of year for me...I’ve started some new medication...So, I think that’s a factor as well and me feeling stronger to go back online. I just felt ready. [IDC; final interview]

Similarly, another participant felt that they were more able to view and contribute to self-harm and suicide content in a positive way when their mental health improved:

When my mental state is better, and I can go back on. I feel like I can share, and I can help someone. [IDM]

Alternatively, some participants (2/14, 14%) described past reengagement with self-harm and suicide content to “punish”

(IDF; baseline interview) themselves for thinking about or carrying out self-harm behaviors:

I think at that time I was kind of trying to make myself feel worse, because it was like, “You need to feel more guilty for what you’re doing.” [IDI; final interview]

However, these participants described differences in their reengagements over time. IDI reported how their reengagement behavior changed during the study. When feeling low, they now went on the web and sought out non-self-harm or suicide content.

Other participants also described attempts to engage with self-harm or suicide content differently during the reengagement period with the aim of regaining control. This included observing interactions rather than actively participating or limiting engagements with specific content on platforms or sites:

Recently I’ve just been viewing [prosuicide threads] and I’ve got to fight the urges [to interact]. [IDH; baseline interview]

However, most participants who disengaged briefly would return to their usual use of web-based content. This reengagement process highlighted weaknesses in participants’ ability to exercise control over web-based actions, leaving users vulnerable to reencountering triggering content on the web and beginning the disengagement-reengagement cycle again:

I basically quit Tik Tok for three weeks because I was like I just can’t deal with it anymore because it’s just so hard to block everything and I was also thinking is it actually good for my mental health and it’s not... [IDK; baseline interview]

...but you are back on TikTok now, is that right? [Interviewer; baseline interview]

I think I was just bored really, and I thought do you know what I’ll just download it for the afternoon, and... [IDK; baseline interview]

Longer Periods of Disengagement

In total, 14% (2/14) of the participants in this study disengaged for up to a month before reengaging with specific platforms. One of these participants disengaged after a second death by suicide in their Twitter community. Notably, an earlier death by suicide of another member of the same community had increased their frequency of accessing the platform.

During their Twitter disengagement, the participant continued their engagement with a self-harm support group on Facebook, where they felt less connected:

I think because I haven’t known them [Facebook users] so long and there’s certain people [on Twitter, subsequently rebranded as X]...who post frequently, several times possibly in a day...I think the more you get to know people and recognise the handles, I know it sounds bizarre, but you feel yourself becoming closer to them. [IDC; final interview]

Despite this, they also reported beginning to reengage with Twitter toward the end of the study:

I think just because I feel a bit better, I wanted to check-in on other things on there on my newsfeed, wall thing. [IDC; final interview]

One other participant disengaged twice from a prosuicide forum. First, they described disengaging following an article on the parents of forum members who had died by suicide. The participant reached out to the parents, and the resulting relationship led to their disengagement:

...they’ve told me I need to get off the site. [IDH; baseline interview]

However, they reengaged shortly after this event after wanting to check whether “they [the site] put the resources [suicide methods literature] back” (IDH; baseline interview) following their removal after the media article publication.

At the midpoint interview, IDH had again disengaged from and reengaged with the forum following the death by suicide of a relative. On describing their reengagement, they reported that “it was to check [that] the sources of getting stuff [suicide materials]...are still available” (IDH; midpoint interview) as they were aware of scams related to sourcing materials and wanted to verify that their plans would still be viable.

Limiting Content Engagement: Strategies

After spending time engaging with web-based self-harm and suicide content, half (7/14, 50%) of the participants began to develop strategies to limit their content engagement. These included less “arbitrary ‘liking’” to curate their feeds (IDI), clearing search histories to “remove temptation” (IDJ), “blocking” or “muting” certain terms or phrases—such as “suicide” and “self-harm” (IDC, IDF, and IDK)—closing their direct messages so that other users were unable to message them (IDI and IDH), “self-banning” so that they were unable to post (IDH), distracting themselves with positive web-based content (IDE, IDI, IDC, and IDF), “starting new accounts” to avoid tailored algorithms (IDK), and distancing themselves from a self-assigned “role” such as being a mental health advocate (IDJ and IDI). In this study, we observed that younger participants predominantly used these strategies, possibly because of the enhanced accessibility to safety features on the platforms or sites they frequented or their proficiency in digital skills. However, it is noteworthy that most participants regardless of age reported learning digital safety methods of limiting engagement over time through their experiences on the web:

As I’ve got older I’ve realised that actually you know what you see online can really impact on you, and that you know no-one’s going to police it for me so I have to be sensible about the types of people that I follow and the types of things that I do online. I think that’s something that came with sort of getting a bit older. [IDF; midpoint interview]

Future Perspectives on Web-Based Self-Harm and Suicide Content Engagement

When asked, none of the participants reported a desire to disengage from web-based self-harm or suicide content entirely in the future. Many alluded to the nonlinear nature of their engagement, recognizing difficulties during previous attempts

to disengage. Some also described a sense of comfort and reassurance knowing that content continued to exist on the web:

It's a cushion for people who need that. [IDA; midpoint interview]

In addition, others reported a desire to “give back” and described having a peer support role themselves as a future goal following their recovery (IDI, IDM, and IDB):

I'm looking forward to where I improve myself, and maybe be able to talk to more people and if possible, reach out to them, and offer that help. [IDM; final interview]

I'm also very passionate about sharing stuff I've learnt. When people are in that place that I remember being in and you can see it from their posts, I think, "I've just learnt about something that will help them. I'll pass that on to them." It's helping to build that confidence back up to do those posts and say those things on there. [IDB; final interview]

Some participants in this study (3/14, 21%) also highlighted that they were unable to find alternative web-based or offline spaces that satisfied their current needs. One participant mentioned that disengaging from their current preferred site or platform could be detrimental and so expressed no wish to “move on”:

What I'm trying to say is that there is nowhere for people when they come off that website. There's no safe space. There's nowhere. If you've been on that particular site [prosuicide forum] for the reason of wanting to die and you didn't, there's nowhere. You'll go on something and just get these silly comments or things where there's lack of understanding that just escalates a situation. [IDH; midpoint interview]

A few participants in this study (3/14, 21%) did recognize the potential costs associated with continuing to engage in web-based spaces with self-harm and suicide content but compromised, stating that “I do feel that the benefits outweigh the risks” (IDC; baseline interview). For these individuals, the draw of the positive aspects of such content was strong enough to justify the potential negative consequences. Other participants (2/14, 14%) struggled to weigh the risks and benefits of engaging with self-harm and suicide content as they felt that the positive and negative aspects of engaging with content were more intertwined, making it difficult to control what they were exposed to:

I'd say that online is very complicated, depending on what you feed your mind, because it has both positive and negative information, so sometimes it's good to your mind, and sometimes not. Also, if you are coming across lots of negative things in a group, that can be harmful, like self-harm pictures. But it's also good to look in those groups for people who are offering help for those things, so that you are learning how to help yourself. [IDM; final interview]

Ultimately, this resulted in both sets of participants remaining vulnerable to the negative effects of harmful content as they

continued to engage with web-based self-harm and suicide material.

Discussion

This study showed that those engaging with web-based self-harm and suicide content experienced nonlinear journeys that were characterized by 5 key themes: “initial engagements with web-based self-harm or suicide content,” “changes in what self-harm or suicide content people engage with and where,” “changes in self-harm or suicide behaviors associated with web-based self-harm or suicide content engagement,” “the disengagement-reengagement cycle,” and “future perspectives on web-based self-harm and suicide content engagement.”

Cognitive Flexibility Versus Cognitive Rigidity

Constructs that may explain behavior change and maintenance within these themes are cognitive flexibility and its counterpart, cognitive rigidity [33]. Cognitive flexibility refers to an openness in thinking and behavior, which allows an individual to consider alternative perspectives and approaches. In contrast, cognitive rigidity is the tendency to adhere to specific thought and behavior patterns, making it challenging to change one's mindset or actions [33]. Previous research has identified a relationship between cognitive rigidity and suicidal ideation [34] and between cognitive rigidity and self-harm [35]. Another study showed that cognitive flexibility can result in engagement in multiple methods of self-harm [36]. This indicates that the construct of cognitive flexibility may provide important insights into the behavior changes over time associated with web-based self-harm and suicide content engagement. This discussion will explore the ways in which cognitive flexibility was impacted by participants' mental health and control over decision-making and how this influenced their web journeys.

Previous research has identified gaps in clinical support as a key motivator for web-based self-harm and suicide content engagement [2]. The causes for initial engagement in this study were consistent with this, with participants reporting a lack of support but also a reluctance to engage with clinical services due to previous experiences. This suggests a high level of cognitive flexibility among participants during their first engagement with web-based content, with mental distress and a lack of alternative resources potentially triggering participants to be more open to web-based options. This emphasizes the critical need for accessible offline options during the early stages of mental health decline, preventing vulnerable people from resorting to web-based avenues where they may lack the control or knowledge to engage safely.

When participants were unable to find content that was immediately desirable to them, they explored different self-harm or suicide-related material on the web. Often, this led to spontaneous browsing of self-harm and suicide-related links or hashtags, a behavior characterized as “pessimistic browsing” [13]. While this reflects a high level of cognitive flexibility among participants, it also indicates what might be a lack of behavioral control, making participants vulnerable to potentially harmful encounters. Later on in web journeys, when browsing routines had been established, some reported similar bouts of

“pessimistic browsing” and harmful behaviors that they considered spontaneous. These episodes of cognitive flexibility were usually triggered by unexpected exposure to web-based self-harm or suicide content, where impulsive tendencies resulted in exploring this novel content further or, in one case, in trying a new self-harm method. This indicates that unexpected engagements with self-harm or suicide content may act as a stimulus for activating cognitive flexibility, resulting in changes in behavior [37]. When experiencing poor control, this cognitive flexibility may lead to a willingness to engage in potentially unhelpful or harmful behaviors when engaging with self-harm or suicide content [38].

Outside of episodes of cognitive flexibility, participants largely accessed web-based self-harm or suicide content in a routine pattern while also reporting a greater feeling of behavioral control. This cognitive rigidity often worked as a coping mechanism allowing for regular engagement with resources of help and support. However, in instances in which content included images or videos of “fresh self-harm,” suicide, or self-harm and suicide method information, repeated engagements were more likely to have negative effects on participant well-being and sometimes led to increased severity of harm to themselves. This shows that, while some perceived their cognitive rigidity as a form of control, it may ultimately have diminished their ability to make decisions to protect themselves and seek alternative positive coping mechanisms [39].

Similarly, participants reported increased engagement with self-harm or suicide content during dips in their mental health, which were prevalent in this study, as indicated by fluctuations in their well-being measure outcomes over time. These engagements, recalled as “habitual” or “addictive,” highlighted a loss of control during these mental health dips. Previous research has shown a relationship between cognitive inflexibility and addictionlike behaviors [40,41], and a more recent study [42] has indicated that distress-driven impulsivity, in which a person is likely to make rash decisions due to a negative mental state, alongside cognitive rigidity, can lead to addictionlike eating behavior. This emphasizes the potential risk of overreliance on web-based self-harm and suicide content as a coping strategy, particularly during periods of mental health decline, when participants may become more vulnerable to the content they are engaging with. The addictive nature of this behavior also has the potential to negatively impact other important aspects of people’s lives, such as social or occupational functioning [43].

Disengaging and Reengaging

Key to self-preservation during the web journey was participants’ ability to disengage from web-based spaces. Most participants recorded disengagements in their web journeys in response to life experiences or stressors, such as work stress, bereavement, or a rapid deterioration in mental health. This indicates a resurgence of cognitive flexibility, which reflects previous research showing that individuals become more open to alternative solutions when their perspective is challenged by a significant life event [44]. Although participants demonstrated disengagement attempts from the content during these times,

they were usually temporary. This represents a brief state of cognitive flexibility, with reengagement often occurring within days. When disengagement was longer, it tended to coincide with more significant life events such as bereavements, which may indicate more prolonged changes to behavior following extreme circumstances and mental health declines.

Participants also reported that their mental state dictated whether they returned to more helpful or harmful content during the reengagement period. Participants experiencing poorer mental health were more likely to reengage with content they described as “negative” as a type of self-punishment or as a preventative measure against potentially worsening self-harm or suicide behavior. They were also more likely to post their own content, which included help-seeking comments, suicide method inquiries, and “depressive” blog posts. This showed that, although some participants attempted to use their online communities for help during mental health dips, others could find themselves returning to potentially unhelpful or harmful situations. This reflects previous research showing that “active” suicidal ideation is associated with greater cognitive rigidity compared to “passive” suicidal ideation [45]. Often, when reengaging during mental health declines, use would also regress to “addictive” or “habitual” engagements. However, when experiencing mental health improvements during reengagement periods, those who had previously engaged with more “positive” or “recovery-based” content would be more likely to return to this material. This indicates that cognitive rigidity is influenced by mental health state and that, when experiencing mental health changes, participants’ well-being is reliant on earlier web-based encounters with self-harm and suicide content.

Upskilling Users

Despite this, some participants did experience lasting adaptations to the ways in which they interacted with the content. These more enduring changes were attributable to the skills that participants reported developing in digital efficacy and metacognition. Digital efficacy skills include the ability to use web-based safety mechanisms such as muting, blocking, and self-banning. Participants with digital efficacy skills in the study felt safer and more protected, which acted as a preventative measure against cognitive rigidity. In this study, these participants were likely to be younger, which reflects research showing that digital literacy skills are significantly better in younger cohorts [46]. Despite this, evidence also shows that digital literacy skills can be built over time [47]. This is consistent with the experiences of some participants in this study who reported that their web experiences prompted them to organically develop digital skills and strategies to stay safe over time. This finding has important implications for industry leaders, who should be encouraged to consider ways in which they can empower users by improving accessibility to safety mechanisms on their platforms and sites.

Metacognition skills, or the ability to reflect on one’s own thoughts and behaviors to change one’s responses, were evident in some of the participants [48]. Specific metacognitive abilities such as self-awareness and self-regulation resulted in greater control over their cognitive flexibility. Some described gaining metacognition skills such as self-awareness before their

participation in the study, which allowed them to recall changing their responses to content from self-harm behavior to vicarious viewing of material. Others identified metacognitive skills gained through therapeutic input as well as through monitoring web-based behavior and reflecting on it during the study. This may reflect a Hawthorne effect [49] in which participant behavior shifts due to their awareness of being observed in a study. Several diary and ecological momentary assessment interventions have resulted in improved metacognitive skills [50] (Haime, Z, unpublished data, January 2024), and it is possible that metacognition was acquired in this study as a result of completing the research diary. In cases in which participants' metacognition developed during the study, they noted improvements in their mental health, also indicated by improvements in their well-being outcomes over time. This resulted in the type of self-harm and suicide material they engaged or reengaged with changing from "negative" or "depressive" to "recovery-based" or "positive" in nature. This shows that self-awareness and control while experiencing mental health improvements lead to positive content engagement during periods of cognitive flexibility in this population and has important implications for the development of future target metacognitive interventions.

Remaining Vulnerable on the Web

However, shifts toward recovery-based content did not necessarily mean that participants were able to fully disengage from their previous self-harm and suicide material. Sometimes, as recovery-based content coexisted alongside more harmful content in web spaces, there was no alternative place in which to access it. On the other hand, some participants expressed a strong connection with the communities they had previously engaged with and reported intentions to remain active in these spaces with a desire to provide support to others. While this altruistic act had benefits, including the ability to continue drawing on support when needed, it left them vulnerable to potentially triggering content. These findings emphasize the strength of web-based self-harm and suicide spaces as a source of comfort and security, which is consistent with previous research on engagement motivators [2,7,10]. Thus, although participants became more aware of the negative outcomes of engaging with web-based self-harm and suicide content and were better able to manage them, the perceived benefits of being involved in a community of like-minded individuals with similar experiences often outweighed the potential costs.

Limitations

Participants in this study used a diverse range of web-based platforms to access self-harm and suicide content, meaning that attempts to identify patterns in behavior related to the sites used were challenging. However, as common behaviors were observed across participants, it was possible to draw conclusions more broadly about how people engage with web-based self-harm and suicide content over time. Diaries in this study were completed daily by participants, but many had missing entries or were filled out retrospectively. This diluted the advantages of "in-the-moment" diary data capture and resulted in some interview topic guides being less informed by participant data. Despite this, participants reported finding the

diaries largely acceptable, and some reported additional benefits to their metacognitive ability related to their completion [51].

While visually observing quantitative data allowed us to identify patterns consistent with participant-reported mental health fluctuations and slight improvements toward the end of the study, our inability to conduct statistical analyses prevented us from identifying any significant differences in participant well-being changes. However, the rich qualitative data and trajectory analysis provided valuable insights into the individual pathways and factors influencing web-based engagement.

In terms of participant characteristics, this study had an underrepresentation of male individuals. Although steps were taken to target male-orientated web spaces for recruitment, uptake remained poor. Furthermore, responses to recruitment were limited, which resulted in possible selection bias and may have affected the representativeness of the sample. In addition, we did not collect data on the educational level or socioeconomic status of the participants involved, limiting our understanding of how demographic characteristics may affect web-based experiences. Half of those recruited at baseline were also lost to follow-up. Strategies were undertaken to limit attrition, including at least 3 attempts to communicate with participants before they were considered lost to follow-up. High attrition rates are consistent with longitudinal studies of self-harm and may represent a selection bias among study completers [52]. Finally, although cognitive flexibility provides a useful framework with which to interpret our findings, it is important to acknowledge that there may be alternative explanations.

Future Implications

The findings of this study have shown that there are ongoing challenges in navigating the web environment for those engaging with self-harm and suicide content. A key priority for future research should be to establish how engaging with web-based content can be better managed in this population. Consequently, the following should be considered:

1. Inaccessibility to offline support was a significant motivator for participants' willingness to explore web-based self-harm and suicide-related resources. Therefore, the availability of offline help and support is necessary to limit or moderate initial web-based engagements.
2. This study offers evidence that greater metacognition and digital efficacy can positively influence web-based behavioral control. As individuals are unlikely to completely disengage from web-based content, it is important to prioritize upskilling users. Therefore, interventions should be developed focusing on improving digital literacy and metacognitive skills, such as the diary-based reflections used in this study.
3. A deeper examination of the perceived benefits of web-based engagement is necessary to ensure that these needs can be met in a safer manner both on the web and offline. In addition, it is crucial to critically evaluate the helpfulness of these perceived benefits, such as the impact of "vicarious living" through observing others self-harm.
4. Web-based industry leaders need to produce more tools that empower individuals to regain control of their

web-based engagement and improve the safety of web-based spaces where self-harm and suicide content is available. This may include changes to the functions of social media, such as providing further control and management options to users over algorithms and hashtags.

Conclusions

A balance between cognitive flexibility and rigidity seems necessary to protect individuals when engaging with self-harm and suicide content on the web. While cognitive flexibility may be helpful in certain situations such as exploring new coping strategies, it can also leave individuals vulnerable to harmful content. On the other hand, cognitive rigidity, or the tendency to repeatedly engage with the same type of content, can lead to

desensitization, potential impairments in functioning, and an increased severity of harm to oneself. Cognitive rigidity can also prevent people from engaging in harmful behaviors and allow them to consistently engage with content that is helpful and positive. Although life events and changes in mental health state could trigger cognitive flexibility resulting in behavior changes, these were unlikely to become permanent unless participants developed skills such as digital efficacy and metacognition that gave them greater control over their behavior. Despite this, even with improved skills for recognizing and managing web-based risks, individuals still perceived that the benefits of web spaces outweighed the costs, making it difficult to fully disengage.

Data Availability

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

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary material including data confidentiality statement, additional tables, and descriptive analysis line graphs.

[[DOCX File, 43 KB - infodemiology_v4i1e47699_app1.docx](#)]

Multimedia Appendix 2

Example safety plan.

[[DOCX File, 64 KB - infodemiology_v4i1e47699_app2.docx](#)]

References

1. Suicide worldwide in 2019: global health estimates. World Health Organization. 2021. URL: <https://www.who.int/publications/i/item/9789240026643> [accessed 2024-02-25]
2. Brown RC, Fischer T, Goldwich DA, Plener PL. "I just finally wanted to belong somewhere"—qualitative analysis of experiences with posting pictures of self-injury on Instagram. *Front. Psychiatry* 2020 Apr 21;11:274 [FREE Full text] [doi: [10.3389/fpsy.2020.00274](https://doi.org/10.3389/fpsy.2020.00274)] [Medline: [32372983](https://pubmed.ncbi.nlm.nih.gov/32372983/)]
3. Lavis A, Winter R. #Online harms or benefits? An ethnographic analysis of the positives and negatives of peer-support around self-harm on social media. *J Child Psychol Psychiatry* 2020 Aug 27;61(8):842-854. [doi: [10.1111/jcpp.13245](https://doi.org/10.1111/jcpp.13245)] [Medline: [32459004](https://pubmed.ncbi.nlm.nih.gov/32459004/)]
4. Marchant A, Hawton K, Stewart A, Montgomery P, Singaravelu V, Lloyd K, et al. Correction: a systematic review of the relationship between internet use, self-harm and suicidal behaviour in young people: the good, the bad and the unknown. *PLoS One* 2018 Mar 1;13(3):e0193937 [FREE Full text] [doi: [10.1371/journal.pone.0193937](https://doi.org/10.1371/journal.pone.0193937)] [Medline: [29494676](https://pubmed.ncbi.nlm.nih.gov/29494676/)]
5. Biddle L, Derges J, Mars B, Heron J, Donovan JL, Potokar J, et al. Suicide and the internet: changes in the accessibility of suicide-related information between 2007 and 2014. *J Affect Disord* 2016 Jan 15;190:370-375. [doi: [10.1016/j.jad.2015.10.028](https://doi.org/10.1016/j.jad.2015.10.028)] [Medline: [26546772](https://pubmed.ncbi.nlm.nih.gov/26546772/)]
6. Miguel EM, Chou T, Golik A, Cornacchio D, Sanchez AL, DeSerisy M, et al. Examining the scope and patterns of deliberate self-injurious cutting content in popular social media. *Depress Anxiety* 2017 Sep 29;34(9):786-793. [doi: [10.1002/da.22668](https://doi.org/10.1002/da.22668)] [Medline: [28661053](https://pubmed.ncbi.nlm.nih.gov/28661053/)]
7. Mars B, Heron J, Biddle L, Donovan JL, Holley R, Piper M, et al. Exposure to, and searching for, information about suicide and self-harm on the internet: prevalence and predictors in a population based cohort of young adults. *J Affect Disord* 2015 Oct 01;185:239-245 [FREE Full text] [doi: [10.1016/j.jad.2015.06.001](https://doi.org/10.1016/j.jad.2015.06.001)] [Medline: [26150198](https://pubmed.ncbi.nlm.nih.gov/26150198/)]
8. Nesi J, Burke TA, Lawrence HR, MacPherson HA, Spirito A, Wolff JC. Online self-injury activities among psychiatrically hospitalized adolescents: prevalence, functions, and perceived consequences. *Res Child Adolesc Psychopathol* 2021 Apr 06;49(4):519-531 [FREE Full text] [doi: [10.1007/s10802-020-00734-4](https://doi.org/10.1007/s10802-020-00734-4)] [Medline: [33404950](https://pubmed.ncbi.nlm.nih.gov/33404950/)]
9. Rodway C, Tham SG, Richards N, Ibrahim S, Turnbull P, Kapur N, et al. Online harms? Suicide-related online experience: a UK-wide case series study of young people who die by suicide. *Psychol Med* 2022 May 19;53(10):4434-4445. [doi: [10.1017/s0033291722001258](https://doi.org/10.1017/s0033291722001258)]

10. Biddle L, Derges J, Goldsmith C, Donovan JL, Gunnell D. Online help for people with suicidal thoughts provided by charities and healthcare organisations: a qualitative study of users' perceptions. *Soc Psychiatry Psychiatr Epidemiol* 2020 Sep 10;55(9):1157-1166 [FREE Full text] [doi: [10.1007/s00127-020-01852-6](https://doi.org/10.1007/s00127-020-01852-6)] [Medline: [32157324](https://pubmed.ncbi.nlm.nih.gov/32157324/)]
11. Tucker IM, Lavis A. Temporalities of mental distress: digital immediacy and the meaning of 'crisis' in online support. *Sociol Health Illn* 2019 Oct 10;41 Suppl 1(S1):132-146. [doi: [10.1111/1467-9566.12943](https://doi.org/10.1111/1467-9566.12943)] [Medline: [31599990](https://pubmed.ncbi.nlm.nih.gov/31599990/)]
12. Mok K, Jorm AF, Pirkis J. Who goes online for suicide-related reasons? *Crisis* 2016 Mar;37(2):112-120. [doi: [10.1027/0227-5910/a000366](https://doi.org/10.1027/0227-5910/a000366)] [Medline: [27232427](https://pubmed.ncbi.nlm.nih.gov/27232427/)]
13. Biddle L, Derges J, Goldsmith C, Donovan JL, Gunnell D. Using the internet for suicide-related purposes: contrasting findings from young people in the community and self-harm patients admitted to hospital. *PLoS One* 2018 May 24;13(5):e0197712 [FREE Full text] [doi: [10.1371/journal.pone.0197712](https://doi.org/10.1371/journal.pone.0197712)] [Medline: [29795637](https://pubmed.ncbi.nlm.nih.gov/29795637/)]
14. Boyce N. Pilots of the future: suicide prevention and the internet. *Lancet* 2010 Dec 04;376(9756):1889-1890. [doi: [10.1016/s0140-6736\(10\)62199-x](https://doi.org/10.1016/s0140-6736(10)62199-x)] [Medline: [21155079](https://pubmed.ncbi.nlm.nih.gov/21155079/)]
15. Brennan CA, Saraiva S, Mitchell E, Melia R, Campbell L, King N, et al. Self-harm and suicidal content online, harmful or helpful? A systematic review of the recent evidence. *J Public Ment Health* 2022 Jan 05;21(1):57-69. [doi: [10.1108/jpmh-09-2021-0118](https://doi.org/10.1108/jpmh-09-2021-0118)]
16. Gunnell D, Derges J, Chang SS, Biddle L. Searching for suicide methods: accessibility of information about helium as a method of suicide on the internet. *Crisis* 2015 Sep;36(5):325-331. [doi: [10.1027/0227-5910/a000326](https://doi.org/10.1027/0227-5910/a000326)] [Medline: [26502782](https://pubmed.ncbi.nlm.nih.gov/26502782/)]
17. Arendt F, Scherr S, Romer D. Effects of exposure to self-harm on social media: Evidence from a two-wave panel study among young adults. *New Media Soc* 2019 May 27;21(11-12):2422-2442. [doi: [10.1177/1461444819850106](https://doi.org/10.1177/1461444819850106)]
18. Picardo J, McKenzie SK, Collings S, Jenkin G. Suicide and self-harm content on instagram: a systematic scoping review. *PLoS One* 2020;15(9):e0238603 [FREE Full text] [doi: [10.1371/journal.pone.0238603](https://doi.org/10.1371/journal.pone.0238603)] [Medline: [32877433](https://pubmed.ncbi.nlm.nih.gov/32877433/)]
19. Eichenberg C, Schott M. An empirical analysis of internet message boards for self-harming behavior. *Arch Suicide Res* 2017 Nov 15;21(4):672-686. [doi: [10.1080/13811118.2016.1259597](https://doi.org/10.1080/13811118.2016.1259597)] [Medline: [27845618](https://pubmed.ncbi.nlm.nih.gov/27845618/)]
20. Harris KM, McLean JP, Sheffield J. Suicidal and online: how do online behaviors inform us of this high-risk population? *Death Stud* 2014 Oct 18;38(6-10):387-394. [doi: [10.1080/07481187.2013.768313](https://doi.org/10.1080/07481187.2013.768313)] [Medline: [24666145](https://pubmed.ncbi.nlm.nih.gov/24666145/)]
21. Scherr S, Reinemann C. First do no harm: cross-sectional and longitudinal evidence for the impact of individual suicidality on the use of online health forums and support groups. *Comput Human Behav* 2016 Aug;61:80-88. [doi: [10.1016/j.chb.2016.03.009](https://doi.org/10.1016/j.chb.2016.03.009)]
22. Till B, Tran US, Voracek M, Niederkrotenthaler T. Beneficial and harmful effects of educative suicide prevention websites: randomised controlled trial exploring Papageno . Werther effects. *Br J Psychiatry* 2017 Aug 02;211(2):109-115. [doi: [10.1192/bjp.bp.115.177394](https://doi.org/10.1192/bjp.bp.115.177394)] [Medline: [28522433](https://pubmed.ncbi.nlm.nih.gov/28522433/)]
23. Cheng Q, Yom-Tov E. Do search engine helpline notices aid in preventing suicide? Analysis of archival data. *J Med Internet Res* 2019 Mar 26;21(3):e12235 [FREE Full text] [doi: [10.2196/12235](https://doi.org/10.2196/12235)] [Medline: [30912753](https://pubmed.ncbi.nlm.nih.gov/30912753/)]
24. Holland J, Thomson R. Gaining a perspective on choice and fate: revisiting critical moments. *Eur Soc* 2009 Jul;11(3):451-469. [doi: [10.1080/14616690902764799](https://doi.org/10.1080/14616690902764799)]
25. Smit D, Vrijzen JN, Broekman T, Groeneweg B, Spijker J. User engagement within an online peer support community (depression connect) and recovery-related changes in empowerment: longitudinal user survey. *JMIR Form Res* 2022 Nov 02;6(11):e39912 [FREE Full text] [doi: [10.2196/39912](https://doi.org/10.2196/39912)] [Medline: [36322110](https://pubmed.ncbi.nlm.nih.gov/36322110/)]
26. Maher CA, Lewis LK, Ferrar K, Marshall S, De Bourdeaudhuij I, Vandelandotte 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](https://doi.org/10.2196/jmir.2952)] [Medline: [24550083](https://pubmed.ncbi.nlm.nih.gov/24550083/)]
27. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. *Qual Health Res* 2016 Nov 10;26(13):1753-1760. [doi: [10.1177/1049732315617444](https://doi.org/10.1177/1049732315617444)] [Medline: [26613970](https://pubmed.ncbi.nlm.nih.gov/26613970/)]
28. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006 May 22;166(10):1092-1097. [doi: [10.1001/archinte.166.10.1092](https://doi.org/10.1001/archinte.166.10.1092)] [Medline: [16717171](https://pubmed.ncbi.nlm.nih.gov/16717171/)]
29. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001 Sep;16(9):606-613 [FREE Full text] [doi: [10.1046/j.1525-1497.2001.016009606.x](https://doi.org/10.1046/j.1525-1497.2001.016009606.x)] [Medline: [11556941](https://pubmed.ncbi.nlm.nih.gov/11556941/)]
30. De Beurs D, Cleare S, Wetherall K, Eschle-Byrne S, Ferguson E, O'Connor DB, et al. Entrapment scale--short-form. *APA PsycTests*. URL: <https://psycnet.apa.org/doiLanding?doi=10.1037%2Ft79989-000> [accessed 2024-02-21]
31. Bass M, Dawkin M, Muncer S, Vigurs S, Bostock J. Validation of Warwick Edinburgh mental well-being scale (WEMWBS) in a population of people using secondary care mental health services. *J Ment Health* 2016 Aug 28;25(4):323-329. [doi: [10.3109/09638237.2015.1124401](https://doi.org/10.3109/09638237.2015.1124401)] [Medline: [26821730](https://pubmed.ncbi.nlm.nih.gov/26821730/)]
32. Grosseohme D, Lipstein E. Analyzing longitudinal qualitative data: the application of trajectory and recurrent cross-sectional approaches. *BMC Res Notes* 2016 Mar 02;9:136 [FREE Full text] [doi: [10.1186/s13104-016-1954-1](https://doi.org/10.1186/s13104-016-1954-1)] [Medline: [26936266](https://pubmed.ncbi.nlm.nih.gov/26936266/)]
33. Dajani DR, Uddin LQ. Demystifying cognitive flexibility: implications for clinical and developmental neuroscience. *Trends Neurosci* 2015 Sep;38(9):571-578 [FREE Full text] [doi: [10.1016/j.tins.2015.07.003](https://doi.org/10.1016/j.tins.2015.07.003)] [Medline: [26343956](https://pubmed.ncbi.nlm.nih.gov/26343956/)]
34. Miranda R, Gallagher M, Bauchner B, Vaysman R, Marroquín B. Cognitive inflexibility as a prospective predictor of suicidal ideation among young adults with a suicide attempt history. *Depress Anxiety* 2012 Mar 06;29(3):180-186. [doi: [10.1002/da.20915](https://doi.org/10.1002/da.20915)] [Medline: [22147587](https://pubmed.ncbi.nlm.nih.gov/22147587/)]

35. Nilsson M, Lundh L, Westrin Å, Westling S. Executive functioning in psychiatric patients with deliberate self-harm, as compared with a psychiatric and a healthy comparison group. *J Clin Exp Neuropsychol* 2021 Apr 05;43(3):225-237. [doi: [10.1080/13803395.2021.1894094](https://doi.org/10.1080/13803395.2021.1894094)] [Medline: [33949907](https://pubmed.ncbi.nlm.nih.gov/33949907/)]
36. Park Y, Ammerman BA. For better or worse?: the role of cognitive flexibility in the association between nonsuicidal self-injury and suicide attempt. *J Psychiatr Res* 2023 Feb;158:157-164. [doi: [10.1016/j.jpsychires.2022.12.040](https://doi.org/10.1016/j.jpsychires.2022.12.040)] [Medline: [36586214](https://pubmed.ncbi.nlm.nih.gov/36586214/)]
37. Cañas JJ, Fajardo I, Salmeron L. Cognitive flexibility. In: Karwowski W, editor. *International Encyclopedia of Ergonomics and Human Factors*. Boca Raton, FL: CRC Press; 2006:297-301.
38. Miranda R, Valderrama J, Tsypes A, Gadol E, Gallagher M. Cognitive inflexibility and suicidal ideation: mediating role of brooding and hopelessness. *Psychiatry Res* 2013 Nov 30;210(1):174-181 [FREE Full text] [doi: [10.1016/j.psychres.2013.02.033](https://doi.org/10.1016/j.psychres.2013.02.033)] [Medline: [23528518](https://pubmed.ncbi.nlm.nih.gov/23528518/)]
39. Kakoschke N, Aarts E, Verdejo-García A. The cognitive drivers of compulsive eating behavior. *Front Behav Neurosci* 2018 Jan 17;12:338 [FREE Full text] [doi: [10.3389/fnbeh.2018.00338](https://doi.org/10.3389/fnbeh.2018.00338)] [Medline: [30705625](https://pubmed.ncbi.nlm.nih.gov/30705625/)]
40. Wu M, Brockmeyer T, Hartmann M, Skunde M, Herzog W, Friederich HC. Set-shifting ability across the spectrum of eating disorders and in overweight and obesity: a systematic review and meta-analysis. *Psychol Med* 2014 Dec;44(16):3365-3385. [doi: [10.1017/S0033291714000294](https://doi.org/10.1017/S0033291714000294)] [Medline: [25066267](https://pubmed.ncbi.nlm.nih.gov/25066267/)]
41. Liu C, Rotaru K, Lee RS, Tiego J, Suo C, Yücel M, et al. Distress-driven impulsivity interacts with cognitive inflexibility to determine addiction-like eating. *J Behav Addict* 2021 Apr 27;10(3):534-539 [FREE Full text] [doi: [10.1556/2006.2021.00027](https://doi.org/10.1556/2006.2021.00027)] [Medline: [33909594](https://pubmed.ncbi.nlm.nih.gov/33909594/)]
42. Aditi B, Arunjyoti B. Addiction severity, social functioning, and life satisfaction of patients diagnosed with substance use disorders. *Indian J Psy Nsg* 2018;15(2):13. [doi: [10.4103/2231-1505.262434](https://doi.org/10.4103/2231-1505.262434)]
43. Dennis JP, Vander Wal JS. The cognitive flexibility inventory: instrument development and estimates of reliability and validity. *Cogn Ther Res* 2009 Nov 11;34(3):241-253. [doi: [10.1007/s10608-009-9276-4](https://doi.org/10.1007/s10608-009-9276-4)]
44. Eshet-Alkali Y, Amichai-Hamburger Y. Experiments in digital literacy. *Cyberpsychol Behav* 2004 Aug;7(4):421-429. [doi: [10.1089/cpb.2004.7.421](https://doi.org/10.1089/cpb.2004.7.421)] [Medline: [15331029](https://pubmed.ncbi.nlm.nih.gov/15331029/)]
45. Lai YJ, Tan HC, Wang CT, Wu WC, Wang LY, Shen YC. Difference in cognitive flexibility between passive and active suicidal ideation in patients with depression. *Neuropsychiatry* 2018;08(04):1182-1185. [doi: [10.4172/neuropsychiatry.1000446](https://doi.org/10.4172/neuropsychiatry.1000446)]
46. Eshet-Alkalai Y, Chajut E. You can teach old dogs new tricks: the factors that affect changes over time in digital literacy. *J Inf Technol Educ Res* 2010;9:173-181. [doi: [10.28945/1186](https://doi.org/10.28945/1186)]
47. Martinez ME. What is metacognition? *Phi Delta Kappan* 2016 Jun 24;87(9):696-699. [doi: [10.1177/003172170608700916](https://doi.org/10.1177/003172170608700916)]
48. Merrett F. Reflections on the Hawthorne effect. *Educ Psychol* 2006 Jan;26(1):143-146. [doi: [10.1080/01443410500341080](https://doi.org/10.1080/01443410500341080)]
49. Derks YP, Klaassen R, Westerhof GJ, Bohlmeijer ET, Noordzij ML. Development of an ambulatory biofeedback app to enhance emotional awareness in patients with borderline personality disorder: multicycle usability testing study. *JMIR Mhealth Uhealth* 2019 Oct 15;7(10):e13479 [FREE Full text] [doi: [10.2196/13479](https://doi.org/10.2196/13479)] [Medline: [31617851](https://pubmed.ncbi.nlm.nih.gov/31617851/)]
50. Kurt M, Kurt S. Improving design understandings and skills through enhanced metacognition: reflective design journals. *Int J Art Design Ed* 2016 Jul 09;36(2):226-238. [doi: [10.1111/jade.12094](https://doi.org/10.1111/jade.12094)]
51. Abdelraheem M, McAloon J, Shand F. Mediating and moderating variables in the prediction of self-harm in young people: a systematic review of prospective longitudinal studies. *J Affect Disord* 2019 Mar 01;246:14-28. [doi: [10.1016/j.jad.2018.12.004](https://doi.org/10.1016/j.jad.2018.12.004)] [Medline: [30572208](https://pubmed.ncbi.nlm.nih.gov/30572208/)]
52. Vogl S, Zartler U. Interviewing adolescents through time: balancing continuity and flexibility in a qualitative longitudinal study. *Longit Life Course Stud* 2021;12:83-97. [doi: [10.1332/175795920x15986464938219](https://doi.org/10.1332/175795920x15986464938219)]

Edited by E Lee; submitted 04.04.23; peer-reviewed by T Officer, R Plackett; comments to author 20.09.23; revised version received 10.10.23; accepted 15.02.24; published 28.03.24.

Please cite as:

Haime Z, Kennedy L, Grace L, Cohen R, Derges J, Biddle L

The Journey of Engaging With Web-Based Self-Harm and Suicide Content: Longitudinal Qualitative Study

JMIR Infodemiology 2024;4:e47699

URL: <https://infodemiology.jmir.org/2024/1/e47699>

doi: [10.2196/47699](https://doi.org/10.2196/47699)

PMID: [38546718](https://pubmed.ncbi.nlm.nih.gov/38546718/)

©Zoë Haime, Laura Kennedy, Lydia Grace, Rachel Cohen, Jane Derges, Lucy Biddle. Originally published in *JMIR Infodemiology* (<https://infodemiology.jmir.org>), 28.03.2024. 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.

Original Paper

Exploring the Impact of the COVID-19 Pandemic on Twitter in Japan: Qualitative Analysis of Disrupted Plans and Consequences

Masaru Kamba¹, PhD; Wan Jou She¹, PhD; Kiki Ferawati¹, MStat; Shoko Wakamiya¹, PhD; Eiji Aramaki¹, PhD

Division of Information Science, Graduate School of Science and Technology, Nara Institute of Science and Technology, Ikoma, Japan

Corresponding Author:

Eiji Aramaki, PhD

Division of Information Science

Graduate School of Science and Technology

Nara Institute of Science and Technology

8916-5 Takayama-cho

Ikoma, 630-0192

Japan

Phone: 81 0743 72 5250

Email: aramaki@is.naist.jp

Abstract

Background: Despite being a pandemic, the impact of the spread of COVID-19 extends beyond public health, influencing areas such as the economy, education, work style, and social relationships. Research studies that document public opinions and estimate the long-term potential impact after the pandemic can be of value to the field.

Objective: This study aims to uncover and track concerns in Japan throughout the COVID-19 pandemic by analyzing Japanese individuals' self-disclosure of disruptions to their life plans on social media. This approach offers alternative evidence for identifying concerns that may require further attention for individuals living in Japan.

Methods: We extracted 300,778 tweets using the query phrase *Corona-no-sei* ("due to COVID-19," "because of COVID-19," or "considering COVID-19"), enabling us to identify the activities and life plans disrupted by the pandemic. The correlation between the number of tweets and COVID-19 cases was analyzed, along with an examination of frequently co-occurring words.

Results: The top 20 nouns, verbs, and noun plus verb pairs co-occurring with *Corona no-sei* were extracted. The top 5 keywords were *graduation ceremony*, *cancel*, *school*, *work*, and *event*. The top 5 verbs were *disappear*, *go*, *rest*, *can go*, and *end*. Our findings indicate that education emerged as the top concern when the Japanese government announced the first state of emergency. We also observed a sudden surge in anxiety about material shortages such as toilet paper. As the pandemic persisted and more states of emergency were declared, we noticed a shift toward long-term concerns, including careers, social relationships, and education.

Conclusions: Our study incorporated machine learning techniques for disease monitoring through the use of tweet data, allowing the identification of underlying concerns (eg, disrupted education and work conditions) throughout the 3 stages of Japanese government emergency announcements. The comparison with COVID-19 case numbers provides valuable insights into the short- and long-term societal impacts, emphasizing the importance of considering citizens' perspectives in policy-making and supporting those affected by the pandemic, particularly in the context of Japanese government decision-making.

(*JMIR Infodemiology* 2024;4:e49699) doi:[10.2196/49699](https://doi.org/10.2196/49699)

KEYWORDS

COVID-19; natural language processing; NLP; Twitter; disrupted plans; concerns

Introduction

Background

The spread of COVID-19 has become a global pandemic, significantly affecting social and economic sectors worldwide [1]. In the early stages of the pandemic, health authorities

recommended social distancing to control the spread of the virus, reduce cases, and avoid overwhelming health care facilities [2-4]. Each country had its own strategy for dealing with COVID-19. A survey conducted across 6 countries illustrated the public's perception of measures taken in response to COVID-19 [5]. Other surveys have been conducted in the United Kingdom and European countries to aid interdisciplinary

research on public health, particularly regarding COVID-19 [6]. Different results were observed owing to social distancing policies, which affected several aspects of life, including economic activities [7] and consumer behavior, such as drops in mobility [8]. Concerns about cybersecurity risks were also raised, as companies might not have been prepared for adequate work-from-home options for employees [9]. The association between implementing some mitigation policies in response to COVID-19 and the outcomes regarding public mobility were noted [10], one of which was also observed in Japan.

After the government confirmed the first COVID-19 case in Japan on January 16, 2020, the number of cases quickly escalated within 3 months, leading to the declaration of a state of emergency to prevent the further spread of the infection. This measure significantly impacted the daily routines and social lives of Japanese residents, forcing them to refrain from going out, close schools, work from home, and be restricted from visiting crowded locations such as department stores and movie theaters. The first state of emergency effectively reduced the number of COVID-19 cases [11], albeit at a high cost to public mental well-being, education quality, and the economy. Furthermore, the number of cases quickly bounced back, peaking at 1762 new daily cases after the state of emergency was lifted, an increase from the peak of 701 new daily cases during the first wave [11]. These numbers suggest that the government was confronted with the dilemma of mitigating the social and economic impact of the lockdown and stopping the spread of COVID-19 [12]. Due to the fluctuations in COVID-19 cases, the government declared other states of emergency, recognizing the profound and deeply rooted impact the COVID-19 lockdown could have on societal and economic levels.

There have been various investigations into the states of emergency. For instance, studies have predicted SARS-CoV-2 infections using state-space models [13] and examined their impact on mental health [14]. In the aspect of mobility, studies have shown the suppression of social activities of the masses [15]. The tourism industry was among the hardest hit sectors, and the arrival of visitors decreased by 93% by March 2020 [16]. Statistics also show that Japan's gross domestic product in 2020 decreased by 4.28%, indicating a substantial impact on the economy [17]. Interestingly, the unemployment rate only slightly increased to 2.8% in 2020, but started declining by 2022 (2.64%), following the gradual recovery of the gross domestic product (2.14% growth by 2021 and 1.03% by 2022) [17,18]. This trend of recovery indicates the strong resilience of the Japanese economy.

Furthermore, it also changed people's behavior, such as following the advisory to stay at home, as confirmed by cell phone location data [19-21]. Such large-scale societal and behavioral changes warrant further investigation through various means to offer a chance to monitor and reflect the short- and long-term impacts of COVID-19 in the future.

Literature Review

The disruption caused by pandemic-related restrictions may be seen as a failure to perform planned activities, but detecting such disruptions was challenging. For example, it is difficult to

obtain behavioral data on trips that individuals could not take or events they could not attend owing to the restriction. Social media, which people use to share their activities, proved to be a great source of information in such cases. Twitter (currently X) is a widely used social media platform in many countries and has a sufficiently large population for social data analysis in health care contexts [22,23]. Japan has a particularly high population density of Twitter users, even when compared to the major countries that use Twitter, such as the United States. Furthermore, owing to language exclusivity, it is easier to filter comments related to Japanese society using Japanese keywords [24]. Twitter has also been frequently used to help summarize peoples' responses about the pandemic and its measures, showing the challenges experienced throughout [25]. Prior studies in Korea and Japan used Twitter to determine public opinion, showing popular words during the pandemic [26]. Because people actively share their daily lives on Twitter, the site has the potential to be a data source for investigating the impact of restrictions on the public. Using Twitter as a resource, this study aims to explore and visualize plans disrupted in Japan due to COVID-19 pandemic measures.

There are many studies on COVID-19 that investigate social media platforms, such as Twitter. Chen et al [27] investigated the levels of anxiety during the COVID-19 pandemic. The adverse effects on the mental health of the public was also one of the impacts of the pandemic, as explained in the research by Li et al [28], who analyzed COVID-19-related tweets into different emotions and investigated the mental health aspects and how they recovered from the COVID-19 crisis. Lyu et al [29] investigated the topics and sentiments in public COVID-19 vaccine-related discussions, whereas Krittanawong et al [30] investigated misinformation dissemination related to COVID-19 on Twitter. Aside from studies focused on the pandemic itself, COVID-19 vaccines have also been highly researched topics on Twitter. Ansari and Khan assessed public responses through sentiment analysis of COVID-19 vaccines using Twitter, revealing an overall negative tone in the tweets [31]. Ferawati et al [32] explored how Twitter reported vaccine-related side effects by comparing the side effects of 2 types of messenger RNA vaccines developed by Pfizer and Moderna in Japan and Indonesia, respectively. Gao et al [33] examined COVID-19 concerns in each Japanese prefecture, and Uehara et al [34] investigated the attitudes toward vaccines or vaccination during the COVID-19 pandemic in different Japanese prefectures using search queries from Yahoo! JAPAN. Our study adopts a unique approach to examine how the COVID-19 pandemic has disrupted everyday activities. Our main focus is on understanding the direct impact of the pandemic on society through the observation of expressions, life disruptions, and plans.

For research on citizen feedback, Ishida et al [35] proposed a method that uses social media data. They implemented a multitask learning framework to estimate the associated viewpoints using bidirectional encoder representations from the transformer model. However, this method requires considerable effort to label the data. This study uses search queries and validates word co-occurrence to infer the themes of topics

discussed during the COVID-19 pandemic in Japan, proposing an efficient and low-resource method for social media analysis.

Objectives and Approach

We aimed to report on the impact of COVID-19 on Japanese society by analyzing public opinions extracted from social media data. Specifically, we focused on the popular term *Corona no-sei* (in Japanese コロナのせい, meaning “due to COVID-19,” “because of COVID-19,” or “considering COVID-19”), which clearly conveyed complaints or concerns about life event disruptions caused by the COVID-19 pandemic. Our study used 2 types of data: the daily COVID-19 case count and Japanese tweets containing the Japanese phrase *Corona no-sei* posted on Twitter between February 1, 2020, and November 30, 2021. We analyzed the trends in the number of tweets and COVID-19 cases to quantitatively explore their relationship and the words frequently used in the tweets to qualitatively explore social needs in the first 2 years of the COVID-19 pandemic.

In conclusion, we critically compared our findings with those identified in other similar studies to provide an alternative evidence base for the impact of COVID-19 on Japanese society.

Methods

COVID-19 Cases

To track the daily rise in COVID-19 cases, we gathered the number of new positive cases in Japan by manually downloading data from a dedicated COVID-19 site maintained by the NHK, Japan’s national broadcaster [36]. Our aim was to investigate the correlation between the number of positive cases and the volume of tweets. A total of 1,726,943 COVID-19–positive cases were recorded between February 1, 2020, and November 30, 2021.

Tweets and Keywords Extraction

Another data set for this study includes 300,778 tweets containing the Japanese phrase *Corona no-sei* during the same period as the recorded COVID-19 cases (between February 1, 2020, and November 30, 2021). We chose this period because by the end of January 2020, the Japanese government had officially established the Japan Anti-Coronavirus National Task Force to actively address the pandemic. In addition, we aimed to include the maximum possible data until the initiation of this study in mid-November 2021. Furthermore, this period also included 3 emergency announcements by the Japanese government, making it a representative period for studying the impact of COVID-19 on Japanese society.

We counted the number of tweets per month and found that the maximum number of tweets was 517,688 in April 2020; the minimum number of tweets was 24,625 in November 2021; and the average number of tweets was 136,717.6. The *Corona no-sei* phrase is frequently used by the public in social media and everyday conversation to express the (often negative) feelings when Japanese individuals’ activities or life plans were interrupted by the COVID-19 outbreak. Although there are several expressions synonymous with *Corona no-sei* (eg, “because of the new coronavirus” and “because of COVID-19”), we chose *Corona no-sei* as a casual expression used by the

public in social media and colloquial speech. The tweet data were provided by the NTT DATA Corporation, which has a real-time backup of Japanese firehose data from X Corporation (formerly known as Twitter). Data access was granted to a few collaborative research institutes, including the University of Tokyo, and one of the authors was granted permission to use the self-adaptive classification system to extract the data and keywords [37].

Although applying a clustering approach for topic modeling can be useful in grasping the topics discussed in the tweets, it does not apply to our context wherein we were targeting *COVID-19* as the main subject and aiming to identify the co-occurrence of events. Instead, we extracted co-occurrence nouns and verbs from the obtained *Corona no-sei* tweets by applying dependency analysis implemented in the system developed by Yoshinaga et al [37–39]. We used the base-phrase chunker to extract all tweets containing the *Corona no-sei* keyword (“keyword” is *bunsetsu* in Japanese). The built-in classifier then extracted the relevant verbs, nouns, and verb-noun-pairs for users based on the nonstack dependency parser, which achieved 99.01% accuracy in base-phrase chunking and 92.23% accuracy in dependency parsing [37]. Researchers who did not use the system and database maintained by the University of Tokyo could use the same tool published by the laboratory Pecco and DepP [37–40]. To avoid overinterpretation, we omitted tweets that described a disruption of plans but did not include COVID-19–related keywords.

Analysis of the Keywords and its Correlation to the COVID-19 Pandemic Trends

The contexts following *Corona no-sei*, which indicate a high level of negative concern about COVID-19, frequently contain verbs in the negative form and nouns associated with them. By aggregating these nouns and verbs, we extracted information on the restrictions imposed and the events or plans canceled owing to the COVID-19 pandemic. This information enabled us to capture the potential social and psychological impact of disrupted life plans. Note that, by events or plans, we refer to the specific *type of occurrences* (eg, university entrance exam) rather than a certain event (eg, a pop singer’s concert in 2019). The frequency of nouns and verbs in tweets containing *Corona no-sei* was counted to identify the restrictions placed on people’s lives.

To investigate the correlation between tweet volumes and COVID-19 cases, we constructed transition diagrams for each. In addition, Pearson correlation coefficients were also calculated. Next, we examined the nouns and verbs co-occurring with *Corona no-sei* over the entire study period and specifically on the day with the highest tweet activity.

The cross-validation of the keywords and tweet contents was performed by randomly extracting 20 tweets from the top 5 verb and noun pairs and other keyword pairs that were deemed worthy of discussion by the researchers. The tweet contents were further annotated to ensure that they were aligned with the researchers’ interpretations of keywords. We then discussed the themes extracted by analyzing and cross-validating the themes and noteworthy keywords.

Ethical Considerations

This study used publicly available data and did not handle identifiable private information, meaning that it was exempt from Institutional Review Board approval according to the Ethical Guidelines for Research of the Japanese National Government [41]. The NTT DATA Corporation obtained tweets according to Twitter terms of service and approved the use of the data for this study.

Results

Figure 1 shows the time trend of *Corona no-sei* tweets (blue line) compared to the trend of positive cases (red line). There were 3 states of emergency announcements within our targeted period between February 1, 2020, and November 30, 2021, which are highlighted in gray in Figure 1. The number of areas under the state of emergency is indicated by the bar graph in the upper part of the figure because the target areas were changed during each state of emergency. The periods during which the states of emergency were imposed roughly corresponded to an increase in case numbers. Interestingly, the announcement of a state of emergency was highly effective in suppressing the number of cases. Regarding the spike caused by the Tokyo Olympics (which took place between July 23, 2021, and August 8, 2021), the case number quickly dropped to below 5000 per day within 3 months.

As the blue line indicates, the *Corona no-sei* tweets peaked in March 2020, roughly before the first state of emergency was

announced and reached the second highest number when the first state of emergency was imposed. After the first announcement of the state of emergency, the number of tweets using *Corona no-sei* showed a downward trend until the end of our data collection period. There were a few instances of small increases in *Corona no-sei* tweets before the second and third states of emergency announcements, but overall, the number of reported plan disruptions never reached the level observed before the first state of emergency announcement. The scatter plot for case numbers and the numbers of *Corona no-sei* tweets is shown in Figure 2, with Pearson correlation coefficients of 0.86, 0.93, and 0.61, respectively, for the first, second, and third states of emergency.

When compared with the number of the *Corona no-sei* tweets during the entire period, the correlation between COVID-19 daily cases and the *Corona no-sei* was not very evident. We were able to observe a slight increase of *Corona no-sei* tweets before the case number started rising, but the extent of increase in case numbers was disproportional to the extent of increase in *Corona no-sei* tweets. Even though the number of cases peaked in September 2021 during the third state of emergency, there was only a slight increase in *Corona no-sei* tweets compared to the high number of complaints at the very beginning of the COVID-19 pandemic. This indicates that Japanese residents might have adapted to the restrictions or disruptions caused by the COVID-19 pandemic lockdown.

Figure 1. Trends in the number of *Corona no-sei* tweets and the number of patients with a COVID-19–positive result. The blue line indicates the number of *Corona no-sei* tweets, and the red line indicates the number of positive COVID-19 cases.

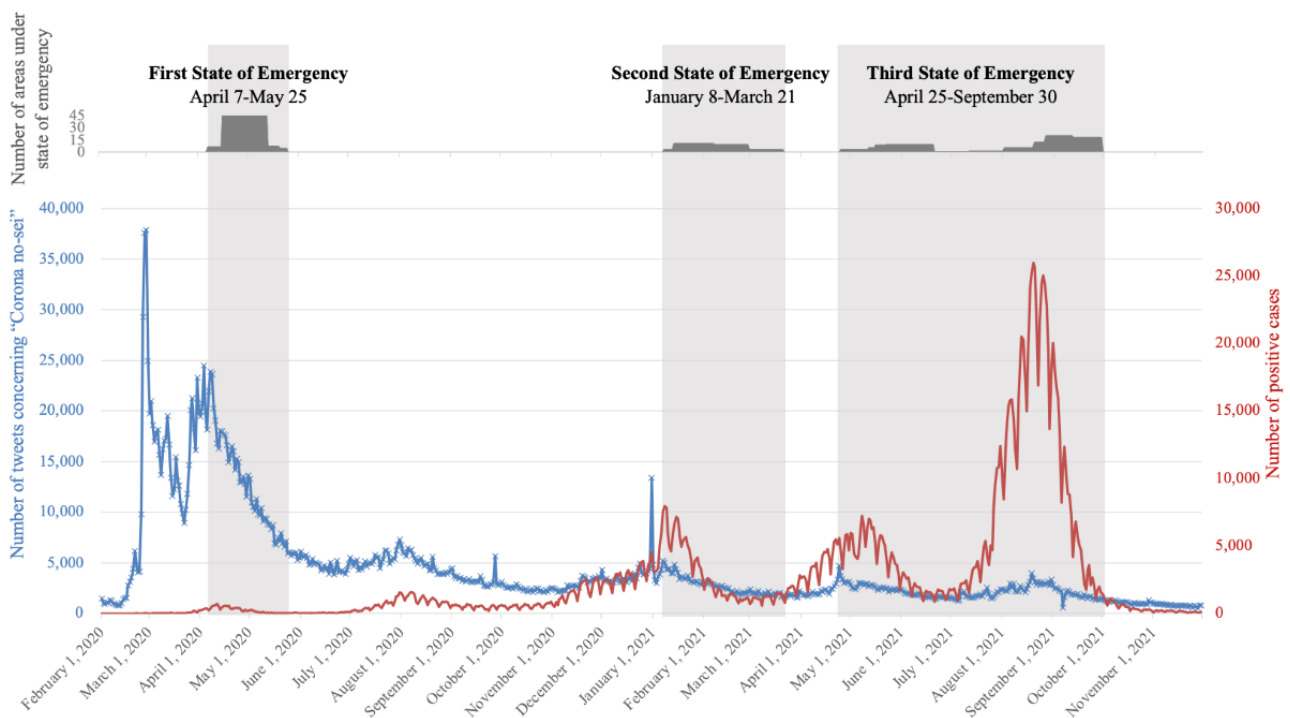
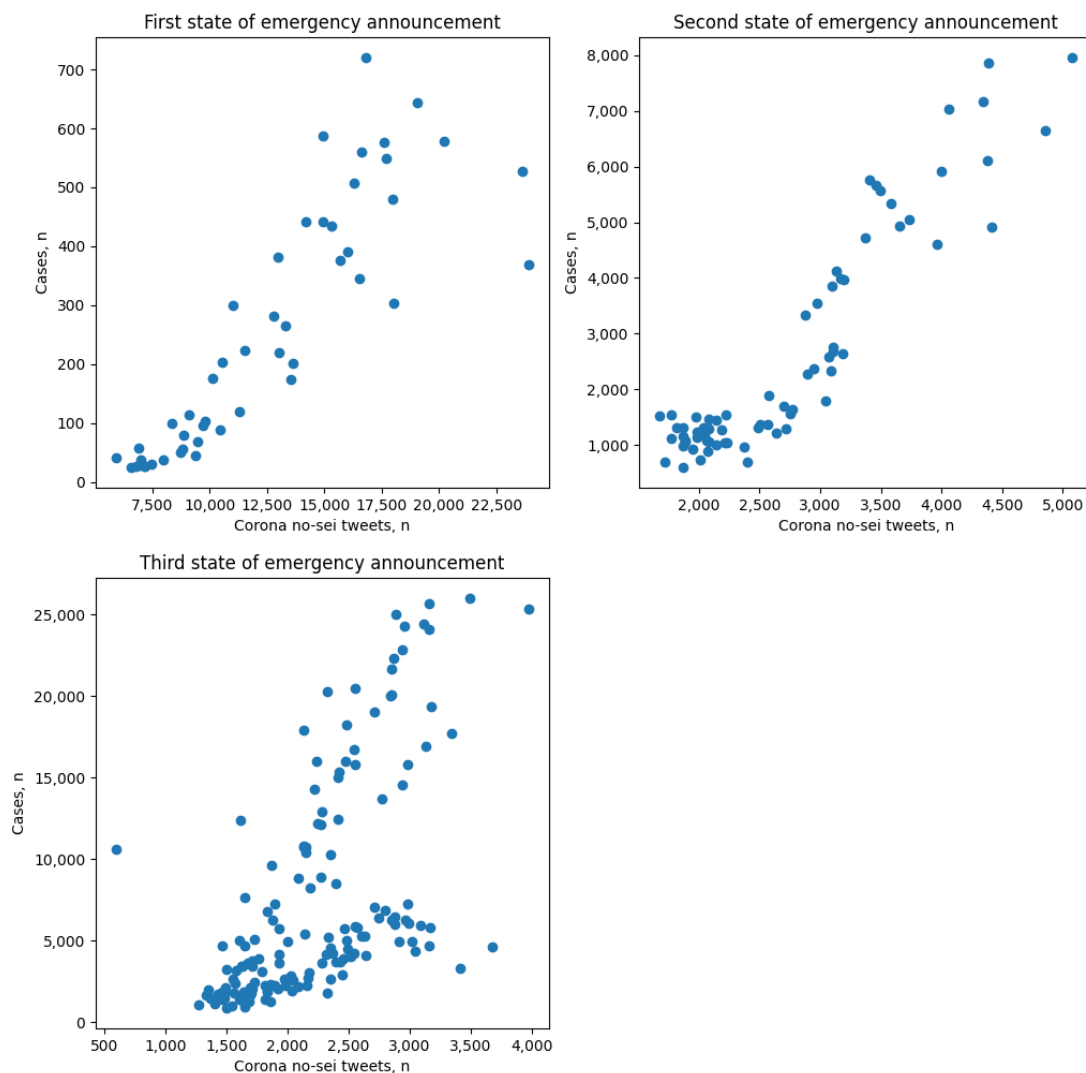


Figure 2. Scatter plot for COVID-19 case numbers and the number of *Corona no-sei* tweets for the first, second, and third state of emergency announcements.



We further investigated the nouns and verbs in the tweets that we sampled. Tables 1 and 2 show the number of tweets for the top 20 nouns and verbs tweeted on February 28, 2020, when the tweet number reached the highest level. Tables 3 and 4 show the top 20 words (nouns and verbs) that co-occurred with *Corona no-sei* tweets in descending order to highlight the most disrupted activities or plans during our data aggregation period. For nouns, here, *Corona* was excluded because it was a word included in the query and was clearly the most frequently detected. For nouns, the top 5 most frequently mentioned words were *work*, *abort*, *home*, *live*, and *friends* after excluding the words that indicate the grammatical tense. These keywords indicated that, over a longer period, Japanese individuals started developing concerns over their disrupted work and social life. For verbs, *go* was the most frequent, but in the actual tweets, it was sometimes used in the negative, and in the context, the verb was unlikely to be used in the affirmative, so the verb was likely used to indicate they *cannot go* even if it is in the affirmative in this paper (Textbox 1). Hence, the top 5 mean *go*, *can go* <negation>, *look*, *meet* <negation>, and *get out*. The results show that there are restrictions on the actions of going, seeing, and meeting as verbs. Compared with the single-day result on February 28, 2020, the concern about work appeared as the top

5 in Tables 1 and 2, suggesting that Japanese individuals placed clear emphasis on their work routines. In addition, the desire for live concerts increased over the long run, making *live concert* the fourth most frequently mentioned keyword in Table 1. Coincidentally, concerns related to friends and missing opportunities to meet up were also observed in both tables, showing the disruption of social relationships and recreational occasions. Both studies indicated that people regarded the COVID-19 pandemic as the main cause of their disrupted plans to hang out with friends or attend large public events. In addition to activities, the keyword finding also reflected the concern of resource shortage, such as toilet paper, masks, and even money, which were critical in supporting daily lives or normal health care practices.

Because the keywords indicated both long-term and short-term concerns, we cross-validated the tweet contents by selecting keyword pairs based on the top 5 keywords related to long-term concerns and those related to short-term anxiety on material shortage. A total of 160 tweets were randomly sampled based on the following keyword pairs (20 tweets each): ライブ+行く / 行けない (*live concert+go/go<negation>*); 家+行く / 行けない (*home+go/go<negation>*); 友達+行く / 行けない

(*friends+go/go<negation>*); 中止+行く / 行けない (toilet paper+vanish); and マスク+なくなる (*mask+vanish*).
 (*cancel+go/go<negation>*); 友達+会う / 会える One of the authors annotated the tweets according to the themes
 (*friends+meet/can meet*); 仕事+行く / 行けない reflected by the keywords. Key findings are discussed in the
 (*work+go/go<negation>*); トイレ+トイレットペーパー+なくなる following section.

Table 1. The number of tweets with co-occurring nouns on February 28, 2020.

Noun	Tweets, n (%)
graduation ceremony	2154 (14.55)
cancel	1813 (12.25)
school	1498 (10.12)
work	1210 (8.17)
event	775 (5.24)
live concert	694 (4.69)
toilet paper	667 (4.51)
part-time job	639 (4.32)
school holiday	636 (4.3)
friends	616 (4.16)
disney	605 (4.09)
postponement	562 (3.8)
home	553 (3.74)
mask	531 (3.59)
new corona	330 (2.23)
test	328 (2.22)
tissue	321 (2.17)
hoax	318 (2.15)
company	304 (2.05)
next month	250 (1.69)

Table 2. The number of tweets with co-occurring verbs on February 28, 2020.

Verb	Tweets, n (%)
disappear	3265 (20.15)
go	1835 (11.33)
rest	1169 (7.22)
can go <negation>	1044 (6.44)
end	862 (5.32)
go out	754 (4.65)
look	734 (4.53)
make effort	711 (4.39)
buy	602 (3.72)
cry	576 (3.56)
vanish	556 (3.43)
can meet <negation>	524 (3.23)
play	505 (3.12)
spring rest	501 (3.09)
dies	499 (3.08)
be told	491 (3.03)
come	441 (2.72)
think	436 (2.69)
return	378 (2.33)
crush	319 (1.97)

Table 3. Noun words co-occurring with *Corona no-sei* in descending order.

Order	Noun
1	work
2	cancel
3	home
4	live concert
5	friends
6	event
7	postponement
8	stress
9	school
10	part-time job
11	company
12	new corona
13	mask
14	graduation ceremony
15	hospital
16	one person
17	opportunity
18	university
19	family
20	money

Table 4. Verb words co-occurring with *Corona no-sei* in descending order.

Order	Verb
1	go
2	can go <negation>
3	look
4	meet <negation>
5	get out
6	lose
7	make effort
8	increase
9	buy
10	lose
11	end
12	come
13	think
14	meet
15	rest
16	can go
17	decrease
18	be told
19	meet
20	play

Textbox 1. Examples of tweets posted on Twitter (Japanese tweets were translated into English).

Verb and example

- Go: “Due to COVID-19, the day I've been looking forward to going out with the guy I love has been postponed... I can't help it now and I'll accept it, but I was looking forward to it.”
- Meet: “It doesn't feel like April at all due to COVID-19, but I can't wait for it to end so that we can all meet, eat, and shop together comfortably. Six years already... I want to quit my job lol.”

Discussion

Principal Findings

Our findings revealed that the COVID-19 pandemic significantly disrupted daily routines in Japan, particularly in terms of work, education, social activities, and material shortages (with regard to the temporary spike of anxiety). The findings from our study correspond with numerous studies conducted in diverse countries, highlighting the extensive impact of the COVID-19 pandemic on social life, economy, public mental health, and education [5]. In this section, we discuss key findings across a temporal spectrum, focusing on 4 crucial aspects: disruption of work routines, public anxiety stemming from perceived resource shortages, concerns regarding social relationships, and interference with the curriculum.

Top Concerns

The impact of the COVID-19 pandemic on the labor market in Japan is unequivocal, mirroring the challenges faced by

numerous countries. The pandemic necessitated a shift in work dynamics with the unintended pilot of remote collaboration. Notably, certain categories of Japanese workers, contingent on their employment contracts, exhibited heightened susceptibility to these alterations in work patterns. In our findings, the keyword *work* demonstrated associations with *part-time*, *abort*, and *money*, indicating that individuals expressing concerns about their work conditions may grapple with job uncertainty, stemming either from the part-time nature of their employment or an abrupt reduction in income. This discovery aligns seamlessly with prior research examining the repercussions of the COVID-19 pandemic on Japan's labor sector. As described by Kikuchi et al [42] in their study, individuals in contingent employment, along with women and those with lower income, were notably susceptible. The shift toward teleworking and the accompanying uncertainty about long-term income during the COVID-19 pandemic had a disproportionately adverse impact on these specific demographic groups [42]. Fukai et al [43] endorsed these findings through extensive government statistical analysis. According to their research, Japanese individuals

employed part-time in service industries or compelled to take leave or face job loss following the declaration of a state of emergency were identified as particularly vulnerable groups significantly affected by the COVID-19 pandemic [43]. Although the use of part-time or contingent workers has traditionally been a standard practice for Japanese companies seeking to optimize budget and resource allocation, the advent of the COVID-19 pandemic has pushed issues related to *work* to the forefront of public concern. Researchers caution that this could potentially exacerbate inequality for susceptible individuals unless actively addressed by government support [44].

In summary, our findings provide substantial evidence for concerns among Japanese internet users regarding job disruption, employment disparities, and inadequate financial resilience. Failing to address these issues during multiple states of emergency, the Japanese government risks compromising the equality within Japan's labor market significantly. Interestingly, a study conducted by Chen et al [27], who sampled 6535 Reddit posts, identified strikingly identical subjects that propelled nationwide anxiety in the United States. Notably, concerns about career, finance, and the future were prevalent. However, our research suggests that health and death concerns were not as prominent in Japan, as observed in the study by Chen et al [27]. We hypothesized that the emphasis on collectivism and harmony in Japanese society could shape individuals' concerns during crises (particularly in the case of a national crisis). For example, apprehensions about not being perceived as "useful" or causing "inconvenience" to others, possibly even relying on government subsidies, were more pronounced than concerns related to health and mortality.

Sudden and Perhaps Excessive Anxiety About Material Shortage

The scarcity of certain items, including toilet paper, masks, and tissues, as outlined in Table 1, emerged as a significant issue in Japan. Our findings closely parallel earlier Twitter studies investigating hoarding behaviors, particularly concerning toilet paper [45]. Although initially observed in the United States, panic buying for household goods rapidly became a global phenomenon. Notably, toilet paper has emerged as a frequently hoarded item, often signaling a surge in demand during natural disasters [46,47]. Although the act of stockpiling toilet paper may seem irrational and has been widely ridiculed on social media, the adverse effects of bulk purchasing have not been as severe. Social scientists may view this behavior as a coping mechanism during a natural disaster [48]. Contrary to the commonly perceived overhoarding of toilet paper, the mask shortage was deemed a more severe public health crisis and a direct threat to well-being. A 2020 agent-based simulation conducted by Tatapudi et al [48] illustrated that universal mask use could potentially reduce infections by 20% [49]. At the time of the study, the total number of people infected by COVID-19 was 541 million, indicating that implementing universal mask use could potentially spare 108 million cases. Numerous studies have indicated a negative correlation between mask use and the COVID-19 infection rate [44,50].

However, the situation in Japan presents a slightly different scenario. The Japanese government faced criticism for a perceived slow response to the awareness of mask shortages, as the pandemic was considered relatively "under control" in its early stages. As the mask crisis unfolded, many Japanese citizens became concerned about their reliance on masks manufactured abroad, prompting the government to take actions to boost domestic mask production. Unfortunately, heightened anxiety also led to the "Abenomasks" incident, wherein the government faced backlash for stockpiling over 82 million unused masks [51]. A crucial lesson learned from this incident is that although social media serves as a critical channel for the dissemination of news and raising public awareness, the emotional contagion and overpromotion of a particular disaster can backfire, impeding the rational coping mechanisms of citizens and the decision-making processes of the government. Our findings, along with those of numerous other studies, indicate that further efforts are needed to develop effective protocols for addressing the widely contagious anxiety stemming from the dissemination of information about natural disasters on social media.

Concerns About Social Relationships

Keywords pertaining to relationships, social life, and collective events were prevalent in our analysis. For instance, the top 20 frequently occurring nouns associated with *Corona no-sei* included friends, family, live, events, and one person. The most frequent verbs in the context of *Corona no-sei* were *go*, *can go* <negation>, *meet* <negation>, *buy*, *meet*, *can go*, and *play*. The example in Table 4 illustrates how Japanese individuals linked *go* and *meet* to their social events. While it may appear that many tweets express concerns about social relationships, these keywords actually reflect people venting their frustration about being unable to meet and engage in activities together, rather than indicating an actual loss of relationships. Interestingly, a study by Goodwin and Takahashi [52] also yielded similar findings. Most Japanese respondents in their survey gauging perceptions of relationship quality during the COVID-19 pandemic indicated that there were no discernible changes in their perceived relationship quality. Only a few reported that their trust and relationship with communities had declined compared to the prepandemic era [52]. There was also a report indicating that students, due to reduced communication with friends, face an increased risk of mental health problems [53].

These findings suggest that events, such as the COVID-19 pandemic, may lead individuals to experience heightened anxiety and stress. While this emotional response could temporarily disrupt their social activities and coping mechanisms against trauma, it may not have a lasting impact on their perceived relationship quality. In fact, the example tweets we analyzed illustrated how individuals, despite feeling frustrated, expressed eagerness to resume their social activities after the pandemic. Hence, we argue that concerns about relationship disruption may be transient and serve as a positive signal prompting individuals in Japan to actively nurture their relationships. As suggested by the study conducted by Goodwin and Takahashi [52], dedicating additional time to communication, particularly in the context of romantic

relationships, could further enhance the quality of these connections [42].

Concern for Education Discontinuation

The peak volume of tweets was recorded on February 28, 2020, coinciding with the government's announcement of the simultaneous closure of all elementary, junior high, and senior high schools in Japan. In fact, in the most frequent nouns and verbs shown in Tables 1 and 2, the top words related to the simultaneous closure of schools were *graduation ceremony*, *cancel*, *lose*, *rest*, and *go* <negation>, all of which reflected Japanese citizens' concerns about the discontinuation of education, the cancellation of the graduation ceremony, and missing school classes. It is essential to note that in Japan, the graduation ceremony typically takes place in March and the new school and work year commences in April. Despite the Japanese government's earnest efforts to mitigate the spread of COVID-19, as scrutinized by scientists, the decision to close schools in Japan did not yield a substantial impact on preventing the spread of COVID-19. Instead, it deprived children of valuable learning and developmental opportunities [54]. Moreover, with the closure of schools, there was a surge in the demand for digital education or internet-based learning platforms. However, many schools and student households were ill-equipped to handle this impromptu shift to an internet-based education system. As discussed in detail by Iwabuchi et al [55], the unequal distribution of resources among schools in Japan further intensified the digital learning disparities brought about by the COVID-19 pandemic-induced school closures. The more well-funded private and prefecture-sponsored schools had often already implemented or could quickly set up the necessary e-learning system to cope with the lack of face-to-face lecturing. However, most public schools were forced to send learning materials to students by mail, risking a huge learning gap between students in private and public schools. The long-term impact on students' physical and mental development remains uncertain, given that most schools were able to resume normalcy after the lifting of the state of emergency. A study conducted by Nishimura et al [56] on medical students clearly indicated a deterioration in subjective mental well-being.

Concerns were also observed regarding web-based alternatives, with growing apprehensions that they fail to adequately substitute the essential in-person learning and hands-on field practice integral to medical education. The diverse concerns reflected in education-related keywords in Table 1 suggest that many Japanese individuals transitioned their focus from one-time events, such as *graduation ceremony* and *school holiday*, to longer-term mental and societal impacts, such as *opportunity*, *stress*, and *university*. This shift implies that the long-term effects would take time to manifest compared to short-term disruptions of specific incidents, such as a graduation ceremony. Further studies are crucial to monitor and unveil a complete picture of this disruption.

Long- and Short-Term Concerns and the Impact on the Society

Following the World Health Organization's official declaration that COVID-19 was no longer considered a global health emergency on May 4, 2023, individuals who survived now faced

a familiar daily life with some changes that were difficult to imagine in the pre-pandemic era. However, there is still an impact on society that can be challenging to trace and monitor. The economic repercussion, such as inflation and tumbling currency values in Japan, are gradually occurring. Schoolchildren who have lost education for almost 1 year are bracing for their future growth. An increasing number of companies are eager to get talent to opt in for remote working styles to attract employees who were reluctant to return to city offices. Individuals are probably no longer worried about toilet paper but will gradually sense the subtle shifting of their workstyles, social styles, and even learning styles. However, due to the limitations of our data, we were not able to speculate about the postpandemic future. Our discussion offers possible clues to further trace the causes of societal changes. The profound effects of the COVID-19 pandemic on society and public health require further investigation and monitoring.

Limitations and Future Work

It should be noted that our study had some limitations in extracting data from social platforms such as Twitter. One limitation is the lack of geolocation metadata. Although we capitalized on the language exclusivity of Japanese tweets and the domestic majority to extract representative samples of Japanese citizens, it is important to note that there may be some minor contributions from Japanese speakers residing outside Japan. This limitation should be considered when interpreting the findings of this study. Another limitation arises from the bias present on Twitter, as its use is lower among older adults compared to the younger population. To mitigate this bias, stratified analysis is necessary to account for the effects of age. However, the current system lacks age data. Consequently, the results should be interpreted with the awareness that the perspectives of the older adults are underrepresented.

Because the purpose of this study was to derive an interruption schedule, we specifically targeted verbs and nouns to better represent social connections (families and friends), locations, events, subjects, and actions, rather than using adjectives or phrases that might focus on emotional descriptions or concrete situations. This approach limited our options for sentiment-related analysis methods or topic modeling, which could reveal emotional reactions instead of generic events and the involvement of close connections. Although people's sentiments were deemed beyond the scope of this study, in future studies, we would like to analyze how people's sentiments have changed through sentiment analysis [57]. With the introduction of transformer-based large language models, such as bidirectional encoder representations from transformers and text-to-text transfer transformers, more contextual and in-depth understanding and analysis might be made available for researchers in social media data. This should be considered in future studies.

We also did not address concerns regarding the safety of cybersecurity during the work-from-home period during the pandemic. We noticed that in the United States, data breaches and the security of the work environment were one of the top concerns [58]; however, based on our current results, there was no direct implication on this aspect in Japan during the

COVID-19 pandemic. This will be considered in our future work.

Conclusions

Overall, by adding the analysis on *Corona no-sei* to the conventional symptom-based monitoring, we were able to identify the underlying concerns at the peak of the disruption and across the whole-time span of the 3 announcements of state of emergency. Our findings and a comparison of the tweets against COVID-19 case numbers yielded rich insights into people's short- and long-term concerns and potential aspects of societal impact caused by the announcements of the state of emergency. Although more studies from different fields would help to reveal the whole landscape of social and psychological impact caused by COVID-19, we believed that the keywords

reflected in *Corona no-sei* tweets provided more nuanced descriptions of real-life problems Japanese individuals faced during the COVID-19 pandemic and revealed the development of different concerns in response to the change of policies.

Timely communication of analysis results is crucial, especially when dealing with issues of significant social impact, such as a global pandemic. A delay in delivering results can hinder decision-making processes and require substantial resources to recover from the initial losses caused by poor decisions. For policy makers, especially the Japanese government, this study reflects the opinions of citizens and should be considered when reviewing the effectiveness and suitability of a policy as well as assessing further measures to support those impacted during the pandemic.

Acknowledgments

This research was conducted as part of the COVID-19 AI & Simulation Project run by Mitsubishi Research Institute commissioned by the Cabinet Secretariat. The methodology used in this study was developed through the support from Japan Science and Technology Agency (JST) Strategic International Collaborative Research Program (SICORP) (grant JPMJSC2107) and Japan Society for the Promotion of Science (JSPS) KAKENHI (grant JP 22K12041) in Japan. The authors acknowledge Dr Naoki Yoshinaga, Dr Masashi Toyoda, and Dr Masaru Kitsuregawa for providing the tweet-analysis system.

Conflicts of Interest

None declared.

References

1. Nicola M, Alsafi Z, Sohrabi C, Kerwan A, Al-Jabir A, Iosifidis C, et al. The socio-economic implications of the coronavirus pandemic (COVID-19): a review. *Int J Surg* 2020 Jun;78:185-193 [FREE Full text] [doi: [10.1016/j.ijssu.2020.04.018](https://doi.org/10.1016/j.ijssu.2020.04.018)] [Medline: [32305533](https://pubmed.ncbi.nlm.nih.gov/32305533/)]
2. Lewnard JA, Lo NC. Scientific and ethical basis for social-distancing interventions against COVID-19. *Lancet Infect Dis* 2020 Jun;20(6):631-633 [FREE Full text] [doi: [10.1016/S1473-3099\(20\)30190-0](https://doi.org/10.1016/S1473-3099(20)30190-0)] [Medline: [32213329](https://pubmed.ncbi.nlm.nih.gov/32213329/)]
3. Qian M, Jiang J. COVID-19 and social distancing. *Z Gesundh Wiss* 2022 May 25;30(1):259-261 [FREE Full text] [doi: [10.1007/s10389-020-01321-z](https://doi.org/10.1007/s10389-020-01321-z)] [Medline: [32837835](https://pubmed.ncbi.nlm.nih.gov/32837835/)]
4. Sen-Crowe B, McKenney M, Elkbuli A. Social distancing during the COVID-19 pandemic: staying home save lives. *Am J Emerg Med* 2020 Jul;38(7):1519-1520 [FREE Full text] [doi: [10.1016/j.ajem.2020.03.063](https://doi.org/10.1016/j.ajem.2020.03.063)] [Medline: [32305155](https://pubmed.ncbi.nlm.nih.gov/32305155/)]
5. Belot M, Choi S, Jamison JC, Papageorge NW, Tripodi E, van den Broek-Altenburg E. Six-country survey on COVID-19. *SSRN J* 2020. [doi: [10.2139/ssrn.3596697](https://doi.org/10.2139/ssrn.3596697)]
6. McBride O, Murphy J, Shevlin M, Gibson-Miller J, Hartman TK, Hyland P, et al. Monitoring the psychological, social, and economic impact of the COVID-19 pandemic in the population: context, design and conduct of the longitudinal COVID-19 psychological research consortium (C19PRC) study. *Int J Methods Psychiatr Res* 2021 Mar;30(1):e1861 [FREE Full text] [doi: [10.1002/mpr.1861](https://doi.org/10.1002/mpr.1861)] [Medline: [33166018](https://pubmed.ncbi.nlm.nih.gov/33166018/)]
7. Andersen L, Hansen T, Johannesen N, Sheridan A. Pandemic, shutdown and consumer spending: lessons from scandinavian policy responses to COVID-19. *arXiv Preprint* posted online May 10, 2020 [FREE Full text] [doi: [10.48550/arXiv.2005.04630](https://doi.org/10.48550/arXiv.2005.04630)]
8. Alexander D, Karger E. Do stay-at-home orders cause people to stay at home? Effects of stay-at-home orders on consumer behavior (working paper). Federal Reserve Bank of Chicago. 2020. URL: <https://ideas.repec.org/p/fip/fedhwp/87999.html> [accessed 2024-03-21]
9. Glorin S. A descriptive study on cybersecurity challenges of working from home during COVID-19 pandemic and a proposed 8 step WFH cyber-attack mitigation plan. *Commun CIBIMA* 2021 Feb 17;2021:1-7 [FREE Full text] [doi: [10.5171/2021.589235](https://doi.org/10.5171/2021.589235)]
10. Hakim AJ, Victory KR, Chevinsky JR, Hast MA, Weikum D, Kazazian L, et al. Mitigation policies, community mobility, and COVID-19 case counts in Australia, Japan, Hong Kong, and Singapore. *Public Health* 2021 May;194:238-244 [FREE Full text] [doi: [10.1016/j.puhe.2021.02.001](https://doi.org/10.1016/j.puhe.2021.02.001)] [Medline: [33965795](https://pubmed.ncbi.nlm.nih.gov/33965795/)]
11. Kuniya T. Evaluation of the effect of the state of emergency for the first wave of COVID-19 in Japan. *Infect Dis Model* 2020;5:580-587 [FREE Full text] [doi: [10.1016/j.idm.2020.08.004](https://doi.org/10.1016/j.idm.2020.08.004)] [Medline: [32844135](https://pubmed.ncbi.nlm.nih.gov/32844135/)]

12. Karako K, Song P, Chen Y, Tang W, Kokudo N. Overview of the characteristics of and responses to the three waves of COVID-19 in Japan during 2020-2021. *Biosci Trends* 2021 Mar 15;15(1):1-8 [FREE Full text] [doi: [10.5582/bst.2021.01019](https://doi.org/10.5582/bst.2021.01019)] [Medline: [33518668](https://pubmed.ncbi.nlm.nih.gov/33518668/)]
13. Kobayashi G, Sugawara S, Tamae H, Ozu T. Predicting intervention effect for COVID-19 in Japan: state space modeling approach. *Biosci Trends* 2020 Jul 17;14(3):174-181 [FREE Full text] [doi: [10.5582/bst.2020.03133](https://doi.org/10.5582/bst.2020.03133)] [Medline: [32461511](https://pubmed.ncbi.nlm.nih.gov/32461511/)]
14. Yamamura E, Tsutsui Y. How does the impact of the COVID-19 state of emergency change? An analysis of preventive behaviors and mental health using panel data in Japan. *Jpn Int Econ* 2022 Jun;64:101194 [FREE Full text] [doi: [10.1016/j.jjie.2022.101194](https://doi.org/10.1016/j.jjie.2022.101194)] [Medline: [35125647](https://pubmed.ncbi.nlm.nih.gov/35125647/)]
15. Katafuchi KY, Kurita K, Managi S. COVID-19 with stigma: theory and evidence from mobility data. *Econ Disaster Clim Chang* 2021;5(1):71-95 [FREE Full text] [doi: [10.1007/s41885-020-00077-w](https://doi.org/10.1007/s41885-020-00077-w)] [Medline: [32984755](https://pubmed.ncbi.nlm.nih.gov/32984755/)]
16. Kitamura Y, Karkour S, Ichisugi Y, Itsuno N. Evaluation of the economic, environmental, and social impacts of the COVID-19 pandemic on the Japanese tourism industry. *Sustainability* 2020 Dec 09;12(24):10302. [doi: [10.3390/su122410302](https://doi.org/10.3390/su122410302)]
17. Japan GDP 1960-2024 by World Bank. Macrotrends LLC. URL: <https://tinyurl.com/4pdrtsub> [accessed 2024-02-15]
18. Japan unemployment rate 1991-2024 by World Bank. Macrotrends LLC. URL: <https://tinyurl.com/bjv6ku2v> [accessed 2024-02-15]
19. Watanabe T, Yabu T. Japan's voluntary lockdown. *PLoS One* 2021 Jun 10;16(6):e0252468 [FREE Full text] [doi: [10.1371/journal.pone.0252468](https://doi.org/10.1371/journal.pone.0252468)] [Medline: [34111163](https://pubmed.ncbi.nlm.nih.gov/34111163/)]
20. Watanabe T, Yabu T. Japan's voluntary lockdown: further evidence based on age-specific mobile location data. *Jpn Econ Rev (Oxf)* 2021;72(3):333-370 [FREE Full text] [doi: [10.1007/s42973-021-00077-9](https://doi.org/10.1007/s42973-021-00077-9)] [Medline: [34177343](https://pubmed.ncbi.nlm.nih.gov/34177343/)]
21. Mizuno T, Ohnishi T, Watanabe T. Visualizing social and behavior change due to the outbreak of COVID-19 using mobile phone location data. *New Gener Comput* 2021 Nov 02;39(3-4):453-468 [FREE Full text] [doi: [10.1007/s00354-021-00139-x](https://doi.org/10.1007/s00354-021-00139-x)] [Medline: [34744249](https://pubmed.ncbi.nlm.nih.gov/34744249/)]
22. Aramaki E, Wakamiya S, Yada S, Nakamura Y. Natural language processing: from bedside to everywhere. *Yearb Med Inform* 2022 Aug 02;31(1):243-253 [FREE Full text] [doi: [10.1055/s-0042-1742510](https://doi.org/10.1055/s-0042-1742510)] [Medline: [35654422](https://pubmed.ncbi.nlm.nih.gov/35654422/)]
23. Mahmoud E, Émilien A, Maxime G, Gilles D. The role of text analytics in healthcare: a review of recent developments and applications. In: *Proceedings of the 14th International Joint Conference on Biomedical Engineering Systems and Technologies*. 2021 Presented at: BIOSTEC '21; February 11-13, 2021; Virtual event p. 825-832 URL: <https://www.scitepress.org/PublishedPapers/2021/104145/104145.pdf> [doi: [10.5220/0010414508250832](https://doi.org/10.5220/0010414508250832)]
24. Izutsu MN, Izutsu K. Why is Twitter so popular in Japan? *Internet Pragmat* 2019 Jul 16;2(2):260-289 [FREE Full text] [doi: [10.1075/ip.00030.izu](https://doi.org/10.1075/ip.00030.izu)]
25. Shanthakumar SG, Seetharam A, Ramesh A. Understanding the societal disruption due to COVID-19 via user tweets. In: *Proceedings of the 2021 IEEE International Conference on Smart Computing*. 2021 Presented at: SMARTCOMP '21; August 23-27, 2021; Irvine, CA p. 137-144 URL: <https://tinyurl.com/5ye73y5j> [doi: [10.1109/smartcomp52413.2021.00039](https://doi.org/10.1109/smartcomp52413.2021.00039)]
26. Lee H, Noh EB, Choi SH, Zhao B, Nam EW. Determining public opinion of the COVID-19 pandemic in South Korea and Japan: social network mining on Twitter. *Healthc Inform Res* 2020 Oct;26(4):335-343 [FREE Full text] [doi: [10.4258/hir.2020.26.4.335](https://doi.org/10.4258/hir.2020.26.4.335)] [Medline: [33190468](https://pubmed.ncbi.nlm.nih.gov/33190468/)]
27. Chen LL, Wilson SR, Lohmann S, Negraia DV. What are you anxious about? Examining subjects of anxiety during the COVID-19 pandemic. *Proc Int AAI Conf Web Soc Media* 2023 Jun 02;17:137-148. [doi: [10.1609/icwsm.v17i1.22133](https://doi.org/10.1609/icwsm.v17i1.22133)]
28. Li I, Li Y, Li T. What are we depressed about when we talk about COVID-19: mental health analysis on tweets using natural language processing. In: *Proceedings of the 40th SGAI International Conference on Artificial Intelligence*. 2020 Presented at: SGAI '20; December 15-17, 2020; Cambridge, UK URL: https://doi.org/10.1007/978-3-030-63799-6_27 [doi: [10.1007/978-3-030-63799-6_27](https://doi.org/10.1007/978-3-030-63799-6_27)]
29. Lyu JC, Han EL, Luli GK. COVID-19 vaccine-related discussion on Twitter: topic modeling and sentiment analysis. *J Med Internet Res* 2021 Jun 29;23(6):e24435 [FREE Full text] [doi: [10.2196/24435](https://doi.org/10.2196/24435)] [Medline: [34115608](https://pubmed.ncbi.nlm.nih.gov/34115608/)]
30. Krittanawong C, Narasimhan B, Virk HU, Narasimhan H, Hahn J, Wang Z, et al. Misinformation dissemination in Twitter in the COVID-19 era. *Am J Med* 2020 Dec;133(12):1367-1369 [FREE Full text] [doi: [10.1016/j.amjmed.2020.07.012](https://doi.org/10.1016/j.amjmed.2020.07.012)] [Medline: [32805227](https://pubmed.ncbi.nlm.nih.gov/32805227/)]
31. Ansari MT, Khan NA. Worldwide COVID-19 vaccines sentiment analysis through Twitter content. *Electron J Gen Med* 2021;18(6):em329. [doi: [10.29333/ejgm/11316](https://doi.org/10.29333/ejgm/11316)]
32. Ferawati K, Liew K, Aramaki E, Wakamiya S. Monitoring mentions of COVID-19 vaccine side effects on Japanese and Indonesian Twitter: infodemiological study. *JMIR Infodemiology* 2022;2(2):e39504 [FREE Full text] [doi: [10.2196/39504](https://doi.org/10.2196/39504)] [Medline: [36277140](https://pubmed.ncbi.nlm.nih.gov/36277140/)]
33. Gao Z, Fujita S, Shimizu N, Liew K, Murayama T, Yada S, et al. Measuring public concern about COVID-19 in Japanese internet users through search queries: infodemiological study. *JMIR Public Health Surveill* 2021 Jul 20;7(7):e29865 [FREE Full text] [doi: [10.2196/29865](https://doi.org/10.2196/29865)] [Medline: [34174781](https://pubmed.ncbi.nlm.nih.gov/34174781/)]
34. Uehara M, Fujita S, Shimizu N, Liew K, Wakamiya S, Aramaki E. Measuring concerns about the COVID-19 vaccine among Japanese internet users through search queries. *Sci Rep* 2022 Sep 03;12(1):15037 [FREE Full text] [doi: [10.1038/s41598-022-18307-4](https://doi.org/10.1038/s41598-022-18307-4)] [Medline: [36057657](https://pubmed.ncbi.nlm.nih.gov/36057657/)]

35. Ishida T, Seki Y, Kashino W, Kando N. Extracting citizen feedback from social media by appraisal opinion type viewpoint. *J Nat Lang Process* 2022;29(2):416-442. [doi: [10.5715/jnlp.29.416](https://doi.org/10.5715/jnlp.29.416)]
36. Special site for coronavirus in Japan. NHK. URL: <https://www3.nhk.or.jp/news/special/coronavirus/data/> [accessed 2023-04-19]
37. Yoshinaga N, Kitsuregawa M. A self-adaptive classifier for efficient text-stream processing. In: Proceedings of the 25th International Conference on Computational Linguistics: Technical Papers. 2014 Presented at: COLING '14; August 23-29, 2014; Dublin, Ireland p. 1091-1102 URL: <https://aclanthology.org/C14-1103.pdf>
38. Yoshinaga N, Kitsuregawa M. Polynomial to linear: Efficient classification with conjunctive features. In: Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing. 2009 Presented at: EMNLP '09; August 6-7, 2009; Singapore p. 1542-1551 URL: <https://aclanthology.org/D09-1160.pdf> [doi: [10.3115/1699648.1699701](https://doi.org/10.3115/1699648.1699701)]
39. Yoshinaga N, Kitsuregawa M. Kernel slicing: scalable online training with conjunctive features. In: Proceedings of the 23rd International Conference on Computational Linguistics. 2010 Presented at: COLING '10; August 23-27, 2010; Beijing, China p. 1245-1253 URL: <https://aclanthology.org/C10-1140.pdf>
40. Oh JH, Torisawa K, Hashimoto C, Iida R, Tanaka M, Kloetzer J. A semi-supervised learning approach to why-question answering. *Proc Int AAAI Conf Web Soc Media* 2016 Mar 05;30(1):10388. [doi: [10.1609/aaai.v30i1.10388](https://doi.org/10.1609/aaai.v30i1.10388)]
41. Ethical guidelines for life sciences and medical research involving human subjects. NHK. URL: <https://www.meti.go.jp/press/2021/03/2022031> [accessed 2022-06-28]
42. Kikuchi S, Kitao S, Mikoshiba M. Who suffers from the COVID-19 shocks? Labor market heterogeneity and welfare consequences in Japan. *J Jpn Int Econ* 2021 Mar;59:101117. [doi: [10.1016/j.jjie.2020.101117](https://doi.org/10.1016/j.jjie.2020.101117)]
43. Fukai T, Ichimura H, Kawata K. Describing the impacts of COVID-19 on the labor market in Japan until June 2020. *Jpn Econ Rev (Oxf)* 2021 Jul 19;72(3):439-470 [FREE Full text] [doi: [10.1007/s42973-021-00081-z](https://doi.org/10.1007/s42973-021-00081-z)] [Medline: [34305434](https://pubmed.ncbi.nlm.nih.gov/34305434/)]
44. Kikuchi S, Kitao S, Mikoshiba M. Heterogeneous vulnerability to the COVID-19 crisis and implications for inequality in Japan. Research Institute of Economy, Trade, and Industry. 2020 Apr. URL: <https://www.rieti.go.jp/jp/publications/dp/20e039.pdf> [accessed 2024-03-05]
45. Leung J, Chung JY, Tisdale C, Chiu V, Lim CC, Chan G. Anxiety and panic buying behaviour during COVID-19 pandemic-a qualitative analysis of toilet paper hoarding contents on Twitter. *Int J Environ Res Public Health* 2021 Jan 27;18(3):1127 [FREE Full text] [doi: [10.3390/ijerph18031127](https://doi.org/10.3390/ijerph18031127)] [Medline: [33514049](https://pubmed.ncbi.nlm.nih.gov/33514049/)]
46. Kaigo M. Social media usage during disasters and social capital: Twitter and the Great East Japan earthquake. *Keio Commun Rev* 2012;34(1):19-35 [FREE Full text]
47. Kirk CP, Rifkin LS. I'll trade you diamonds for toilet paper: consumer reacting, coping and adapting behaviors in the COVID-19 pandemic. *J Bus Res* 2020 Sep;117:124-131 [FREE Full text] [doi: [10.1016/j.jbusres.2020.05.028](https://doi.org/10.1016/j.jbusres.2020.05.028)] [Medline: [32834208](https://pubmed.ncbi.nlm.nih.gov/32834208/)]
48. Tatapudi H, Das R, Das TK. Impact assessment of full and partial stay-at-home orders, face mask usage, and contact tracing: an agent-based simulation study of COVID-19 for an urban region. *Glob Epidemiol* 2020 Nov;2:100036 [FREE Full text] [doi: [10.1016/j.gloepi.2020.100036](https://doi.org/10.1016/j.gloepi.2020.100036)] [Medline: [33103108](https://pubmed.ncbi.nlm.nih.gov/33103108/)]
49. Catching A, Capponi S, Yeh MT, Bianco S, Andino R. Examining the interplay between face mask usage, asymptomatic transmission, and social distancing on the spread of COVID-19. *Sci Rep* 2021 Aug 06;11(1):15998 [FREE Full text] [doi: [10.1038/s41598-021-94960-5](https://doi.org/10.1038/s41598-021-94960-5)] [Medline: [34362936](https://pubmed.ncbi.nlm.nih.gov/34362936/)]
50. Adjodah D, Dinakar K, Chinazzi M, Fraiberger SP, Pentland A, Bates S, et al. Association between COVID-19 outcomes and mask mandates, adherence, and attitudes. *PLoS One* 2021 Jun 23;16(6):e0252315 [FREE Full text] [doi: [10.1371/journal.pone.0252315](https://doi.org/10.1371/journal.pone.0252315)] [Medline: [34161332](https://pubmed.ncbi.nlm.nih.gov/34161332/)]
51. Osaki T. Abenomask? Prime minister's 'two masks per household' policy spawns memes on social media. *The Japan Times*. 2020 Apr 2. URL: <https://www.japantimes.co.jp/news/2020/04/02/national/> [accessed 2024-03-05]
52. Goodwin R, Takahashi M. Anxiety, past trauma and changes in relationships in Japan during COVID-19. *J Psychiatr Res* 2022 Jul;151:377-381 [FREE Full text] [doi: [10.1016/j.jpsychires.2022.04.032](https://doi.org/10.1016/j.jpsychires.2022.04.032)] [Medline: [35567802](https://pubmed.ncbi.nlm.nih.gov/35567802/)]
53. Tahara M, Mashizume Y, Takahashi K. Mental health crisis and stress coping among healthcare college students momentarily displaced from their campus community because of COVID-19 restrictions in Japan. *Int J Environ Res Public Health* 2021 Jul 06;18(14):7245 [FREE Full text] [doi: [10.3390/ijerph18147245](https://doi.org/10.3390/ijerph18147245)] [Medline: [34299694](https://pubmed.ncbi.nlm.nih.gov/34299694/)]
54. Fukumoto K, McClean CT, Nakagawa K. No causal effect of school closures in Japan on the spread of COVID-19 in spring 2020. *Nat Med* 2021 Dec 27;27(12):2111-2119 [FREE Full text] [doi: [10.1038/s41591-021-01571-8](https://doi.org/10.1038/s41591-021-01571-8)] [Medline: [34707318](https://pubmed.ncbi.nlm.nih.gov/34707318/)]
55. Iwabuchi K, Hodama K, Onishi Y, Miyazaki S, Nakae S, Suzuki KH. COVID-19 and education on the front lines in Japan: what caused learning disparities and how did the government and schools take initiative? In: Reimers FM, editor. *Primary and Secondary Education During Covid-19: Disruptions to Educational Opportunity During a Pandemic*. Cham, Switzerland: Springer; 2022:125-151.
56. Nishimura Y, Ochi K, Tokumasu K, Obika M, Hagiya H, Kataoka H, et al. Impact of the COVID-19 pandemic on the psychological distress of medical students in Japan: cross-sectional survey study. *J Med Internet Res* 2021 Feb 18;23(2):e25232 [FREE Full text] [doi: [10.2196/25232](https://doi.org/10.2196/25232)] [Medline: [33556033](https://pubmed.ncbi.nlm.nih.gov/33556033/)]

57. Ainley E, Witwicki C, Tallett A, Graham C. Using Twitter comments to understand people's experiences of UK health care during the COVID-19 pandemic: thematic and sentiment analysis. *J Med Internet Res* 2021 Oct 25;23(10):e31101 [FREE Full text] [doi: [10.2196/31101](https://doi.org/10.2196/31101)] [Medline: [34469327](https://pubmed.ncbi.nlm.nih.gov/34469327/)]
58. Glorin S. Cyber kill chain analysis of five major US data breaches: lessons learnt and prevention plan. *Int J Cyber Warf Terror* 2022;12(1):1-15 [FREE Full text] [doi: [10.4018/ijcwt.315651](https://doi.org/10.4018/ijcwt.315651)]

Edited by T Mackey; submitted 06.06.23; peer-reviewed by L Chen, Y Pachankis, G Sebastian, M Elbattah; comments to author 29.06.23; revised version received 11.08.23; accepted 06.03.24; published 01.04.24.

Please cite as:

Kamba M, She WJ, Ferawati K, Wakamiya S, Aramaki E

Exploring the Impact of the COVID-19 Pandemic on Twitter in Japan: Qualitative Analysis of Disrupted Plans and Consequences
JMIR Infodemiology 2024;4:e49699

URL: <https://infodemiology.jmir.org/2024/1/e49699>

doi: [10.2196/49699](https://doi.org/10.2196/49699)

PMID: [38557446](https://pubmed.ncbi.nlm.nih.gov/38557446/)

©Masaru Kamba, Wan Jou She, Kiki Ferawati, Shoko Wakamiya, Eiji Aramaki. Originally published in JMIR Infodemiology (<https://infodemiology.jmir.org>), 01.04.2024. 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.

Original Paper

Perceptions of Health Misinformation on Social Media: Cross-Sectional Survey Study

Anna Gaysynsky^{1,2}, MPH; Nicole Senft Everson¹, PhD; Kathryn Heley¹, PhD, MPH; Wen-Ying Sylvia Chou¹, PhD, MPH

¹Health Communication and Informatics Research Branch, Behavioral Research Program, National Cancer Institute, Rockville, MD, United States

²ICF Next, ICF, Rockville, MD, United States

Corresponding Author:

Anna Gaysynsky, MPH

Health Communication and Informatics Research Branch

Behavioral Research Program

National Cancer Institute

9609 Medical Center Drive

Rockville, MD, 20850

United States

Phone: 1 240 276 5284

Email: anna.gaysynsky@nih.gov

Abstract

Background: Health misinformation on social media can negatively affect knowledge, attitudes, and behaviors, undermining clinical care and public health efforts. Therefore, it is vital to better understand the public's experience with health misinformation on social media.

Objective: The goal of this analysis was to examine perceptions of the social media information environment and identify associations between health misinformation perceptions and health communication behaviors among US adults.

Methods: Analyses used data from the 2022 Health Information National Trends Survey (N=6252). Weighted unadjusted proportions described respondents' perceptions of the amount of false or misleading health information on social media ("perceived misinformation amount") and how difficult it is to discern true from false information on social media ("perceived discernment difficulty"). Weighted multivariable logistic regressions examined (1) associations of sociodemographic characteristics and subjective literacy measures with misinformation perceptions and (2) relationships between misinformation perceptions and health communication behaviors (ie, sharing personal or general health information on social media and using social media information in health decisions or in discussions with health care providers).

Results: Over one-third of social media users (35.61%) perceived high levels of health misinformation, and approximately two-thirds (66.56%) reported high perceived discernment difficulty. Odds of perceiving high amounts of misinformation were lower among non-Hispanic Black/African American (adjusted odds ratio [aOR] 0.407, 95% CI 0.282-0.587) and Hispanic (aOR 0.610, 95% CI 0.449-0.831) individuals compared to White individuals. Those with lower subjective health literacy were less likely to report high perceived misinformation amount (aOR 0.602, 95% CI 0.374-0.970), whereas those with lower subjective digital literacy were more likely to report high perceived misinformation amount (aOR 1.775, 95% CI 1.400-2.251). Compared to White individuals, Hispanic individuals had lower odds of reporting high discernment difficulty (aOR 0.620, 95% CI 0.462-0.831). Those with lower subjective digital literacy (aOR 1.873, 95% CI 1.478-2.374) or numeracy (aOR 1.465, 95% CI 1.047-2.049) were more likely to report high discernment difficulty. High perceived misinformation amount was associated with lower odds of sharing general health information on social media (aOR 0.742, 95% CI 0.568-0.968), using social media information to make health decisions (aOR 0.273, 95% CI 0.156-0.479), and using social media information in discussions with health care providers (aOR 0.460, 95% CI 0.323-0.655). High perceived discernment difficulty was associated with higher odds of using social media information in health decisions (aOR 1.724, 95% CI 1.208-2.460) and health care provider discussions (aOR 1.389, 95% CI 1.035-1.864).

Conclusions: Perceptions of high health misinformation prevalence and discernment difficulty are widespread among social media users, and each has unique associations with sociodemographic characteristics, literacy, and health communication behaviors. These insights can help inform future health communication interventions.

KEYWORDS

social media; misinformation; health communication; health literacy; patient-provider communication

Introduction

Background

The Pew Research Center estimates that approximately 72% of Americans use social media [1], and research suggests that social media is widely used for health-related purposes specifically [2]. Social media has become an important venue for the exchange of health-related information and advice [3]. In 2019, 41% of internet users in the United States reported watching health-related YouTube videos, and 17% reported sharing health information on social networking sites [4]. Social media can help people find and access more useful and personally relevant information, facilitate the exchange of social support, and aid with disease management efforts [3]. However, while social media can make health information more accessible, the use of social media for health information seeking can also create the risk of harm through exposure to misinformation.

Defined as “health-related information that is false, inaccurate, or misleading according to the best available evidence at the time,” health misinformation is increasingly recognized as a threat to public health [5,6] (note that this definition includes *disinformation*, or false information that is created and spread with the intent to deceive, as a subset of misinformation [7]). Although health misinformation is not a new phenomenon, social media facilitates the rapid spread of falsehoods [6], thereby exacerbating the potential negative impact of misinformation on both individual and population health. Certain features of social media platforms, such as incentives that reward the sharing of content that receives more engagement, can result in a focus on sharing emotionally charged or provocative content rather than accurate content [6,8]. Additionally, algorithms that suggest content to users are often based on past engagement behavior, which can reinforce echo chambers, whereby users who engage with misinformation increasingly encounter further misinformation [6,8].

Many studies have documented substantial health-related misinformation on social media across a range of topics (including tobacco products, drugs, and vaccines) [9], and research increasingly suggests that social media misinformation can have a negative impact on health-related attitudes, behaviors, and outcomes. For example, Pierri et al [10] found that the prevalence of COVID-19 vaccine misinformation on Twitter (now rebranded as X) was related to higher levels of vaccine hesitancy and lower vaccination uptake rates in the United States even after accounting for political and sociodemographic factors. Furthermore, their causality analysis suggested a directional relationship between social media misinformation and vaccine hesitancy, with a lag of approximately 2 to 6 days from misinformation being posted in a county to a corresponding increase in vaccine hesitancy in that county [10]. Further evidence of a causal relationship between exposure to misinformation and health-related attitudes and intentions is provided by a randomized controlled trial conducted in the

United States and the United Kingdom, which showed that exposure to misinformation in the form of social media posts decreased the number of respondents who said that they would “definitely” take the COVID-19 vaccine by approximately 6 percentage points relative to the control group [11].

While there is a growing body of research examining the prevalence of misinformation on social media, as well as the association between social media misinformation and health-related outcomes, little work to date has focused on understanding how individuals perceive misinformation on social media or how these perceptions impact behavior. Surveys conducted over the past few years show that many US adults believe that much of the news they see on social media is false or inaccurate [12,13]. This is significant because studies have shown that misinformation perceptions impact communication behaviors. For example, a study conducted in Germany found higher self-perceived exposure to “fake news” to be associated with more frequent engagement in information verification behaviors on Facebook [14]. Meanwhile, a study conducted on Amazon Mechanical Turk found that perceiving fake news to have a greater influence on others than on oneself (ie, the third-person effect) was associated with lower intent to share news obtained from social media (either online or offline) [15].

Studies also suggest that perceptions of misinformation prevalence are associated with attitudes toward health issues—for example, one cross-sectional study found perceptions of high misinformation prevalence to be correlated with worry about COVID-19 [16]. Notably, the study found neither a significant association between actual misinformation prevalence (as measured using the “Infodemic Risk Index,” which produces frequency estimates of misinformation on Twitter by country) and worry about COVID-19 nor an interaction between actual misinformation prevalence and perceived prevalence in explaining pandemic worry [16]. This suggests critical psychological and cognitive effects of misinformation perceptions independent of actual misinformation prevalence (and separate from misinformation endorsement or belief) [16].

In addition to assessing perceptions of misinformation prevalence, assessing people’s confidence in their ability to detect misinformation is important because confidence can affect the way people make subsequent judgments [17]. For example, confidence can determine whether an individual acts on their initial judgment or seeks out additional information [18]. Confidence levels also affect a person’s willingness and ability to defend their assessments such that individuals who are able to discern true from false information—and are confident about their judgments—are more resistant to misinformation [18]. A person’s perceptions about their ability to accurately detect misinformation can also influence their perceptions about their capacity to manage health issues or make health decisions. For example, Park et al [19] found that people who had higher confidence in their ability to distinguish between

true and false COVID-19 information also had higher COVID-19 risk readiness perceptions (ie, felt that they had a “handle on the issues and developments surrounding the coronavirus outbreak”). Unfortunately, confidence can also be easily undermined, particularly when an individual is unsure about the validity of the material they are considering or lacks the necessary skills or literacy competencies to feel secure in their assessment [17].

Beyond obtaining a better understanding of misinformation perceptions and how they impact cognitive and behavioral processes and outcomes, it is also important to assess whether these perceptions vary by sociodemographic or other characteristics to identify groups that may be more vulnerable to misinformation and in need of more targeted efforts. For example, the trial conducted by Loomba et al [11] showed that some groups were differentially affected by exposure to misinformation—in the United States, female individuals were found to be less resistant to misinformation than male individuals, whereas those with lower incomes were found to be more resistant. Additionally, a survey conducted by the Pew Research Center in 2016 found that White individuals were more likely than Black and Hispanic individuals to say that they often saw fake political news online, and those with annual incomes of at least US \$75,000 were more likely to report seeing fake news compared to those who made less than US \$75,000 per year [20]. Findings regarding demographic differences in perceived ability to discern true from false information are more mixed. The Pew survey found that confidence in detecting fake political news did not differ significantly by sociodemographic characteristics (such as age, gender, income, or race) [20], whereas the study conducted by Park et al [19] found education and income to be significant predictors of confidence in distinguishing true from false information about COVID-19. Furthermore, a large survey experiment conducted by Sirlin et al [21] found digital literacy (as measured by familiarity with internet-related terms and attitudes toward technology as well as understanding of social media algorithms) to be an important predictor of the ability to discern truths from falsehoods when judging headline accuracy for both political and COVID-19 articles. Taken together, these findings demonstrate the importance of assessing how perceptions of health misinformation and misinformation discernment vary by sociodemographic characteristics and literacy. Although research regarding vulnerability to misinformation remains mixed (eg, with regard to age, sex, and income) [22], the potential impact of social media misinformation on health disparities is an important issue that requires ongoing attention.

Study Aims

Because misinformation perceptions can affect attitudes and behaviors, a better understanding of the public’s perceptions of health misinformation on social media and their ability to detect it, as well as possible subgroup differences in such perceptions, is needed. Toward that end, this study analyzed data from the National Cancer Institute’s 2022 Health Information National Trends Survey (HINTS 6) to (1) assess the prevalence of 2 distinct misinformation-related perceptions—perceived amount of health misinformation on social media and perceived ability to distinguish true from false health information on social

media—(2) identify sociodemographic factors associated with these health misinformation perceptions; and (3) explore associations between these misinformation perceptions and health communication behaviors, including information sharing, health decision-making, and communicating with health care providers.

Methods

Data and Sample Selection

HINTS is a nationally representative, cross-sectional, self-administered survey of civilian, noninstitutionalized US adults aged ≥ 18 years. Data for HINTS 6 (N=6252) were collected between March 7, 2022, and November 8, 2022, using questionnaires administered via mailed paper or web-based surveys. The overall response rate for HINTS 6 was 28.1%. Respondents who reported that they did not use social media (1211/6252, 19.37%) were excluded from the analyses, resulting in a starting analytic sample of 5041. Details regarding the design of HINTS 6, including methodology, sampling, and weighting procedures, have been published elsewhere [23].

Ethical Considerations

HINTS 6 received approval from the Westat Institutional Review Board on May 10, 2021 (6632.03.51), and was designated as non-human subjects research by the National Institutes of Health Office of Human Subjects Research on August 16, 2021 (000626). Respondents’ return of the completed survey indicated consent to participate.

Measures

Social Media Health Misinformation Perceptions

A total of 2 social media misinformation-related perceptions were measured. Perceived amount of misinformation on social media (“perceived misinformation amount”) was assessed with the following item: “How much of the health information that you see on social media do you think is false or misleading?” Response options were *none*, *a little*, *some*, *a lot*, and *I do not use social media* (as noted previously, those who selected “I do not use social media” in response to this item were excluded from the analyses).

Perceived difficulty distinguishing true from false information on social media (“perceived discernment difficulty”) was measured by assessing agreement with the following statement—“I find it hard to tell whether health information on social media is true or false”—among respondents who reported social media use. Response options were *strongly agree*, *somewhat agree*, *somewhat disagree*, and *strongly disagree*.

Health Communication Behaviors Related to Social Media Use

Information sharing on social media was assessed using two items that asked how often in the previous 12 months respondents (1) “share[d] personal health information on social media” and (2) “share[d] general health-related information on social media (for example, a news article).” Response options were *almost every day*, *at least once a week*, *a few times a month*, *less than once a month*, and *never*.

Respondents' use of information encountered on social media was assessed through reported agreement with 2 items: "I use information from social media to make decisions about my health" and "I use information from social media in discussions with my healthcare provider." Response options were *strongly agree*, *somewhat agree*, *somewhat disagree*, and *strongly disagree*.

Sociodemographic Characteristics

Sociodemographic variables included (1) educational level (categorized as high school degree or lower, some college or vocational training, and college graduate or higher), (2) sex (male or female), (3) age (18-24 years, 25-34 years, 35-44 years, 45-54 years, 55-64 years, and ≥ 65 years), (4) race or ethnicity (non-Hispanic White; non-Hispanic Black/African American; Hispanic; and non-Hispanic other, which included non-Hispanic American Indian or Alaska Native, non-Hispanic Asian, non-Hispanic Native Hawaiian or other Pacific Islander, and non-Hispanic multiple races), (5) annual household income (<US \$20,000, US \$20,000-<\$35,000, US \$35,000-<\$50,000, US \$50,000-<\$75,000, US \$75,000-<\$100,000, and \geq US \$100,000), and (6) geographic residence (urban or rural based on the 2013 Rural-Urban Continuum Codes).

Literacy Measures

Subjective health literacy was assessed using the following item: "How confident are you filling out medical forms by yourself?" Response options were *very* [confident], *somewhat* [confident], *a little* [confident], and *not at all* [confident]. This measure is one of the brief screening questions identified by Chew et al [24] for detecting inadequate or marginal health literacy among adults. Subjective digital literacy was assessed using the following item: "How confident are you that you can find helpful health resources on the Internet?" Response options were *completely confident*, *very confident*, *somewhat confident*, *a little confident*, and *not at all confident*. This measure was adapted from the eHealth Literacy Scale [25]. Subjective numeracy was assessed using the following item: "In general, how easy or hard do you find it to understand medical statistics?" Response options were *very easy*, *easy*, *hard*, and *very hard*. This item, which is part of the STAT-Confidence scale developed by Woloshin et al [26], has been shown to be a strong predictor of scores on the Newest Vital Sign measure (an objective measure of health literacy and numeracy) [27].

Statistical Analysis

To account for the complex sampling design of HINTS, analyses were conducted in SAS (version 9.4; SAS Institute) using final sample weights to obtain population-level point estimates and a set of 50 replicate weights to compute accurate variance estimates [23]. Frequencies and survey-weighted unadjusted proportions were used to describe the distributions of perceived misinformation amount and perceived discernment difficulty.

In total, 2 weighted multivariable logistic regression models examined associations of sociodemographic characteristics and literacy measures with perceived misinformation amount and perceived discernment difficulty. For these analyses, perceived misinformation amount was dichotomized to reflect high

perceived misinformation amount (*a lot*) versus low perceived misinformation amount (*none*, *a little*, or *some*) to facilitate comparison between those who perceived misinformation to be a significant problem in the information environment and those who did not. Furthermore, only a relatively small proportion of respondents felt that "none" or only "a little" of the information they saw on social media was false or misleading, whereas over a third of the sample reported that "a lot" of the information they saw was false or misleading. Perceived discernment difficulty was dichotomized as high (*strongly agree* or *somewhat agree*) versus low (*somewhat disagree* or *strongly disagree*). Additionally, subjective health literacy was dichotomized as high (*very* [confident] or *somewhat* [confident]) versus low (*a little* [confident] or *not at all* [confident]), digital literacy was dichotomized as high (*completely confident* or *very confident*) versus low (*somewhat confident*, *a little confident*, or *not at all confident*), and numeracy was dichotomized as high (*very easy* or *easy*) versus low (*hard* or *very hard*).

A total of 4 additional weighted multivariable logistic regression models tested associations of high versus low perceived misinformation amount and high versus low perceived discernment difficulty with communication behaviors related to social media use (ie, sharing personal health information on social media, sharing general health information on social media, using information from social media to make health decisions, and using information from social media in discussions with health care providers) adjusted for sociodemographic characteristics and dichotomized literacy measures. The 2 information-sharing behavior measures were dichotomized as ever shared (*almost every day*, *at least once a week*, *a few times a month*, or *less than once a month*) versus never shared, whereas the 2 social media information use items were dichotomized as agreement (*strongly agree* or *somewhat agree*) versus disagreement (*somewhat disagree* or *strongly disagree*). Sensitivity analyses tested the interaction of perceived misinformation amount and discernment difficulty in predicting these communication outcomes, but this interaction was not statistically significant in any of the models.

Adjusted analyses used complete case analysis, with valid analytic samples reported in tables corresponding to each analysis. Descriptive information on missing data for each variable is publicly available on the HINTS website [28]. Tests of significance were conducted at the $P < .05$ level.

Results

Prevalence of Social Media Health Misinformation Perceptions

As shown in Figure 1, over one-third of American social media users (35.61%) perceived "a lot" of misinformation on social media (ie, expressed high perceived misinformation amount), whereas only a very small percentage (1.54%) perceived that "none" of the health information they see is false or misleading. Figure 2 shows that approximately two-thirds of American social media users agreed that they find it hard to tell whether health information on social media is true or false (ie, endorsed high discernment difficulty).

Figure 1. Weighted proportions of perceived health misinformation amount among American social media users.

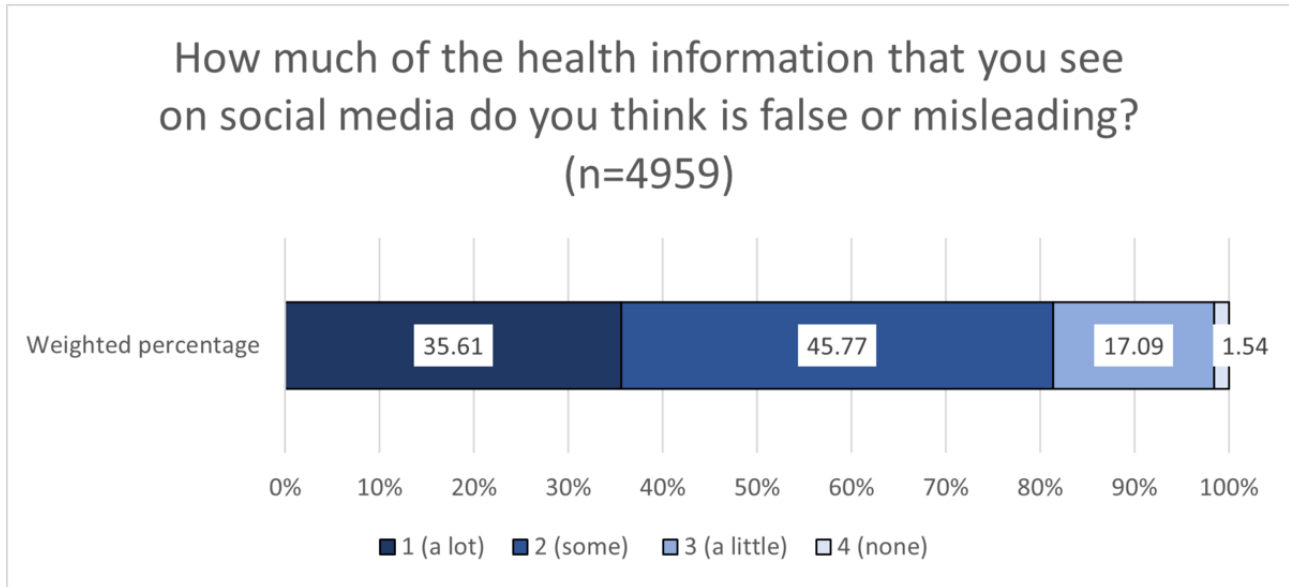
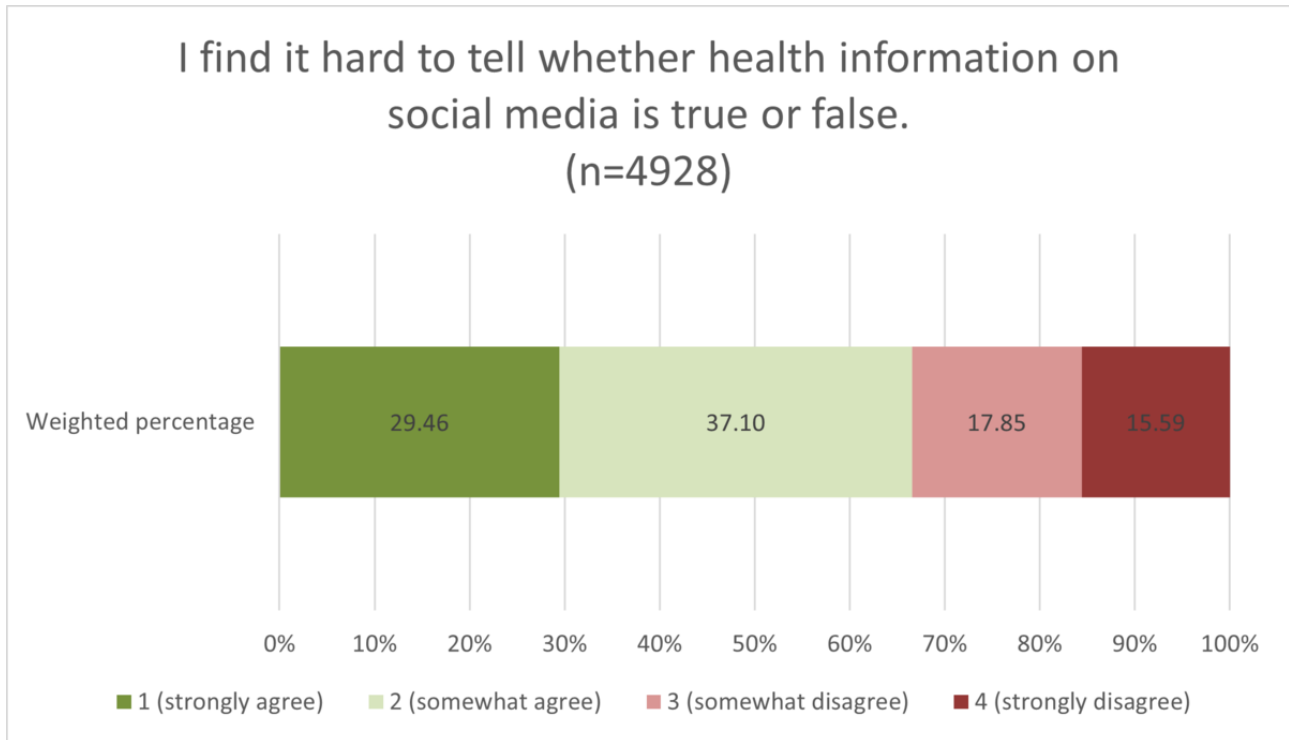


Figure 2. Weighted proportions of perceived discernment difficulty among American social media users.



Predictors of Social Media Health Misinformation Perceptions

Perceived Misinformation Amount

As shown in Table 1, individuals who were non-Hispanic Black/African American (compared to non-Hispanic White individuals; adjusted odds ratio [aOR] 0.407, 95% CI 0.282-0.587) or Hispanic (compared non-Hispanic White individuals; aOR 0.610, 95% CI 0.449-0.831) or who had lower

subjective health literacy (vs those with higher health literacy; aOR 0.602, 95% CI 0.374-0.970) were less likely to report high perceived misinformation amount. Comparatively, respondents with lower subjective digital literacy were more likely to report high misinformation amount (vs those with higher digital literacy; aOR 1.775, 95% CI 1.400-2.251). Age, sex, educational level, income, geographic residence, and numeracy were not statistically significantly related to perceived amount of misinformation.

Table 1. Predictors of social media health misinformation perceptions.

Independent variable	Perceived misinformation amount ^a (n=4218), adjusted odds ratio (95% CI)	Perceived discernment difficulty ^b (n=4205), adjusted odds ratio (95% CI)
Age (y; reference: 18-24)		
25-34	0.745 (0.403-1.376)	<i>0.485</i> (0.289-0.816) ^c
35-44	0.859 (0.495-1.491)	<i>0.487</i> (0.311-0.763)
45-54	0.772 (0.410-1.454)	0.624 (0.384-1.014)
55-64	0.790 (0.425-1.468)	<i>0.605</i> (0.369-0.990)
≥65	0.688 (0.384-1.232)	0.841 (0.525-1.346)
Sex (reference: male)		
Female	1.037 (0.791-1.358)	1.077 (0.871-1.331)
Educational level (reference: high school or lower)		
Some college or vocational training	1.206 (0.867-1.677)	1.210 (0.882-1.660)
College graduate or higher	1.144 (0.809-1.618)	0.871 (0.632-1.200)
Race or ethnicity (reference: non-Hispanic White)		
Hispanic	<i>0.610</i> (0.449-0.831)	<i>0.620</i> (0.462-0.831)
Non-Hispanic Black/African American	<i>0.407</i> (0.282-0.587)	0.830 (0.596-1.156)
Non-Hispanic other	0.977 (0.662-1.442)	1.126 (0.788-1.607)
Annual household income (reference: <US \$20,000)		
US \$20,000-\$34,999	1.283 (0.747-2.202)	1.016 (0.640-1.614)
US \$35,000-\$49,999	1.039 (0.613-1.760)	1.061 (0.678-1.659)
US \$50,000-\$74,999	1.619 (0.968-2.709)	1.058 (0.740-1.512)
US \$75,000-\$99,999	1.693 (0.996-2.880)	1.459 (0.932-2.283)
≥US \$100,000	1.469 (0.910-2.369)	1.245 (0.870-1.780)
Geographic residence (reference: urban)		
Rural	1.012 (0.770-1.331)	1.109 (0.769-1.600)
Health literacy (reference: high health literacy)		
Low health literacy	<i>0.602</i> (0.374-0.970)	1.230 (0.829-1.824)
Digital literacy (reference: high digital literacy)		
Low digital literacy	<i>1.775</i> (1.400-2.251)	<i>1.873</i> (1.478-2.374)
Numeracy (reference: high numeracy)		
Low numeracy	1.030 (0.771-1.376)	<i>1.465</i> (1.047-2.049)

^aThe probability modeled was odds of high perceived misinformation amount (*a lot*) in reference to low perceived misinformation amount (*none, a little, or some*).

^bThe probability modeled was odds of reporting high perceived discernment difficulty (*strongly agree or somewhat agree*) in reference to low perceived discernment difficulty (*somewhat disagree or strongly disagree*).

^cItalicized values are statistically significant ($P < .05$).

Perceived Discernment Difficulty

As shown in Table 1, there were differences in perceived discernment difficulty by age—adults aged 25 to 34 years, 35 to 44 years, and 55 to 64 years were less likely to report high discernment difficulty compared to those aged 18 to 24 years, whereas adults aged 45 to 54 years and those aged ≥65 years did not differ significantly from the youngest age group. Hispanic individuals (vs non-Hispanic White individuals; aOR 0.620, 95% CI 0.462-0.831) were less likely to report high

discernment difficulty. Those with lower (vs higher) subjective digital literacy (aOR 1.873, 95% CI 1.478-2.374) or lower (vs higher) subjective numeracy (aOR 1.465, 95% CI 1.047-2.049) were more likely to report high discernment difficulty. The associations between perceived discernment difficulty and sex, educational level, income, geographic residence, and health literacy were not statistically significant.

Associations Between Social Media Health Misinformation Perceptions and Communication Behaviors

After adjusting for sociodemographic characteristics and literacy measures, individuals who perceived high (vs low) levels of social media misinformation were less likely to report sharing general health information on social media (aOR 0.742, 95% CI 0.568-0.968), using social media information to make health decisions (aOR 0.273, 95% CI 0.156-0.479), and using social media information in discussions with health care providers

(aOR 0.460, 95% CI 0.323-0.655). Perceived misinformation amount was not significantly associated with sharing personal health information on social media (Table 2).

Individuals with high (vs low) perceived discernment difficulty were more likely to report using information from social media to make health decisions (aOR 1.724, 95% CI 1.208-2.460) and in discussions with health care providers (aOR 1.389, 95% CI 1.035-1.864). Perceived discernment difficulty was not significantly associated with sharing personal or general health information on social media.

Table 2. Adjusted odds ratio (aOR) and 95% CI of health information sharing and social media information use by social media health misinformation perceptions^a.

Social media misinformation perception	Sharing personal health information ^b (n=4136), aOR (95% CI)	Sharing general health information ^c (n=4159), aOR (95% CI)	Using social media information to make health decisions ^d (n=4177), aOR (95% CI)	Using social media information in discussions with health care providers ^e (n=4174), aOR (95% CI)
High perceived misinformation amount ^f	0.803 (0.591-1.092)	<i>0.742 (0.568-0.968)^g</i>	0.273 (0.156-0.479)	0.460 (0.323-0.655)
High perceived discernment difficulty ^h	1.163 (0.862-1.570)	1.100 (0.878-1.379)	1.724 (1.208-2.460)	1.389 (1.035-1.864)

^aAnalyses were adjusted for age, sex, educational level, race or ethnicity, income, geographic residence, health literacy, digital literacy, and numeracy.

^bThe probability modeled was odds of having ever shared personal information on social media (*shared almost every day, at least once a week, a few times a month, or less than once a month in the past 12 months*) in reference to having never shared.

^cThe probability modeled was odds of having ever shared general information on social media (*shared almost every day, at least once a week, a few times a month, or less than once a month in the past 12 months*) in reference to having never shared.

^dThe probability modeled was odds of using social media information for making health decisions (*strongly agree or somewhat agree*) in reference to not using social media information for making health decisions (*strongly disagree or somewhat disagree*).

^eThe probability modeled was odds of using social media information in discussions with health care providers (*strongly agree or somewhat agree*) in reference to not using social media information in discussions with health care providers (*strongly disagree or somewhat disagree*).

^fHigh perceived misinformation amount=thinking that *a lot* of the health information on social media is false or misleading; low perceived misinformation amount=thinking that *none, a little, or some* of the health information on social media is false or misleading.

^gItalicized values are statistically significant ($P<.05$).

^hHigh perceived discernment difficulty=strongly or somewhat agreeing that it is hard to tell whether health information on social media is true or false; low perceived discernment difficulty=strongly or somewhat disagreeing that it is hard to tell whether health information on social media is true or false.

Discussion

Principal Findings

This study examined 2 misinformation-related perceptions among social media users (perception of the amount of health misinformation on social media and perceived ability to distinguish true from false health information on social media) to better understand the prevalence of these perceptions, subgroup differences in these perceptions, and how these perceptions are related to health communication behaviors. The study found that over one-third of social media users perceived their information environment to contain “a lot” of misleading or false content, and two-thirds expressed difficulty discerning true from false information on social media, with significant variation in these perceptions by sociodemographic characteristics and self-reported literacy skills. The analysis also showed that perceiving a high amount of misinformation on social media was related to lower information sharing on social media and lower use of social media information in discussions with providers and in health decisions, whereas difficulty distinguishing true from false information was

associated with higher use of social media information in discussions with providers and health decisions. These results suggest that understanding misinformation perceptions could help inform health communication interventions and efforts to mitigate the impact of web-based misinformation, and that different approaches may be needed in response to each of these misinformation perceptions.

A substantial proportion of American social media users reported that “a lot” of the health information they see on social media is false or misleading, and this perception varied by race and ethnicity, as well as subjective measures of literacy. Non-Hispanic Black/African American and Hispanic individuals were less likely to say that “a lot” of the health information they see on social media is false or misleading. Because this analysis relied on self-report measures, it is not possible to ascertain whether minority groups are actually less exposed to social media misinformation (eg, due to the nature of their web-based networks) or if they are less aware that the information they are seeing is, in fact, false. However, the reasons behind these differences in misinformation perceptions and the potential for these differences to exacerbate health disparities deserve further

attention given that Black and Hispanic individuals use social media at higher rates than White individuals [1] and substantial proportions of individuals in these groups report regularly obtaining their news from social media platforms [29]. Additional research that attempts to triangulate user perceptions with the social media content they encounter [30] could help shed light on the unique impact of objective and subjective social media experiences.

The analysis also found that individuals with lower digital literacy were more likely to report that “a lot” of the health information they see on social media is false or misleading, whereas those with lower health literacy were less likely to do so. This may be because individuals who self-report low confidence in their ability to find helpful resources on the web are more aware of content quality issues on the internet (including on social media), whereas reporting low subjective health literacy (eg, expressing difficulty filling out medical forms) may not be similarly related to concerns about the online information environment. In fact, a small study conducted in Europe found that participants with low health literacy (as measured using the Newest Vital Sign) had higher scores on the eHealth Literacy Scale, suggesting that they perceived themselves to have higher digital literacy than those in the high health literacy group [31]. The authors hypothesized that this finding might reflect differences in awareness of the issue of web-based health information quality between those with high versus low health literacy as well as differences in knowledge and use of established information evaluation criteria [31].

This study also revealed that approximately two-thirds of American social media users find it hard to tell whether health information on social media is true or false. High levels of discernment difficulty among the public are concerning. Low confidence in one’s ability to distinguish true from false information could result in lower motivation to seek additional information [19], apathy, and confusion, which could lead to negative health outcomes not just because people might act on misinformation but also because they may fail to act on accurate information or adhere to public health recommendations. Research suggests that self-efficacy (ie, judgments regarding how well one can execute a course of action required to deal with a prospective situation) plays an important role in how people select and evaluate information in web-based environments [32]. Individuals with higher self-efficacy may be better able to make accurate credibility assessments because they are more motivated to engage in deep cognitive processing and critical thinking [32], whereas those with lower self-efficacy may avoid engaging in extensive evaluations of information credibility, especially in contexts characterized by uncertainty and ambiguity, as they may not feel that they have a high likelihood of achieving desirable outcomes and, therefore, may experience negative affect (eg, anxiety, frustration, and confusion) in response to these situations [32]. However, although some research suggests that confidence in one’s ability to spot misinformation is associated with better performance in accurately distinguishing false from accurate news [32], the evidence is somewhat limited, and further research combining both subjective perceptions of ability and objective measures of ability is needed in order to investigate the impact of

confidence on the way in which individuals navigate health information on social media.

Beyond generally high rates of discernment difficulty, this analysis also identified differences in perceived discernment ability in certain demographic subgroups. Specifically, adults aged 25 to 34 years, 35 to 44 years, and 55 to 64 years were less likely than those in the youngest age group to report discernment difficulty, and Hispanic individuals reported less discernment difficulty compared to non-Hispanic White individuals. Higher confidence in discernment ability among these groups could be justified (eg, slightly older adults may be just as technologically savvy as young adults but also have more experience and therefore may be better equipped to make accurate credibility assessments); however, it is also possible that discernment confidence in these groups is misplaced, which would be a cause for concern as it might mean that individuals in these groups are less likely to verify information that might be false (eg, through additional research or by speaking to a health care provider), potentially putting them at greater risk of acting on false information. Additional research is needed to better understand why these groups express higher levels of confidence in their discernment ability.

In contrast, those with lower digital literacy and those with lower numeracy were more likely to report high discernment difficulty. This is perhaps not surprising as research has shown lack of digital literacy to be associated with lower objective ability to successfully judge the accuracy of news stories [21], suggesting that individuals with lower digital literacy may be aware of their limitations in this area. Therefore, digital literacy skills as well as health information evaluation abilities may be important targets for interventions seeking to increase resiliency against misinformation—particularly among more susceptible groups.

The results of this analysis also indicated an association between health misinformation perceptions and distinct behavioral patterns. For example, individuals who perceived high levels of misinformation were less likely to share general health information on social media (perhaps because they are more aware of the problem and are more hesitant to share information that could be false), whereas self-reported discernment difficulty was not similarly associated with sharing behaviors on social media. This finding is in line with the results of previous research showing that confidence in one’s ability to identify factually incorrect information is not significantly associated with likelihood of sharing misinformation [32,33]. Some studies suggest that accuracy may not be the most important factor that people consider when making sharing decisions [34,35], which may help explain why uncertainty about the veracity of information does not significantly influence sharing behavior. Research has shown that interventions that prime individuals to consider accuracy when making sharing decisions on social media could be a promising way to mitigate the spread of misinformation [35], and the results of this analysis suggest that this strategy deserves further attention.

Additionally, the results of this study showed that individuals who perceived high levels of misinformation were less likely to use social media information in making health decisions or

in discussions with health care providers. It is possible that, because these individuals perceive high amounts of health misinformation on social media, they are skeptical of the information they encounter on these platforms and, therefore, do not rely on it to inform either their conversations with health care providers or their health decision-making. In contrast, individuals who reported difficulty distinguishing between true and false information on social media were more likely to use information from these platforms in making health decisions and in discussions with health care providers, perhaps because they seek assistance from their health care providers in assessing the credibility of the information. These findings were somewhat counterintuitive, and future research exploring *how* and *why* individuals who report high discernment difficulty use the information they encounter on social media in health-related decisions and discussions could help provide important insights that are beyond the scope of this analysis. For example, exploring whether these individuals are asking for clarification about social media information in discussions with providers versus seeking a “second opinion” on social media after speaking to their clinicians would provide important context to these findings and could help inform how providers can best communicate with patients about information obtained from social media.

Significance

This study offers a unique contribution to our understanding of social media misinformation by focusing on perceptions of the issue rather than objective assessments of misinformation prevalence, exposure, endorsement, or discernment. Assessing perceptions is important because perception of widespread misinformation on social media, as well as perceptions of personal ability to navigate misinformation in web-based spaces, can affect attitudes and behaviors—over and above the impact of actual exposure or ability [16]. In fact, individuals who report high perceived misinformation are likely less susceptible to the direct effects of misinformation (as individuals who characterize a claim as “misinformation” are unlikely to accept it or act on it); however, as demonstrated in this study as well as in previous work, misinformation perceptions can still shape their responses and behaviors [16].

Notably, there are limitations to using self-reported measures of perception—for example, it is impossible to know whether people’s perceptions are an accurate reflection of “the ground truth” (ie, whether a lot of the social media information they are exposed to really is or is not false and whether they are really capable of discerning the veracity of social media information) [20]. However, there is still value in assessing these perceptions to obtain a high-level understanding of the public’s views on the scope of the problem and the extent to which it affects them as well as their judgment of their own capacity to cope with the problem. In the context of political misinformation, individuals who perceived a lot of exposure to misinformation were more likely to believe that misinformation is a serious problem that creates a lot of confusion about the basic facts of current issues and events and were also more confident in their ability to identify misinformation [20]. Perceptions of the information environment can also impact attitudes and behaviors in ways that are important to health [19]—for example, people may feel

overwhelmed and discouraged from seeking additional information about a health topic or develop inaccurate risk perceptions. Additionally, the differences in misinformation perceptions by demographics and literacy levels identified in this study are concerning as they threaten to increase disparities among vulnerable populations. However, while perceptions are important in and of themselves, future research could benefit from including both subjective and objective measures of the information environment to better understand the unique contribution of each construct and provide a more comprehensive understanding of how people respond to social media information.

Nonetheless, the results of this study suggest several practical measures that could help mitigate the impact of misinformation on social media. First, they point to specific populations that may benefit from targeted interventions. For example, those who perceived “a lot” of misinformation on social media were less likely to use this information in health decision-making, suggesting that interventions that raise awareness of information quality issues on social media could limit the extent to which individuals rely on questionable information from social media to make health decisions. These efforts may be especially impactful among groups who are less likely to report perceiving high amounts of misinformation on social media (eg, Black/African American and Hispanic individuals and individuals with lower health literacy). Additionally, the finding that those who express high discernment difficulty still use information from social media to make health decisions suggests that these individuals should be targeted for training interventions that can increase their ability to discern misinformation to (1) ensure that they are relying on accurate information to make these decisions and (2) increase their confidence in their ability to navigate the social media information environment. For example, instructional programs that train people to recognize misinformation techniques have been shown to increase their awareness of these tactics as well as confidence in their ability to successfully deal with misinformation [36]. Furthermore, increasing confidence in discernment ability may itself be a viable target for encouraging careful evaluation of information and increasing resilience to misinformation. For example, Ferrucci and Hopp [37] found that a short intervention providing positive verbal persuasion regarding participants’ ability to identify false information on social media increased fake news self-efficacy and that higher self-efficacy beliefs were in turn associated with ability to correctly classify both credible and “fake” news headlines in an information accuracy assessment task.

Second, the finding that those who express high discernment difficulty are more likely to have discussions with health care providers regarding social media health information suggests a need for training aimed at providers to support them in effectively helping patients navigate the information they encounter on the web (eg, teaching providers how to invite these conversations, address misinformation with empathy, and empower patients by recommending accurate sources of information) [38,39]. Research suggests that providers rarely initiate conversations about web-based health information seeking with patients [40,41], but the results of this study

indicate that asking about patients' perceptions and use of social media health information could be helpful to incorporate into patient-provider conversations.

Although this study looks at individual-level perceptions and has implications for individual-level interventions (eg, increasing digital literacy), the onus should not be solely on individuals (or providers) to address the problem of social media misinformation. Social media platforms could take steps to decrease the amount of misinformation that users are exposed to in the first place and make it easier for them to discern true from false information (eg, through the use of fact-checking labels and account verification). However, in the absence of these types of more systematic changes in the social media environment, individuals will likely be left to navigate the increasingly confusing information landscape on their own and will need to be supported in their efforts, for example, through campaigns to raise awareness of the issue (particularly among vulnerable populations), training on information evaluation strategies and common misinformation techniques, and encouragement to discuss social media health information with providers and others with relevant expertise. These interventions can be deployed in both web-based and offline contexts (eg, through video advertisements on social media platforms [42] or through educational services delivered in health care settings [43]).

Limitations

This study has several limitations. First, the cross-sectional nature of HINTS data precludes causal inferences from being drawn about observed relationships between variables. Second, the misinformation measures included in this analysis are subjective perception items. As such, there is no way to determine the objective truth about a respondent's actual misinformation exposure or their ability to differentiate true from false information. However, even if they do not reflect objective reality, perceptions are valuable to assess because they enable a better understanding of the public's views on the scope of the misinformation problem and their capacity to cope with it and can help shed light on the way in which perceptions of the information environment shape health-related attitudes and behaviors. Third, the lack of information on certain aspects of respondents' social media experiences and behaviors (eg, the specific platforms they use) is a limitation of this analysis—and reflects a disadvantage of using a national health communication survey that includes only a limited number of items regarding

social media use due to space constraints. Finally, the response rate for HINTS 6 (28.1%) was relatively low, which may introduce bias into the data [44]. However, methodological research suggests that the impact of low response rates on data quality may be less significant than previously assumed [44]. Despite these limitations, this analysis provides an important contribution to the broader health misinformation literature as there has been limited research to date focusing on perceptions of misinformation, particularly outside the context of COVID-19.

Conclusions

Many social media users in the United States perceive high levels of misinformation on social media and report difficulty discerning true from false information. This is concerning because perceptions of high misinformation prevalence could increase negative affect (eg, anxiety and worry) regarding health issues, whereas low discernment confidence could result in apathy, confusion, and lower motivation to seek additional information. The fact that health misinformation perceptions were found to vary across race, ethnicity, age, and literacy levels may suggest a need to raise awareness about misinformation and provide training for certain populations (eg, those with low health literacy) to ensure that they approach the information environment with sufficient skepticism and are better able to verify the health claims they see on social media. Finally, the associations between misinformation perceptions and social media-related communication behaviors found in this study can help inform future research as well as health communication interventions and misinformation mitigation efforts. For example, the finding that individuals who have low confidence in their discernment ability are more likely to use social media information to make health decisions and in discussions with health care providers suggests that they may benefit from providers assisting them in navigating and verifying web-based information.

Although a growing body of literature focusing on social media misinformation has emerged in recent years, to date, very little work has been done to look at subjective assessments of the problem of misinformation. This study provides initial insights into the prevalence, disparities, and potential impact of social media misinformation perceptions. However, more research is needed to understand how perceptions of misinformation affect the public's health-related cognitions, attitudes, communication behaviors, and outcomes.

Disclaimer

Any opinions expressed by the authors are their own, and this material should not be interpreted as representing the official viewpoint of the US Department of Health and Human Services, the National Institutes of Health, or the National Cancer Institute.

Conflicts of Interest

None declared.

References

1. Auxier B, Anderson M. Social media use in 2021. Pew Research Center. 2021 Apr 07. URL: <https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/> [accessed 2023-12-15]

2. Chen J, Wang Y. Social media use for health purposes: systematic review. *J Med Internet Res* 2021 May 12;23(5):e17917 [FREE Full text] [doi: [10.2196/17917](https://doi.org/10.2196/17917)] [Medline: [33978589](https://pubmed.ncbi.nlm.nih.gov/33978589/)]
3. Zhao Y, Zhang J. Consumer health information seeking in social media: a literature review. *Health Info Libr J* 2017 Dec;34(4):268-283 [FREE Full text] [doi: [10.1111/hir.12192](https://doi.org/10.1111/hir.12192)] [Medline: [29045011](https://pubmed.ncbi.nlm.nih.gov/29045011/)]
4. Chou WYS, Gaysynsky A, Trivedi N, Vanderpool RC. Using social media for health: national data from HINTS 2019. *J Health Commun* 2021 Mar 04;26(3):184-193. [doi: [10.1080/10810730.2021.1903627](https://doi.org/10.1080/10810730.2021.1903627)] [Medline: [33856286](https://pubmed.ncbi.nlm.nih.gov/33856286/)]
5. Chou WYS, Gaysynsky A, Cappella JN. Where we go from here: health misinformation on social media. *Am J Public Health* 2020 Oct;110(S3):S273-S275. [doi: [10.2105/AJPH.2020.305905](https://doi.org/10.2105/AJPH.2020.305905)] [Medline: [33001722](https://pubmed.ncbi.nlm.nih.gov/33001722/)]
6. Confronting health misinformation: the U.S. surgeon general's advisory on building a healthy information environment. US Department of Health and Human Services. 2021. URL: https://www.ncbi.nlm.nih.gov/books/NBK572169/pdf/Bookshelf_NBK572169.pdf [accessed 2023-12-15]
7. Treen KM, Williams HT, O'Neill SJ. Online misinformation about climate change. *Wiley Interdiscip Rev Clim Change* 2020 Jun 18;11(5):e665. [doi: [10.1002/wcc.665](https://doi.org/10.1002/wcc.665)]
8. Chou WY, Gaysynsky A, Lama Y. The health misinformation ecosystem on social media: emerging evidence and research gaps. In: Keselman A, Smith CA, Wilson AJ, editors. *Combating Online Health Misinformation: A Professional's Guide to Helping the Public*. Lanham, MD: Rowman & Littlefield; Apr 01, 2023:31-44.
9. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. *J Med Internet Res* 2021 Jan 20;23(1):e17187 [FREE Full text] [doi: [10.2196/17187](https://doi.org/10.2196/17187)] [Medline: [33470931](https://pubmed.ncbi.nlm.nih.gov/33470931/)]
10. Pierri F, Perry BL, DeVerna MR, Yang KC, Flammini A, Menczer F, et al. Online misinformation is linked to early COVID-19 vaccination hesitancy and refusal. *Sci Rep* 2022 Apr 26;12(1):5966. [doi: [10.1038/s41598-022-10070-w](https://doi.org/10.1038/s41598-022-10070-w)] [Medline: [35474313](https://pubmed.ncbi.nlm.nih.gov/35474313/)]
11. Loomba S, de Figueiredo A, Piatek SJ, de Graaf K, Larson HJ. Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. *Nat Hum Behav* 2021 Mar;5(3):337-348. [doi: [10.1038/s41562-021-01056-1](https://doi.org/10.1038/s41562-021-01056-1)] [Medline: [33547453](https://pubmed.ncbi.nlm.nih.gov/33547453/)]
12. Americans' views of misinformation in the news and how to counteract it. Gallup Inc. 2018. URL: <https://knightfoundation.org/reports/americans-views-of-misinformation-in-the-news-and-how-to-counteract-it/> [accessed 2023-12-15]
13. Shearer E, Matsa KE. News use across social media platforms 2018. Pew Research Center. 2018 Sep 10. URL: <https://www.pewresearch.org/journalism/2018/09/10/news-use-across-social-media-platforms-2018/> [accessed 2023-12-15]
14. Müller P, Schulz A. Facebook or Fakebook? How users' perceptions of 'fake news' are related to their evaluation and verification of news on Facebook. *Stud Commun Media* 2019;8(4):547-559. [doi: [10.5771/2192-4007-2019-4-547](https://doi.org/10.5771/2192-4007-2019-4-547)]
15. Yang F, Horning M. Reluctant to share: how third person perceptions of fake news discourage news readers from sharing "real news" on social media. *Soc Media Soc* 2020 Sep 28;6(3):205630512095517. [doi: [10.1177/2056305120955173](https://doi.org/10.1177/2056305120955173)]
16. Matthes J, Corbu N, Jin S, Theocharis Y, Schemer C, van Aelst P, et al. Perceived prevalence of misinformation fuels worries about COVID-19: a cross-country, multi-method investigation. *Inf Commun Soc* 2022 Nov 29;26(16):3133-3156. [doi: [10.1080/1369118x.2022.2146983](https://doi.org/10.1080/1369118x.2022.2146983)]
17. Hinsley A, Ju I, Park T, Ohs J. Credibility in the time of COVID-19: cues that audiences look for when assessing information on social media and building confidence in identifying 'fake news' about the virus. *Open Inf Sci* 2022;6(1):61-73. [doi: [10.1515/opis-2022-0132](https://doi.org/10.1515/opis-2022-0132)]
18. Roozenbeek J, Maertens R, Herzog SM, Geers M, Kurvers R, Sultan M, et al. Susceptibility to misinformation is consistent across question framings and response modes and better explained by myside bias and partisanship than analytical thinking. *Judgm Decis Mak* 2023 Jan 01;17(3):547-573. [doi: [10.1017/s1930297500003570](https://doi.org/10.1017/s1930297500003570)]
19. Park YJ, Chung JE, Kim JN. Social media, misinformation, and cultivation of informational mistrust: cultivating COVID-19 mistrust. *Journalism* 2022 Apr 08;23(12):2571-2590. [doi: [10.1177/14648849221085050](https://doi.org/10.1177/14648849221085050)]
20. Barthel M, Mitchell A, Holcomb J. Many Americans believe fake news is sowing confusion. Pew Research Center. 2016 Dec 15. URL: <https://www.pewresearch.org/journalism/2016/12/15/many-americans-believe-fake-news-is-sowing-confusion/> [accessed 2023-12-15]
21. Sirlin N, Epstein Z, Arechar AA, Rand DG. Digital literacy is associated with more discerning accuracy judgments but not sharing intentions. *Misinformation Review* 2021 Dec 06;2(6):1-13 [FREE Full text] [doi: [10.37016/mr-2020-83](https://doi.org/10.37016/mr-2020-83)]
22. Nan X, Wang Y, Thier K. Why do people believe health misinformation and who is at risk? A systematic review of individual differences in susceptibility to health misinformation. *Soc Sci Med* 2022 Dec;314:115398. [doi: [10.1016/j.socscimed.2022.115398](https://doi.org/10.1016/j.socscimed.2022.115398)] [Medline: [36327631](https://pubmed.ncbi.nlm.nih.gov/36327631/)]
23. HINTS 6 methodology report. Westat. 2023. URL: https://hints.cancer.gov/docs/methodologyreports/HINTS_6_MethodologyReport.pdf [accessed 2023-12-15]
24. Chew LD, Bradley KA, Boyko EJ. Brief questions to identify patients with inadequate health literacy. *Fam Med* 2004 Sep;36(8):588-594 [FREE Full text] [Medline: [15343421](https://pubmed.ncbi.nlm.nih.gov/15343421/)]
25. Norman CD, Skinner HA. eHEALS: the eHealth literacy scale. *J Med Internet Res* 2006 Nov 14;8(4):e27 [FREE Full text] [doi: [10.2196/jmir.8.4.e27](https://doi.org/10.2196/jmir.8.4.e27)] [Medline: [17213046](https://pubmed.ncbi.nlm.nih.gov/17213046/)]
26. Woloshin S, Schwartz LM, Welch HG. Patients and medical statistics: interest, confidence, and ability. *J Gen Intern Med* 2005 Nov;20(11):996-1000 [FREE Full text] [doi: [10.1111/j.1525-1497.2005.00179.x](https://doi.org/10.1111/j.1525-1497.2005.00179.x)] [Medline: [16307623](https://pubmed.ncbi.nlm.nih.gov/16307623/)]

27. Stagliano V, Wallace LS. Brief health literacy screening items predict newest vital sign scores. *J Am Board Fam Med* 2013;26(5):558-565 [FREE Full text] [doi: [10.3122/jabfm.2013.05.130096](https://doi.org/10.3122/jabfm.2013.05.130096)] [Medline: [24004707](https://pubmed.ncbi.nlm.nih.gov/24004707/)]
28. Data. National Cancer Institute: Health Information National Trends Survey. 2023. URL: <https://hints.cancer.gov/data/Default.aspx> [accessed 2023-12-15]
29. Appendix: demographic profiles and party identification of regular social media news consumers in the United States. Pew Research Center. 2023 Nov 15. URL: <https://www.pewresearch.org/journalism/2023/11/15/appendix-demographic-and-party-identification-of-regular-social-media-news-consumers-in-the-u-s/> [accessed 2023-12-15]
30. Rivera YM, Moran MB, Thrul J, Joshu C, Smith KC. Contextualizing engagement with health information on Facebook: using the social media content and context elicitation method. *J Med Internet Res* 2022 Mar 04;24(3):e25243 [FREE Full text] [doi: [10.2196/25243](https://doi.org/10.2196/25243)] [Medline: [35254266](https://pubmed.ncbi.nlm.nih.gov/35254266/)]
31. Diviani N, van den Putte B, Meppelink CS, van Weert JC. Exploring the role of health literacy in the evaluation of online health information: insights from a mixed-methods study. *Patient Educ Couns* 2016 Jun;99(6):1017-1025. [doi: [10.1016/j.pec.2016.01.007](https://doi.org/10.1016/j.pec.2016.01.007)] [Medline: [26817407](https://pubmed.ncbi.nlm.nih.gov/26817407/)]
32. Hopp T. Fake news self-efficacy, fake news identification, and content sharing on Facebook. *J Inf Technol Politics* 2021 Aug 12;19(2):229-252. [doi: [10.1080/19331681.2021.1962778](https://doi.org/10.1080/19331681.2021.1962778)]
33. Paciello M, Corbelli G, D'Errico F. The role of self-efficacy beliefs in dealing with misinformation among adolescents. *Front Psychol* 2023;14:1155280 [FREE Full text] [doi: [10.3389/fpsyg.2023.1155280](https://doi.org/10.3389/fpsyg.2023.1155280)] [Medline: [37275715](https://pubmed.ncbi.nlm.nih.gov/37275715/)]
34. Pennycook G, Epstein Z, Mosleh M, Arechar AA, Eckles D, Rand DG. Shifting attention to accuracy can reduce misinformation online. *Nature* 2021 Apr;592(7855):590-595. [doi: [10.1038/s41586-021-03344-2](https://doi.org/10.1038/s41586-021-03344-2)] [Medline: [33731933](https://pubmed.ncbi.nlm.nih.gov/33731933/)]
35. Pennycook G, McPhetres J, Zhang Y, Lu JG, Rand DG. Fighting COVID-19 misinformation on social media: experimental evidence for a scalable accuracy-nudge intervention. *Psychol Sci* 2020 Jul;31(7):770-780 [FREE Full text] [doi: [10.1177/0956797620939054](https://doi.org/10.1177/0956797620939054)] [Medline: [32603243](https://pubmed.ncbi.nlm.nih.gov/32603243/)]
36. Rapp DN, Withall MM. Confidence as a metacognitive contributor to and consequence of misinformation experiences. *Curr Opin Psychol* 2024 Feb;55:101735. [doi: [10.1016/j.copsyc.2023.101735](https://doi.org/10.1016/j.copsyc.2023.101735)] [Medline: [38041918](https://pubmed.ncbi.nlm.nih.gov/38041918/)]
37. Ferrucci P, Hopp T. Let's intervene: how platforms can combine media literacy and self-efficacy to fight fake news. *Commun Public* 2023 Oct 31;8(4):367-389. [doi: [10.1177/20570473231203081](https://doi.org/10.1177/20570473231203081)]
38. Wood JL, Lee GY, Stinnett SS, Southwell BG. A pilot study of medical misinformation perceptions and training among practitioners in North Carolina (USA). *Inquiry* 2021;58:469580211035742 [FREE Full text] [doi: [10.1177/00469580211035742](https://doi.org/10.1177/00469580211035742)] [Medline: [34399597](https://pubmed.ncbi.nlm.nih.gov/34399597/)]
39. Southwell BG, Wood JL, Navar AM. Roles for health care professionals in addressing patient-held misinformation beyond fact correction. *Am J Public Health* 2020 Oct;110(S3):S288-S289. [doi: [10.2105/AJPH.2020.305729](https://doi.org/10.2105/AJPH.2020.305729)] [Medline: [33001723](https://pubmed.ncbi.nlm.nih.gov/33001723/)]
40. Gantenbein L, Navarini AA, Maul LV, Brandt O, Mueller SM. Internet and social media use in dermatology patients: search behavior and impact on patient-physician relationship. *Dermatol Ther* 2020 Nov;33(6):e14098. [doi: [10.1111/dth.14098](https://doi.org/10.1111/dth.14098)] [Medline: [32725746](https://pubmed.ncbi.nlm.nih.gov/32725746/)]
41. Shen MJ, Dyson RC, D'Agostino TA, Ostroff JS, Dickler MN, Heerdt AS, et al. Cancer-related internet information communication between oncologists and patients with breast cancer: a qualitative study. *Psychooncology* 2015 Nov;24(11):1439-1447 [FREE Full text] [doi: [10.1002/pon.3752](https://doi.org/10.1002/pon.3752)] [Medline: [25631285](https://pubmed.ncbi.nlm.nih.gov/25631285/)]
42. Roozenbeek J, van der Linden S, Goldberg B, Rathje S, Lewandowsky S. Psychological inoculation improves resilience against misinformation on social media. *Sci Adv* 2022 Aug 26;8(34):eabo6254 [FREE Full text] [doi: [10.1126/sciadv.abo6254](https://doi.org/10.1126/sciadv.abo6254)] [Medline: [36001675](https://pubmed.ncbi.nlm.nih.gov/36001675/)]
43. Smith C, Martin-Lillie C, Higano JD, Turner L, Phu S, Arthurs J, et al. Challenging misinformation and engaging patients: characterizing a regenerative medicine consult service. *Regen Med* 2020 Mar;15(3):1427-1440 [FREE Full text] [doi: [10.2217/rme-2020-0018](https://doi.org/10.2217/rme-2020-0018)] [Medline: [32319855](https://pubmed.ncbi.nlm.nih.gov/32319855/)]
44. Blake KD, Ottenbacher AJ, Finney Rutten LJ, Grady MA, Kobrin SC, Jacobson RM, et al. Predictors of human papillomavirus awareness and knowledge in 2013: gaps and opportunities for targeted communication strategies. *Am J Prev Med* 2015 Apr;48(4):402-410 [FREE Full text] [doi: [10.1016/j.amepre.2014.10.024](https://doi.org/10.1016/j.amepre.2014.10.024)] [Medline: [25700651](https://pubmed.ncbi.nlm.nih.gov/25700651/)]

Abbreviations

aOR: adjusted odds ratio

HINTS: Health Information National Trends Survey

Edited by T Mackey; submitted 21.07.23; peer-reviewed by S Ganesh, B Hoffman; comments to author 28.11.23; revised version received 15.12.23; accepted 20.03.24; published 30.04.24.

Please cite as:

Gaysynsky A, Senft Everson N, Heley K, Chou WYS

Perceptions of Health Misinformation on Social Media: Cross-Sectional Survey Study

JMIR Infodemiology 2024;4:e51127

URL: <https://infodemiology.jmir.org/2024/1/e51127>

doi: [10.2196/51127](https://doi.org/10.2196/51127)

PMID: [38687591](https://pubmed.ncbi.nlm.nih.gov/38687591/)

©Anna Gaysynsky, Nicole Senft Everson, Kathryn Heley, Wen-Ying Sylvia Chou. Originally published in JMIR Infodemiology (<https://infodemiology.jmir.org>), 30.04.2024. 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.

Original Paper

TikTok as a Source of Health Information and Misinformation for Young Women in the United States: Survey Study

Ciera E Kirkpatrick¹, PhD; LaRissa L Lawrie², MA

¹College of Journalism & Mass Communications, University of Nebraska-Lincoln, Lincoln, NE, United States

²School of Journalism, University of Missouri, Columbia, MO, United States

Corresponding Author:

Ciera E Kirkpatrick, PhD

College of Journalism & Mass Communications

University of Nebraska-Lincoln

331 Andersen Hall

Lincoln, NE, 68588

United States

Phone: 1 3165594397

Email: ciera.kirkpatrick@unl.edu

Abstract

Background: TikTok is one of the most-used and fastest-growing social media platforms in the world, and recent reports indicate that it has become an increasingly popular source of news and information in the United States. These trends have important implications for public health because an abundance of health information exists on the platform. Women are among the largest group of TikTok users in the United States and may be especially affected by the dissemination of health information on TikTok. Prior research has shown that women are not only more likely to look for information on the internet but are also more likely to have their health-related behaviors and perceptions affected by their involvement with social media.

Objective: We conducted a survey of young women in the United States to better understand their use of TikTok for health information as well as their perceptions of TikTok's health information and health communication sources.

Methods: A web-based survey of US women aged 18 to 29 years (N=1172) was conducted in April-May 2023. The sample was recruited from a Qualtrics research panel and 2 public universities in the United States.

Results: The results indicate that the majority of young women in the United States who have used TikTok have obtained health information from the platform either intentionally (672/1026, 65.5%) or unintentionally (948/1026, 92.4%). Age (959/1026, 93.47%; $r=0.30$; $P<.001$), education (959/1026, 93.47%; $\rho=0.10$; $P=.001$), and TikTok intensity (ie, participants' emotional connectedness to TikTok and TikTok's integration into their daily lives; 959/1026, 93.47%; $r=0.32$; $P<.001$) were positively correlated with overall credibility perceptions of the health information. Nearly the entire sample reported that they think that misinformation is prevalent on TikTok to at least some extent (1007/1026, 98.15%), but a third-person effect was found because the young women reported that they believe that other people are more susceptible to health misinformation on TikTok than they personally are ($t_{1025}=21.16$; $P<.001$). Both health professionals and general users were common sources of health information on TikTok: 93.08% (955/1026) of the participants indicated that they had obtained health information from a health professional, and 93.86% (963/1026) indicated that they had obtained health information from a general user. The respondents showed greater preference for health information from health professionals (vs general users; $t_{1025}=23.75$; $P<.001$); the respondents also reported obtaining health information from health professionals more often than from general users ($t_{1025}=8.13$; $P<.001$), and they were more likely to act on health information from health professionals (vs general users; $t_{1025}=12.74$; $P<.001$).

Conclusions: The findings suggest that health professionals and health communication scholars need to proactively consider using TikTok as a platform for disseminating health information to young women because young women are obtaining health information from TikTok and prefer information from health professionals.

(JMIR Infodemiology 2024;4:e54663) doi:[10.2196/54663](https://doi.org/10.2196/54663)

KEYWORDS

credibility perceptions; health information; health misinformation; information seeking; misinformation perceptions; public health; social media; strategic communication; third-person effect; TikTok

Introduction

Background

As one of the most-used and fastest-growing social media platforms in the world, TikTok has drastically changed American culture [1,2]. The social media platform, which allows users to create and watch short-form videos ranging in length from 15 seconds to 10 minutes, has >150 million active users in the United States and is expected to reach 955 million users worldwide by 2025 [3,4]. This rapid popularity has caught the attention of health communication scholars and practitioners because the platform is a vehicle for finding and disseminating information, including health-related content [5,6]. According to the Pew Research Center, the number of adults in the United States who regularly get news from TikTok has more than tripled (from 3% in 2020 to 10% in 2022). Adults aged <30 years are the most likely group, with a third (32%) of adults aged 18 to 29 years saying that they regularly get their news from TikTok [7].

Americans turned to the internet to find health information during the COVID-19 pandemic, and medical professionals and health institutions met them on TikTok, delivering pandemic-related information and recommendations [5]. The presence of health-related content has extended beyond COVID-19-related information and includes a wide range of health topics such as cervical cancer screening, chronic pulmonary obstructive disease, diabetes, mental health, and more [8-11]. Studies examining the engagement with this content [10,11] suggest that users like finding health information on the platform. A recent survey of 2000 Americans conducted by the prescription savings company CharityRx found that 1 in 5 Americans turns to TikTok for advice before going to their physician. Of the participants belonging to Generation Z, specifically, 1 in 3 reported obtaining health information from TikTok [12].

TikTok's concise video format encourages users to convey their message in a brief yet engaging way, while the relaxed atmosphere and the capacity to engage with viewers facilitate a more intimate and authentic form of communication [13]. These videos may be especially memorable (and thus influential) because viewers can retrieve both visual and verbal information that they have processed and stored while watching the videos [14]. Social media platforms have the power to spread credible, useful health information [8]. However, recent research has indicated that 1 in 5 TikTok videos likely contains misinformation [15], and fact-checking has been found to be uncommon on social media [16]. Consequently, as young women turn to TikTok for health information, they may encounter both the beneficial aspects and drawbacks of this accessible platform. For these reasons, we suggest the need for a better empirical understanding of the extent to which young women are obtaining health information on TikTok and their associated perceptions and behaviors related to the information they encounter.

Objectives

In light of the popularity of TikTok, we conducted a survey of young women (assigned female sex at birth) in the United States

to better understand their use of the social media platform for health information as well as their perceptions of the platform's health information and related communication sources. We focus on women aged 18 to 29 years for this study because women make up the majority of TikTok users in the United States and because survey data have revealed that users aged 18 to 19 years and 20 to 29 years were the 2 largest age groups using TikTok during the time period that this study was conducted [17,18]. Furthermore, it has been suggested that women may be especially affected by health information on TikTok. Mainstream news programs (eg, Good Morning America on the ABC network) have anecdotally reported that many young women turn to TikTok for health information and that obstetrician-gynecologists and other physicians have developed TikTok brands specifically to reach these users [6]. According to prior research, young women's involvement with social media significantly influences their perceptions and behaviors concerning their health [19]; women are more likely to look for health information, including via the internet [20,21]; and women tend to have a leading role in the majority of decisions for their families' health [22].

Specifically, we first sought to explore how often young women are intentionally and unintentionally obtaining health information from TikTok and their top reasons for obtaining health information from the platform. Second, we explored their perceptions of credibility (ie, the perceived credibility of TikTok health information overall) and misinformation in relation to the health information they see on TikTok, as well as their frequency of verifying the health information they see. Within this, we asked questions about perceived susceptibility to health misinformation on TikTok to see if a third-person effect might exist. The third-person effect is a communication theory that suggests that people tend to perceive that messages in the media have a greater effect on other people than on themselves [23], which, in the context of misinformation on TikTok, could cause young women to underestimate the potential impact of misinformation on their own health-related decisions and behaviors. Third, we explored perceptions and behaviors related to the top 2 types of sources that share health information on TikTok (*health professionals* and *general users*). In terms of perceptions, we examined how often young women obtain information from these sources, how much they prefer to obtain information from these sources, and how credible they perceive the information from these sources to be (ie, the perceived credibility of TikTok health information from health professionals and the perceived credibility of TikTok health information from general users). For behaviors related to these source types, we examined whether the young women have acted on health information they obtained from these sources, their likelihood of acting on health information from these sources in the future, and their likelihood of fact-checking information from these sources. In exploring each of these 3 areas, we also examined whether the women's age, highest level of education, and level of TikTok intensity (defined as their emotional connectedness to TikTok and TikTok's integration into their daily lives) had a relationship with their use of TikTok as a source of health information.

Altogether, this study seeks to help both health communication researchers and practitioners by illuminating the role that TikTok plays in young women's acquisition of health information in the United States.

Methods

Design and Sample

A web-based survey of US women (assigned female sex at birth) aged 18 to 29 years (N=1172) was conducted between April and May 2023. While the sample was focused on individuals who were assigned female sex at birth, we also asked about their gender identity (refer to the Results section). The sample was recruited using a Qualtrics research panel as well as convenience sampling at 2 public universities in the United States.

Ethical Considerations

The institutional review boards at the University of Nebraska-Lincoln (IRB 20230122526EX) and the University of Missouri (IRB 2095651) approved the study. Respondents recruited via the Qualtrics panel were compensated in agreement with their Qualtrics contract, and the respondents recruited at the universities were compensated with course credit.

Procedure

The study began with a web-based informed consent form that briefly explained the purpose of the study and gave the survey respondents information about the study's investigator, the expected length of the survey, and how their data would be used (for the reporting of aggregate data) and stored (in a password-protected electronic format). Respondents then answered screening questions, and respondents who were not assigned female sex at birth and not aged 18 to 29 years were excluded from the survey. After passing the eligibility criteria, respondents were asked whether they had ever used TikTok to either watch or post videos. The respondents who had used TikTok were then asked about their average amount of use, whether they had ever intentionally used TikTok to look for advice or information about their health or health care (and whether they had done so in the past 3 months), and whether they had ever unintentionally been exposed to health information on TikTok. From here, the questions they saw depended upon whether they had ever seen health information (intentionally or unintentionally) on TikTok.

All respondents who had ever used TikTok (n=1026) responded to items measuring reasons for health-related TikTok use; the perceptions of health misinformation on TikTok; the use of, and preference for, particular sources (health professionals and general users) of health information on TikTok; the perceived credibility of health information from health professionals and general users on TikTok; the likelihood of acting on health information obtained from health professionals and general users on TikTok; and the likelihood of fact-checking health information from health professionals and general users on TikTok.

Respondents who had seen health information on TikTok (n=959) additionally responded to items measuring their

perceived credibility of the health information they have seen overall on TikTok and their verification of the health information they have seen on TikTok. Respondents who had ever received health information on TikTok from either source of interest (health professionals or general users) were also asked about whether they had acted on health information from these sources.

Finally, all participants (N=1172) responded to items measuring TikTok intensity (ie, their emotional connectedness to TikTok and TikTok's integration into their daily lives) and answered demographic questions, including their highest level of education, race, and ethnicity.

Measures

Using TikTok as a Health Information Source

Frequency of Use

Respondents were asked whether they had ever used TikTok (either to watch or to post videos). The respondents who had used TikTok were asked to indicate their average amount of use, using the following options: less than once a month, once a month, once a week, a few times a week, once a day, more often than once a day. The respondents who had ever used TikTok were also asked whether they had ever used TikTok to look for advice or information about their health or health care (selecting *yes* or *no*). Those who selected *yes* were asked how often they *intentionally* use TikTok to obtain health information, and all respondents who had ever used TikTok were asked how often they *unintentionally* obtain health information on TikTok (hourly, daily, weekly, monthly, less often, or not at all).

Reasons for Health-Related TikTok Use

Respondents indicated their reasons for health-related TikTok use by indicating their level of agreement (ranging from 1=*strongly disagree* to 7=*strongly agree*) with 10 different statements (eg, "I like to get health information from TikTok because it can help me to maintain a healthy lifestyle") adapted from prior research [24-26] (refer to the Results section).

Credibility, Misinformation, and Verification of Health Information on TikTok

Perceived Credibility of TikTok Health Information Overall

Using the 7-point scale (ranging from 1=*not at all* to 7=*extremely*) for media credibility developed by Flanagan and Metzger [27], respondents rated how believable, accurate, trustworthy, biased (reverse coded), and complete they perceive health information on TikTok, overall, to be. Specifically, the respondents were asked, "To what degree do you rate the health information provided on TikTok?" The 5 items were averaged to create a perceived credibility score for each respondent (mean 4.48, SD 1.28; Cronbach $\alpha=0.90$).

Misinformation Perceptions

For the perceptions of misinformation, respondents were asked to indicate on a 7-point scale how prevalent they think health misinformation is on TikTok (ranging from 1=*not prevalent at all* to 7=*very prevalent*; mean 5.14, SD 1.42), how serious they think the impact of health misinformation on TikTok is (ranging from 1=*not serious at all* to 7=*very serious*; mean 5.57, SD

1.46), and how susceptible they think they are to the influence of health misinformation on TikTok (ranging from 1=*not susceptible at all* to 7=*very susceptible*; mean 4.07, SD 1.75; adapted from the study by Chang [28]). In addition, to explore the possibility of a third-person effect, the respondents were asked to indicate how susceptible (ranging from 1=*not susceptible at all* to 7=*very susceptible*) they think others (eg, the public; mean 5.26, SD 1.46) are to the influence of health misinformation on TikTok (adapted from the study by van der Meer et al [23]). Participants were also asked about their self-perceived direct experience with health information on TikTok with the following item (adapted from the study by Chang [28]): “Have you encountered health misinformation on TikTok in the past? (*yes, no, or unsure*).”

Verification of Health Information on TikTok

Using a scale adapted from the study by Flanagin and Metzger [27], respondents indicated on a 7-point scale (ranging from 1=*never* to 7=*always*) how often they performed 6 different verification behaviors (eg, “Check to see if the information is current”) when seeing health information on TikTok. The scores for the 6 items were averaged to create a *verification of TikTok health information* score for each respondent (mean 4.83, SD 1.53; Cronbach $\alpha=0.92$).

Health Professionals and General Users as Sources of Health Information on TikTok

Source Preferences

Using a 7-point scale (ranging from 1=*not at all* to 7=*very often*), respondents indicated how often they obtain health information from health professionals on TikTok (mean 5.04, SD 1.83) and how often they obtain health information from general users on TikTok (mean 4.55, SD 1.89). In addition, using a 7-point scale (ranging from 1=*don't prefer them at all* to 7=*prefer them a lot*), respondents were asked to indicate how much they prefer to obtain health information from health professionals on TikTok (mean 5.65, SD 1.75) and how much they prefer to obtain health information from general users on TikTok (mean 4.08, SD 1.96).

Perceived Credibility of TikTok Health Information From Health Professionals and General Users

In addition to measuring the respondents' perceived credibility of TikTok health information *overall*, we also measured the respondents' perceived credibility of the 2 sources of interest (health professionals and general users) using the 7-point scale for media credibility developed by Flanagin and Metzger [27]. Specifically, we asked, “To what degree do you rate the health information provided by *health professionals* (eg, a doctor or nurse) on TikTok?” and “To what degree do you rate the health information provided by *general users* (someone like you) on TikTok?” Respondents rated how believable, accurate, trustworthy, biased (reverse coded), and complete they believe the health information on TikTok to be from each of these sources. The scores for each of the 5 items were averaged for each of the source types such that each respondent had a score for the perceived credibility of TikTok health information from health professionals (mean 5.16, SD 1.18; Cronbach $\alpha=0.90$) and the perceived credibility of TikTok health information from general users (mean 3.95, SD 1.54; Cronbach $\alpha=0.94$).

Acting on Health Information

Respondents' likelihood of acting on health information was measured with items adapted from the study by Hu and Shyam Sundar [29]. Using a 7-point scale (ranging from 1=*not at all likely* to 7=*extremely likely*), respondents indicated how likely they are to act on health information from a health professional on TikTok (mean 4.50, SD 1.79) and from a general user on TikTok (mean 3.96, SD 1.89). Respondents were also asked whether they ever have acted on health information provided on TikTok by a health professional or general user.

Fact-Checking Information

Respondents were asked, on a scale ranging from 1=*not at all likely* to 7=*very likely*, to rate how likely they are to fact-check health information on TikTok from a health professional (mean 4.88, SD 1.80) and a general user (mean 5.37, SD 1.83).

Audience Characteristics

TikTok Intensity

Scores for TikTok intensity were created using an adapted form of the scale for Facebook intensity developed by Ellison et al [30]. The scale was created to measure how emotionally connected participants are to the social media platform as well as the extent to which the platform is part of their everyday lives. This TikTok-modified version of the scale (ranging from 1=*strongly disagree* to 7=*strongly agree*) asked respondents to rate their agreement with the following six items: (1) “TikTok is part of my everyday life,” (2) “I am proud to tell people I'm on TikTok,” (3) “TikTok has become part of my daily routine,” (4) “I feel out of touch when I haven't logged into TikTok for a while,” (5) “I feel I am part of the TikTok community,” and (6) “I would be sorry if TikTok shut down.” The scores of the 6 items were averaged to create TikTok intensity scores for each respondent who had reported ever having used TikTok (mean 4.91, SD 1.49; Cronbach $\alpha=0.90$).

Survey Questionnaire and Descriptive Statistics

The full survey questionnaire and descriptive statistics for each variable across the student and Qualtrics samples can be found in [Multimedia Appendix 1](#) and [Multimedia Appendix 2](#), respectively.

Data Analysis

Statistical analyses were performed using SPSS software (version 29.0; IBM Corp). Descriptive analyses were conducted to describe the respondents' frequency of TikTok use, frequency of intentional and unintentional exposure to health information on TikTok, reasons for health-related TikTok use, beliefs about encountering health misinformation, perceptions of misinformation on TikTok, frequency of performing verification behaviors on TikTok, frequency and preferences related to obtaining health information from health professionals and general users on TikTok, and frequency of acting on TikTok health information.

Bivariate correlational analyses were conducted to examine the relationships that credibility perceptions (the perceived credibility of TikTok health information overall, the perceived credibility of TikTok health information from health professionals, and the perceived credibility of TikTok health

information from general users), misinformation perceptions, verification behaviors, the likelihood of acting on TikTok health information, and the likelihood of fact-checking TikTok health information had with the respondents' age, highest level of education, and level of TikTok intensity. In addition, bivariate correlational analyses were conducted to examine the relationship between respondents' perceived credibility of TikTok health information overall and their likelihood of acting on the health information. For these correlational analyses, a correlation was considered weak if the correlation coefficient was between -0.4 and 0.4 . A correlation was considered moderate if the correlation coefficient was between -0.8 and -0.4 or between 0.4 and 0.8 . A correlation was considered strong if the correlation coefficient was between -1 and -0.8 or between 0.8 and 1 .

Finally, paired samples *t* tests (2-tailed) were conducted to observe the statistical differences between respondents' perceived susceptibility of health misinformation on TikTok for themselves versus for others, frequency of obtaining health information on TikTok from health professionals versus general users, preference for obtaining health information on TikTok

from health professionals versus general users, perceived credibility of TikTok health information from health professionals versus general users, likelihood of acting on TikTok health information from health professionals versus general users, and likelihood of fact-checking TikTok health information from a health professional versus a general user.

Results

Overview

A total of 1172 qualified responses were collected, with the average age of the sample being 22.82 (SD 3.15) years. A little more than half of the participants came from the Qualtrics panel (636/1172, 54.27%), and the rest were recruited through the universities (536/1172, 45.73%). The majority of the sample identified as White (910/1172, 77.65%), and most of the sample reported having used TikTok (1026/1172, 87.54%). Approximately half of the participants reported using TikTok more often than once a day (615/1172, 52.47%). Further demographic information is included in [Table 1](#), and the full list of demographic questions can be found in [Multimedia Appendix 1](#).

Table 1. Respondent demographics (N=1172).

Characteristics	Participants, n (%)
Race	
American Indian or Alaska Native	3 (0.26)
Asian	16 (1.37)
Black or African American	117 (9.98)
Native Hawaiian or Pacific Islander	57 (4.86)
White	910 (77.65)
Other	53 (4.52)
Prefer not to answer	16 (1.37)
Ethnicity	
Cuban	13 (1.11)
Mexican, Mexican American, or Chicana	109 (9.3)
Puerto Rican	11 (9.39)
Other Spanish, Hispanic, or Latina	98 (8.36)
Not Spanish, Hispanic, or Latina	919 (78.41)
Prefer not to answer	22 (1.88)
Education	
Less than high school	17 (1.45)
High school graduate or equivalent (eg, GED ^a)	207 (17.66)
Some college	509 (43.43)
2-year degree	90 (7.68)
4-year degree	224 (19.11)
Professional or master's degree	68 (5.8)
Doctorate	57 (4.86)
Self-identified gender	
Woman	1154 (98.46)
Transgender	4 (0.34)
Nonbinary	11 (0.94)
Gender fluid	2 (0.17)
Other	1 (0.09)
Frequency of TikTok use	
Never	146 (12.46)
Less than once a month	58 (4.95)
Once a month	50 (4.27)
Once a week	60 (5.12)
A few times a week	108 (9.22)
Once a day	135 (11.52)
More often than once a day	615 (52.47)

^aGED: General Educational Development test.

Using TikTok as a Health Information Source

Frequency of Use

Of the 1026 respondents who had used TikTok before, 672 (65.5%) reported that they had *intentionally* used TikTok to look for advice or information about their health or health care, while 948 (92.4%) reported that they had *unintentionally*

received health information or advice on TikTok. Of the 1026 respondents who had ever used TikTok, 582 (56.73%) reported having intentionally used TikTok to look for advice or information about their health or health care in the last 3 months. A breakdown of the frequency of intentional and unintentional exposure to health information on TikTok is provided in [Table 2](#).

Table 2. Frequency of using TikTok intentionally and unintentionally as a source of health information among respondents who had ever used TikTok (n=1026).

Questions and responses	Participants, n (%)
Frequency of intentional use of TikTok to obtain health information	
Hourly	29 (2.83)
Daily	108 (10.53)
Weekly	143 (13.94)
Monthly	168 (16.37)
Less often	190 (18.52)
Not at all	388 (37.82)
Frequency of unintentional exposure to health information on TikTok	
Hourly	38 (3.7)
Daily	232 (22.61)
Weekly	363 (35.38)
Monthly	172 (16.76)
Less often	143 (13.94)
Not at all	78 (7.6)

Reasons for Health-Related TikTok Use

Of the 10 reasons presented for health-related TikTok use, obtaining advice from others with the same disease or health condition (mean 5.29, SD 1.54), receiving social support from others (mean 5.29, SD 1.57), and gaining knowledge about a

disease they had been diagnosed with (mean 5.01, SD 1.59) were the most agreed upon reasons. The least agreed upon reason for health-related TikTok use was communicating with physicians (mean 4.06, SD 1.86). [Table 3](#) shows the mean and SD of the level of agreement for each of the 10 reasons.

Table 3. Reasons for health-related TikTok use among respondents who had ever used TikTok (n=1026).

I like to get health information from TikTok because...	Level of agreement ^a , mean (SD)
It can help me to maintain a healthy lifestyle.	4.92 (1.60)
It can help me determine whether I need to see a doctor.	4.81 (1.67)
It can provide me with more information after I've seen my doctor.	4.65 (1.72)
It can help me find different options for treatment or maintenance of my health condition(s).	4.79 (1.65)
I can gain knowledge about a disease I've been diagnosed with.	5.01 (1.59)
I can obtain advice from other patients with the same disease or health condition as me.	5.29 (1.54)
I can receive social support from others.	5.29 (1.57)
I can communicate with physicians.	4.06 (1.86)
I can interact in real time with TikTok users.	4.59 (1.81)
I can obtain immediate health information and make use of it.	4.53 (1.71)

^aRespondents indicated their level of agreement using a 7-point scale (ranging from 1=*strongly disagree* to 7=*strongly agree*).

Credibility, Misinformation, and Verification of Health Information on TikTok

Perceived Credibility of TikTok Health Information Overall

With a mean of 4.48 (SD 1.28) on a 7-point scale, the credibility perceptions were moderate. A positive correlation was found (959/1026, 93.47%; $r=0.30$; $P<.001$) between the perceived credibility of health information on TikTok overall and the respondents' age. Older participants tended to perceive the content as more credible. A positive correlation was also found (959/1026, 93.47%; $\rho=0.10$; $P=.001$) between the perceived credibility of health information on TikTok overall and the respondents' highest level of education, with respondents with greater education tending to rate the health information on TikTok as more credible. There was a positive correlation (959/1026, 93.47%; $r=0.32$; $P<.001$) between TikTok intensity and the perceived credibility of TikTok health information overall. Respondents with higher TikTok intensity scores tended to have greater perceived credibility of health information on TikTok overall.

Misinformation Perceptions

Approximately half (563/1026, 54.87%) of the respondents who had used TikTok indicated that they believe that they have personally encountered health misinformation on the platform at some point, and only 1.85% (19/1026) of the sample stated that they think that health misinformation is not prevalent at all on TikTok. Table 4 shows the levels of perceived prevalence,

severity, and susceptibility of health misinformation on TikTok reported by the respondents who had used TikTok.

There was a weak but significant negative correlation between age and the perceived seriousness of health misinformation on social media (1026/1172, 87.54%; $r=-0.07$; $P=.04$). There was not a significant relationship between age and perceived prevalence (1026/1172, 87.54%; $r=-0.06$; $P=.05$) or perceived susceptibility (1026/1172, 87.54%; $r=-0.02$; $P=.50$). There was not a significant relationship between the respondents' highest level of education and the perceived prevalence (1026/1172, 87.54%; $\rho=-0.008$; $P=.81$), the perceived seriousness (1026/1172, 87.54%; $\rho=0.01$; $P=.66$), or the perceived susceptibility of health misinformation on TikTok (1026/1172, 87.54%; $\rho=0.04$; $P=.17$).

There was a weak but significant positive relationship between respondents' TikTok intensity scores and perceived seriousness (1026/1172, 87.54%; $\rho=0.07$; $P=.02$) and perceived susceptibility (1026/1172, 87.54%; $\rho=0.07$; $P=.02$) such that respondents with high TikTok intensity scores tended to perceive greater seriousness and susceptibility of health misinformation on TikTok. There was not a significant relationship between TikTok intensity scores and the perceived prevalence of health misinformation on TikTok (1026/1172, 87.54%; $\rho=0.03$; $P=.29$).

The results also showed that respondents perceive other people (mean 5.26, SD 1.46) as more susceptible to health misinformation on TikTok than they personally are (mean 4.07, SD 1.75; $t_{1025}=21.16$; $P<.001$).

Table 4. Perceived prevalence, seriousness, and susceptibility of health misinformation on TikTok among respondents who had ever used TikTok (n=1026).

Questions	Level of agreement ^a , mean (SD)
How prevalent is health misinformation on TikTok?	5.14 (1.42)
How serious do you think the impact of health misinformation on TikTok is?	5.57 (1.46)
How susceptible are you to the influence of health misinformation on TikTok?	4.07 (1.75)
How susceptible are other people to the influence of health misinformation on TikTok?	5.26 (1.46)

^aRespondents indicated their level of agreement using a 7-point scale (ranging from 1=*not at all prevalent, serious, or susceptible* to 7=*very prevalent, serious, or susceptible*).

Verification of Health Information on TikTok

The most frequently used form of verification was considering whether the information represented was opinion or fact and the least frequently used form of verification was verifying the TikTok users' qualifications or credentials. Table 5 shows the respondents' frequency of each verification behavior.

There was a weak but significant positive correlation found between the respondents' age and their likelihood of verifying

TikTok health information (959/1026, 93.47%; $r=0.20$; $P<.001$) as well as between the respondents' highest level of education and their likelihood of verifying TikTok health information (959/1026, 93.47%; $\rho=0.09$; $P<.001$). A weak but significant positive relationship existed between respondents' TikTok intensity scores and their likelihood of verifying TikTok health information (959/1026, 93.47%; $r=0.21$; $P<.001$). Participants with high TikTok intensity scores were more likely to verify the health information.

Table 5. Verification of TikTok health information among participants who had seen health information on TikTok (n=959).

Verification behavior	Participant response ^a						
	Never, n (%)	Almost never, n (%)	Rarely, n (%)	About half of the time, n (%)	Most of the time, n (%)	Almost always, n (%)	Always, n (%)
Verify the TikTok users' qualifications or credentials	121 (12.6)	68 (7.1)	89 (9.3)	131 (13.7)	193 (20.1)	178 (18.6)	179 (18.7)
Consider the TikTok users' goals and objectives for posting information on the web	74 (7.7)	87 (9.1)	70 (7.3)	165 (17.2)	208 (21.7)	199 (20.8)	156 (16.3)
Check to see whether the information is current	72 (7.5)	61 (6.4)	66 (6.9)	143 (14.9)	223 (23.3)	198 (20.7)	196 (20.4)
Seek out other sources to validate the information	63 (6.6)	52 (5.4)	76 (7.9)	121 (12.6)	200 (20.9)	221 (23)	226 (23.6)
Consider whether the information represented is opinion or fact	53 (5.5)	50 (5.2)	49 (5.1)	125 (13)	199 (20.8)	251 (26.2)	232 (24.2)
Check to see that the information is complete and comprehensive	73 (7.6)	55 (5.7)	65 (6.8)	148 (15.4)	193 (20.1)	225 (23.5)	200 (20.9)

^aParticipants were asked to indicate how often they perform each of the 6 different behaviors when seeing health information on TikTok.

Health Professionals and General Users as Sources of Health Information on TikTok

Source Preferences

Of the respondents who had ever used TikTok, 93.08% (955/1026) indicated that they had obtained health information from a health professional on the platform, while 93.86% (963/1026) indicated that they had obtained health information from a general user on the platform. That said, respondents reported obtaining health information from health professionals on TikTok (mean 5.04, SD 1.83) significantly more often than they obtain health information from general users on TikTok (mean 4.55, SD 1.89; $t_{1025}=8.13$; $P<.001$). This was in line with their preferences for health information sources because the respondents' preference for obtaining health information from health professionals (mean 5.65, SD 1.75) was significantly greater than their preference for obtaining health information from general users (mean 4.08, SD 1.96; $t_{1025}=23.75$; $P<.001$).

Perceived Credibility of TikTok Health Information From Health Professionals and General Users

Respondents perceived health information from health professionals on TikTok (mean 5.16, SD 1.18) to be significantly more credible than health information provided by general users on TikTok (mean 3.95, SD 1.54; $t_{958}=26.737$; $P<.001$).

Acting on Health Information

Of the respondents who had received health information from a health professional on TikTok, 43.35% (414/955) reported that they had acted on health information they obtained from a health professional on TikTok. In comparison, 37.8% (364/963) of the respondents who had received health information from a general user on TikTok reported that they had acted on health information they obtained on TikTok from a general user. When asked about their likelihood of acting on health information on TikTok in the future, the respondents' likelihood of acting on health information from a health professional on TikTok (mean

4.50, SD 1.79) was significantly higher than their likelihood of acting on health information from a general user on TikTok (mean 3.96, SD 1.89; $t_{1025}=12.74$; $P<.001$).

Likewise, the respondents' perceived credibility of TikTok health information overall was positively correlated with their likelihood of acting on the health information. Respondents who perceived health information on TikTok overall as credible were more likely to act on health information they obtained from a health professional on TikTok (959/1026, 93.47%; $r=0.47$; $P<.001$) than from a general user on TikTok (959/1026, 93.47%; $r=0.53$; $P<.001$).

Age was weakly positively correlated with the respondents' likelihood of acting on health advice found on TikTok both from a health professional (1026/1172, 87.54%; $r=0.11$; $P<.001$) and from a general user (1026/1172, 87.54%; $r=0.23$; $P<.001$). Likewise, education was weakly positively correlated with their likelihood of acting on health information found on TikTok both from a health professional (1026/1172, 87.54%; $\rho=0.09$; $P=.003$) and from a general user (1026/1172, 87.54%; $\rho=0.08$; $P=.01$).

There was a significant positive relationship between TikTok intensity and the respondents' likelihood of acting on health information from health professionals on TikTok (1026/1172, 87.54%; $r=0.41$; $P<.001$). Likewise, there was a significant positive relationship between TikTok intensity and their likelihood of acting on health information from general users on TikTok (1026/1172, 87.54%; $r=0.39$; $P<.001$). Respondents with higher TikTok intensity scores tended to be more likely to act on health information from both health professionals and general users on TikTok.

Fact-Checking Information

There was a statistically significant difference in the respondents' likelihood of fact-checking health information on TikTok from a health professional versus a general user. The likelihood of fact-checking health information from a general user (mean 5.37, SD 1.83) was higher than the likelihood of

fact-checking health information from a health professional (mean 4.88, SD 1.80; $t_{1025}=9.71$; $P<.001$).

Age was weakly positively correlated with the respondents' likelihood of fact-checking TikTok health information both from a health professional (1026/1172, 87.54%; $r=0.18$; $P<.001$) and from a general user (1026/1172, 87.54%; $r=0.07$; $P=.02$). Education was also weakly positively correlated with their likelihood of fact-checking TikTok health information from a health professional (1026/1172, 87.54%; $\rho=0.11$; $P<.001$), but education was not correlated with their likelihood of fact-checking TikTok health information for a general user (1026/1172, 87.54%; $\rho=0.04$; $P=.26$).

There was a significant but weak positive relationship between TikTok intensity and the respondents' likelihood of fact-checking TikTok health information from both health professionals (1026/1172, 87.54%; $r=0.07$; $P=.03$) and general users on TikTok (1026/1172, 87.54%; $r=0.10$; $P=.002$). Respondents with higher TikTok intensity scores tended to be more likely to fact-check health information from both health professionals and general users on TikTok.

Discussion

Principal Findings

TikTok has generated substantial attention due to recent reports suggesting its emergence as a significant source of information for many Americans [7]. For some users, TikTok has replaced traditional news networks as well as widely used search engines such as Google [31]. Given this emergence of TikTok as an information source and the presence of health information available on the platform [5,6], we surveyed 1172 women aged 18 to 29 years to understand their use of TikTok as a source of health information. Of the 1172 respondents, 1026 (87.54%) had used TikTok in some capacity.

The findings provide evidence that TikTok has become a source of health information for young women in the United States. More than half of the respondents who had ever used TikTok (672/1026, 65.5%) reported that they had intentionally used TikTok to look for advice or information about their health or health care, and nearly the entire sample of TikTok users (948/1026, 92.4%) reported having unintentionally obtained health information on TikTok. The popularity of health-related hashtags on TikTok (eg, as of November 2023, #medicaltiktok and #healthtok had 7.6 billion and 2.4 billion views, respectively) has illuminated some of TikTok's popularity as a commonly searched platform for information related to health, but the findings of this study provide a greater empirical understanding of the extent to which young women actually obtain health information from the platform.

TikTok's popularity as a source of health information may, in part, be the result of how technology has influenced human beings' desire for immediate information. Rather than having to wait for a physician's appointment to ask about one's symptoms or health condition, one can take to the internet (eg, TikTok) and find related information in a matter of minutes [12]. This phenomenon of individuals seeking immediate information has important implications for health professionals.

By knowing that individuals turn to platforms such as TikTok to find health information, health professionals can proactively create content so that credible health information is available when users go to find it. Social media platforms were heavily relied upon for health information during the COVID-19 pandemic [16], and since then, social media, and TikTok specifically, have been recommended as a tool for health promotion [5,32,33]. Given that TikTok is easily accessible and allows anyone to consume information without judgment, it may especially be helpful for populations with barriers to care and for communicating about taboo or stigmatized topics that users may be less comfortable asking about in a traditional setting [6,34].

To better understand why young women are using TikTok as a source of health information, we asked our respondents about their agreement with various reasons for health-related TikTok use. Our findings showed that the most agreed upon reasons were obtaining advice from others with the same disease or health condition, receiving social support from others, and gaining knowledge about a disease they had been diagnosed with. In an examination of how the current digital landscape has affected Americans' consumer behavior, CharityRx found "relatability to a shared personal experience" to be a top reason why people go to health influencers for information [12]. In addition to TikTok having the ability to provide immediate information related to users' health inquiries, it has the capacity to help users locate other people who are similar to them. This may be especially relevant for women. As women have experienced gender bias and poorer treatment in health care settings [35,36], it is possible that they may be especially motivated to seek social support and health information from others like them. Prior research has also indicated that social support is especially beneficial for women [37]. When individuals perceive similarity to a source of information, this can cause the message recipient to feel a stronger sense of connection with the message, which can have important implications in terms of the effects of a message [38]. Perceiving similarity to a source of information can also enhance a user's perceptions of the message and overall acceptance [39]. In terms of obtaining a further understanding of young women's motivations for using TikTok as a source of health information, future research could examine what topics (eg, health conditions) young women are most interested in and likely to search for.

Given the potential for misinformation to rapidly pervade the social media landscape, it has been recommended that experts in medical science, public health, and social sciences collaborate to better understand health misinformation on social media, including its reach and influence [40]. Our findings help show the degree to which young women on TikTok perceive an issue of misinformation and to what extent they try to verify or fact-check the information they consume. In first examining the perceived credibility of TikTok health information overall, we found that credibility perceptions were moderate among the respondents who had reported ever having obtained health-related advice or information (intentionally or unintentionally) from TikTok. Notably, health information on TikTok was perceived to be more credible by participants who were older, more educated, or had higher TikTok intensity scores

(ie, were more emotionally connected to TikTok and had greater integration of TikTok into their daily lives). While our findings are able to show these positive correlations, we do not know whether the information they are seeing on TikTok (and thus reflecting on when indicating their credibility perceptions) truly is credible. It may be the case that older age, higher levels of education, and more experience with TikTok lead to following more credible users and being delivered more credible content via the TikTok algorithm. In this case, the content may truly be more credible for these users (leading to their greater perceptions of credibility). Future research could explore whether this is the case.

Nearly all participants (1007/1026, 98.15%) indicated that they believe that misinformation is prevalent on TikTok to at least some extent. This may be the result of mainstream news commonly communicating that misinformation is a problem on social media platforms, including TikTok [15]. A great deal of health misinformation reached social media users during the COVID-19 pandemic [8]. It is possible that respondents in this study were among these users or that they heard about this problematic phenomenon. However, despite nearly all participants (1007/1026, 98.15%) stating that they think that misinformation is prevalent on TikTok to at least some extent, only approximately half of the participants (563/1026, 54.87%) indicated that they believe that they have personally encountered health misinformation on TikTok at some point. This discrepancy could stem from a few factors. It could be that the users know that there is a misinformation epidemic but have not been exposed to misinformation because of their commitment to only following credible users (thus leading the TikTok algorithm to feed them more credible content). However, given the large amounts of misinformation that have been identified on the platform [15,41,42] and Americans' inability to identify most forms of misinformation [43-45], it is more likely that this discrepancy is the result of some respondents not having recognized that they have been exposed to health misinformation. This possibility is further supported by our results that showed that respondents perceive other people as more susceptible to health misinformation on TikTok than they personally are. This finding demonstrates what seems to be a third-person effect, in which the young women perceive that media messages have a greater influence on others than on themselves [23]. As evidence of a third-person effect was provided by the results (with respondents perceiving that other people are more susceptible to health misinformation on TikTok than they personally are), it might be that some of the respondents are naive about their susceptibility to misinformation.

The most common form of verification (to verify the accuracy of the health information found on TikTok) was considering whether the information presented was opinion or fact. The least frequently reported form of verification was verifying the TikTok users' qualifications or credentials. Age, education, and TikTok intensity were each found to have a weak positive correlation with the likelihood of verifying TikTok health information. Participants who were older, more educated, or had higher TikTok intensity scores were more likely to verify health information on TikTok. As discussed in the Results

section, these demographics (age, education, and TikTok intensity) were also found to each be positively correlated with the perceived credibility of TikTok health information overall. Thus, as both credibility perceptions and the verification of the information increase with age, education, and TikTok intensity, perhaps these users truly are seeing more credible content and appropriately perceiving it to be credible. Future research should further explore these relationships because this is beyond the scope of this study. In prior research, fact-checking has typically been found to not be very common on social media. In a survey by Neely et al [16] of Americans' reliance on social media during the COVID-19 pandemic, it was found that three-fourths of those surveyed had relied on social media to some extent to stay informed about the pandemic, but the majority of them were unlikely to fact-check the information they found.

Finally, this study investigated young women's perceptions and behaviors related to the top 2 types of sources that share health information on TikTok: health professionals and general users. Prior research has identified that health professionals and general users are 2 of the most prevalent sources on TikTok communicating health information [10,11], and our findings indicate that young women are obtaining health information on TikTok from each of these 2 source types. Of the respondents who had used TikTok, a majority indicated that they had obtained health information both from a health professional on the platform (955/1026, 93.08%) and from a general user on the platform (963/1026, 93.86%). Given that the respondents' top reasons for health-related TikTok use were obtaining advice from others with the same disease or health condition, receiving social support from others, and gaining knowledge about a disease they had been diagnosed with, it is reasonable that the respondents would perceive both health professionals and general users as valuable sources of information. Prior research exploring the effects of communication sources on social media [9] has explained that both expert-type and peer-type sources provide value. While health professionals have formal training and credentialed experience, general users (eg, peers) can have a form of "experiential credibility" from their own personal experiences (such as that of living with a particular health condition) [9,27]. While most of the respondents had obtained health information from each of the 2 source types, their preference for obtaining health information from a health professional was significantly greater than their preference for obtaining health information from general users ($t_{1025}=23.75$; $P<.001$), and, in line with their preferences, the young women reported obtaining health information from health professionals on TikTok significantly more often than they reported obtaining health information from general users on TikTok ($t_{1025}=8.13$; $P<.001$). As medical professionals and health institutions delivered COVID-19-related information during the pandemic [5], it was found that Americans who used social media as a source of COVID-19-related information expanded their social media networks to include credible sources (eg, medical institutions and scientific sources) [16]. In addition, CharityRx's survey found "medical accreditation and certification" to be the top reported reason why people go to influencers for health information [12]. Together, the prior and current findings seem to indicate a preference for obtaining health information from

health professionals on platforms such as TikTok. However, it is important to recognize that both of the common source types—health professionals and general users—are providing health information to these users.

Our findings showed that the perceived credibility of TikTok health information from health professionals was significantly higher than the perceived credibility of TikTok health information from general users ($t_{958}=26.737$; $P<.001$). This is promising for health professionals who choose to invest in creating a TikTok presence because studies have shown that web-based information is more likely to be attended to when it is perceived as credible [46]. There was also a significant difference in the young women's likelihood of acting on health information from a health professional versus a general user on TikTok, with their likelihood of acting on the information being greater when the information was from a health professional ($t_{1025}=12.74$; $P<.001$). As the respondents had greater perceived credibility of TikTok health information from health professionals (vs general users), it is logical that they would also be more likely to act on health information from health professionals, given that credibility perceptions enhance the likelihood of persuasion [47,48], including in the context of social media messaging [49]. We found that, of the respondents who had received health information from a health professional on TikTok, 43.35% (414/955) reported that they had acted on health information they obtained from a health professional on TikTok. In comparison, 37.8% (364/963) of the respondents who had received health information from a general user on TikTok reported that they had acted on health information they obtained on TikTok from a general user. The respondents' perceived credibility of health information on TikTok overall was found to be positively correlated with their likelihood of acting on health information from both of the source types, which is in alignment with the relationship between credibility perceptions and persuasive effects [47,48]. On the one hand, these findings are promising in the sense that young women perceive health professionals on TikTok to be more credible and are more influenced by them, further suggesting that it is worthwhile for health professionals to use TikTok as a strategic communication tool. On the other hand, this could mean that young women are more susceptible to being influenced by individuals who give the impression of being qualified health professionals. Medical professionals sometimes provide information that is outside of their scope of expertise [50], and uncredentialed users are often confused for credentialed health professionals [51]. Furthermore, it is important to note that medical credentials and titles vary. A search of the hashtag #womenshealth on TikTok results in videos from a number of different types of health professionals, including nurses, nurse practitioners, obstetrician-gynecologists, medical doctors (with MD or DO credentials), and midwives, and because prior research has shown that many individuals do not understand medical roles and titles or how to differentiate between them [52], this could have profound implications. It is also important to note that anyone on TikTok can present themselves as though they have the necessary credentials for the information they are sharing (eg, adding credentials to their username, presenting themselves with a formal title, wearing a laboratory coat or

surgical scrubs, and communicating information in a persuasive manner). An authoritative title, on its own, can be enough to capture an individual's attention and generate respect [53]. Therefore, with anyone being able to add credentials to their TikTok username, this could be problematic, especially given that credibility is hard to distinguish on social media. Users are more likely to rely on heuristic cues (such as the titles included in a username) to determine a user's credibility [9,54]. The aforementioned findings also showed that the least frequently reported form of information verification was verifying the TikTok users' qualifications or credentials, further illustrating that this could be vastly problematic.

The study's findings showed that the young women in our study are more likely to fact-check information from a general user than fact-check information from a health professional. Again, this is promising in terms of the fact that content from general users may be more likely to include misinformation, but this could be problematic if the credibility cues of a health professional lead users to automatically assume that they can believe and trust any of the information. It would be worthwhile for future research to investigate this further, uncovering whether users trust misinformation from health professionals on TikTok simply because of the creators' credentials.

We also found that both age and education are positively correlated with young women's likelihood of acting on health information found on TikTok—both from a health professional and from a general user. However, these correlations were quite weak. A stronger positive correlation was found between the users' emotional connection to TikTok (ie, TikTok intensity) and their likelihood of acting on health information from both health professionals and general users. It may be logical to assume that users who heavily engage more with TikTok will have a greater propensity to act on the information they receive. Social influence theory suggests that individuals are influenced by those around them [55]. This may extend to the web-based environment, such that as TikTok becomes more integrated into one's life, it is more likely to affect one's behaviors.

Limitations

The findings of this survey research should be interpreted in light of some limitations. First, the sample of survey respondents was recruited through convenience sampling methods. While the sample is only a segment of the total population, we tried to ensure that we had a large sample that was representative of the population of interest (women aged 18-29 years throughout the United States) by using both Qualtrics and 2 large public universities to recruit individuals who fit the sample parameters. This study does not discuss differences between individuals recruited from the universities and those recruited from the Qualtrics sample, but we have provided descriptive statistics for each variable across each sample as a means of allowing for some comparison between the 2 samples (Multimedia Appendix 2). In addition, we made sure to include the respondents' highest level of education as a variable in this study to see how education is associated with the variables of interest in this study, which can help to provide some understanding of how one's educational experience might be related to one's use and perceptions of TikTok health information. Second, because this

study relies on self-reporting from the survey respondents, there is a chance that the results do not truly capture the real behaviors of the respondents. As we asked questions about fact-checking web-based information (behaviors that individuals likely know they should engage in), the respondents may have answered some questions in a more socially desirable or acceptable manner (ie, what they expect would be an “appropriate” response) rather than being truthful in their responses. Hopefully, though, because participants knew that their responses would be anonymous, this helped to lessen social desirability bias. Finally, this study focused on young women (assigned female sex at birth) as the population of interest for this study. This narrow focus allows us to better understand the implications of TikTok use for this demographic, but it is also important to explore other populations’ behaviors and perceptions.

Conclusions and Practical Implications

This study provides a greater understanding of the extent to which TikTok is serving as a source of health information for young women in the United States. With nearly all young women in this study (who had used TikTok) having been exposed to health information on TikTok (948/1026, 92.4%), and more than half of them (582/1026, 56.73%) having actively sought health-related information on the platform in the last 3 months, it is imperative for health professionals and health communication scholars to prioritize the consideration of TikTok as a platform that is influencing health information acquisition and dissemination in the United States. While the popularity and accessibility of TikTok may change, short-form video social media sites are likely to remain a common form of communication [56].

The findings of this study illustrate the potential value that TikTok can have for disseminating health information to an audience of young women. As the respondents of this survey reported a preference for information from health professionals and were more likely to act on the information from these

sources, it is worthwhile for health professionals to use TikTok to disseminate health information to this audience, especially given the large number of women on the platform and prior research illustrating that social media significantly influence women’s health-related behaviors and perceptions [17,19]. In doing so, health professionals may want to consider how they can align their content with young women’s most common motivations for using TikTok as a source of health information, which we found to be obtaining advice from others with the same disease or health condition, receiving social support from others, and gaining knowledge about a disease they had been diagnosed with. Given that young women want advice and support from others who are experiencing similar health conditions, it may be useful for health professionals to consider working with individuals who are willing to share their personal experience with a health condition. Incorporating the stories of patients and other experienced individuals who have similar characteristics to those searching for information on TikTok could be especially influential for increasing attention to, and engagement with, health information on TikTok.

Furthermore, given that our findings indicate that young women have a preference for obtaining health information on TikTok from health professionals and that they are less likely to fact-check information from these sources, it is imperative that future initiatives address the proliferation of individuals sharing information beyond their scope of expertise and the problem of social media users confusing uncredentialed users as credentialed health professionals [50,51]. Future researchers and practitioners should also work on media literacy and education initiatives, given the third-person effect found in this research. It seems that young women know that misinformation is an issue on TikTok, but it seems that they may not be recognizing that they have been exposed to misinformation and that they perceive themselves as less susceptible. It may be beneficial for future interventions to address this perception and help young women to have better recognition of when they are being exposed to health misinformation.

Acknowledgments

This study was supported by an Emerging Scholars grant from the Association for Education in Journalism and Mass Communication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Full survey questionnaire.

[DOCX File, 31 KB - [infodemiology_v4i1e54663_app1.docx](#)]

Multimedia Appendix 2

Descriptive statistics for study measures across 2 samples (student sample and Qualtrics sample).

[DOCX File, 41 KB - [infodemiology_v4i1e54663_app2.docx](#)]

References

1. Dellatto M. TikTok hits 1 billion monthly active users. Forbes. URL: <https://www.forbes.com/sites/marisadellatto/2021/09/27/tiktok-hits-1-billion-monthly-active-users/> [accessed 2023-11-02]

2. Willingham A, Asmelash L, Andrew S. The biggest ways TikTok has changed American culture. CNN. URL: <https://www.cnn.com/2023/04/02/us/tiktok-american-culture-effects-cec/index.html> [accessed 2023-11-02]
3. Shepardson D. TikTok hits 150 mln U.S. monthly users, up from 100 million in 2020. Reuters. URL: <https://www.reuters.com/technology/tiktok-tell-congress-it-has-150-million-monthly-active-us-users-2023-03-20/> [accessed 2023-11-02]
4. Yuen M. TikTok users worldwide (2020-2025). Insider Intelligence. URL: <https://www.emarketer.com/chart/c/T11778/tiktok-users-worldwide-2020-2025-millions-change-of-internet-users-1> [accessed 2023-05-10]
5. Comp G, Dyer S, Gottlieb M. Is TikTok the next social media frontier for medicine? AEM Educ Train 2021 Jul 21;5(3):1-4 [FREE Full text] [doi: [10.1002/aet2.10532](https://doi.org/10.1002/aet2.10532)] [Medline: [34095694](https://pubmed.ncbi.nlm.nih.gov/34095694/)]
6. Kindelan K. Women turn to TikTok for health information and OBGYNs are there to meet them. Good Morning America. URL: <https://www.goodmorningamerica.com/wellness/story/women-turn-tiktok-health-information-obgyns-meet-68824442> [accessed 2023-11-02]
7. Matsa KE. More Americans are getting news on TikTok, bucking the trend on other social media sites. Pew Research Center. URL: <https://tinyurl.com/34avwsnk> [accessed 2023-10-30]
8. Basch CH, Hillyer GC, Jaime C. COVID-19 on TikTok: harnessing an emerging social media platform to convey important public health messages. Int J Adolesc Med Health 2022 Oct 01;34(5):367-369 [FREE Full text] [doi: [10.1515/ijamh-2020-0111](https://doi.org/10.1515/ijamh-2020-0111)] [Medline: [32776899](https://pubmed.ncbi.nlm.nih.gov/32776899/)]
9. Kirkpatrick CE, Lawrie LL. Can videos on TikTok improve pap smear attitudes and intentions? Effects of source and autonomy support in short-form health videos. Health Commun 2023 Sep 10:1-13. [doi: [10.1080/10410236.2023.2254962](https://doi.org/10.1080/10410236.2023.2254962)] [Medline: [37691164](https://pubmed.ncbi.nlm.nih.gov/37691164/)]
10. Kong W, Song S, Zhao YC, Zhu Q, Sha L. TikTok as a health information source: assessment of the quality of information in diabetes-related videos. J Med Internet Res 2021 Sep 01;23(9):e30409 [FREE Full text] [doi: [10.2196/30409](https://doi.org/10.2196/30409)] [Medline: [34468327](https://pubmed.ncbi.nlm.nih.gov/34468327/)]
11. Song S, Xue X, Zhao YC, Li J, Zhu Q, Zhao M. Short-video apps as a health information source for chronic obstructive pulmonary disease: information quality assessment of TikTok videos. J Med Internet Res 2021 Dec 20;23(12):e28318 [FREE Full text] [doi: [10.2196/28318](https://doi.org/10.2196/28318)] [Medline: [34931996](https://pubmed.ncbi.nlm.nih.gov/34931996/)]
12. CharityRx Blog. URL: <https://tinyurl.com/yck583pj> [accessed 2023-11-02]
13. Montenegro L. The rise of short-form video: TikTok is changing the game. Forbes. URL: <https://www.forbes.com/sites/forbesagencycouncil/2021/08/27/the-rise-of-short-form-video-tiktok-is-changing-the-game/?sh=78ccf7c15083> [accessed 2023-11-02]
14. Paivio A. Imagery and Verbal Processes. New York, NY: Psychology Press; 2013.
15. Brewster J, Arvanitis L, Pavidonis V, Wang M. Misinformation monitor: September 2022. NewsGuard. URL: <https://www.newsguardtech.com/misinformation-monitor/september-2022> [accessed 2023-11-02]
16. Neely S, Eldredge C, Sanders R. Health information seeking behaviors on social media during the COVID-19 pandemic among American social networking site users: survey study. J Med Internet Res 2021 Jun 11;23(6):e29802 [FREE Full text] [doi: [10.2196/29802](https://doi.org/10.2196/29802)] [Medline: [34043526](https://pubmed.ncbi.nlm.nih.gov/34043526/)]
17. Ceci L. U.S. TikTok audience by gender 2023. Statista. URL: <https://www.statista.com/statistics/1398182/tiktok-us-audience-by-gender/> [accessed 2023-11-02]
18. Ceci L. U.S. TikTok users by age 2022. Statista. URL: <https://www.statista.com/statistics/1095186/tiktok-us-users-age/> [accessed 2023-11-02]
19. Davies C, Mann A. Factors influencing women to accept diet and exercise messages on social media during COVID-19 lockdowns: a qualitative application of the health belief model. Health Mark Q 2023;40(4):415-433. [doi: [10.1080/07359683.2023.2193076](https://doi.org/10.1080/07359683.2023.2193076)] [Medline: [37021625](https://pubmed.ncbi.nlm.nih.gov/37021625/)]
20. Dluhos-Sebesto C, Jethwa TE, Bertasi TG, Bertasi RA, Nishi LY, Pantin SA, et al. Women's health information survey: common health concerns and trusted sources of health information among different populations of female patients. Womens Health Rep (New Rochelle) 2021;2(1):173-181 [FREE Full text] [doi: [10.1089/whr.2020.0118](https://doi.org/10.1089/whr.2020.0118)] [Medline: [34235504](https://pubmed.ncbi.nlm.nih.gov/34235504/)]
21. Manierre MJ. Gaps in knowledge: tracking and explaining gender differences in health information seeking. Soc Sci Med 2015 Mar;128:151-158. [doi: [10.1016/j.socscimed.2015.01.028](https://doi.org/10.1016/j.socscimed.2015.01.028)] [Medline: [25618604](https://pubmed.ncbi.nlm.nih.gov/25618604/)]
22. General facts on women and job based health. U.S. Department of Labor. 2018. URL: <https://www.dol.gov/sites/dolgov/files/ebsa/about-ebsa/our-activities/resource-center/fact-sheets/women-and-job-based-health.pdf> [accessed 2024-04-29]
23. van der Meer TG, Brosius A, Hameleers M. The role of media use and misinformation perceptions in optimistic bias and third-person perceptions in times of high media dependency: evidence from four countries in the first stage of the COVID-19 pandemic. Mass Commun Soc 2022 Feb 25;26(3):438-462. [doi: [10.1080/15205436.2022.2039202](https://doi.org/10.1080/15205436.2022.2039202)]
24. Caiata-Zufferey M, Abraham A, Sommerhalder K, Schulz PJ. Online health information seeking in the context of the medical consultation in Switzerland. Qual Health Res 2010 Aug;20(8):1050-1061. [doi: [10.1177/1049732310368404](https://doi.org/10.1177/1049732310368404)] [Medline: [20442347](https://pubmed.ncbi.nlm.nih.gov/20442347/)]
25. Huo J, Desai R, Hong YR, Turner K, Mainous AG, Bian J. Use of social media in health communication: findings from the health information national trends survey 2013, 2014, and 2017. Cancer Control 2019 Apr 18;26(1):1073274819841442 [FREE Full text] [doi: [10.1177/1073274819841442](https://doi.org/10.1177/1073274819841442)] [Medline: [30995864](https://pubmed.ncbi.nlm.nih.gov/30995864/)]

26. Jia X, Pang Y, Liu LS. Online health information seeking behavior: a systematic review. *Healthcare (Basel)* 2021 Dec 16;9(12):1740 [FREE Full text] [doi: [10.3390/healthcare9121740](https://doi.org/10.3390/healthcare9121740)] [Medline: [34946466](https://pubmed.ncbi.nlm.nih.gov/34946466/)]
27. Flanagin AJ, Metzger MJ. Perceptions of internet information credibility. *Mass Commun Q Journal* 2016 Jun 25;77(3):515-540. [doi: [10.1177/107769900007700304](https://doi.org/10.1177/107769900007700304)]
28. Chang C. Fake news: audience perceptions and concerted coping strategies. *Digit Journal* 2021 May 25;9(5):636-659. [doi: [10.1080/21670811.2021.1923403](https://doi.org/10.1080/21670811.2021.1923403)]
29. Hu Y, Shyam Sundar S. Effects of online health sources on credibility and behavioral intentions. *Commun Res* 2009 Nov 25;37(1):105-132. [doi: [10.1177/0093650209351512](https://doi.org/10.1177/0093650209351512)]
30. Ellison N, Steinfield C, Lampe C. The benefits of Facebook "friends": social capital and college students' use of online social network sites. *J Comput Mediat Commun* 2007;12(4):1143-1168. [doi: [10.1111/j.1083-6101.2007.00367.x](https://doi.org/10.1111/j.1083-6101.2007.00367.x)]
31. Perez S. Google exec suggests Instagram and TikTok are eating into Google's core products, Search and Maps. *TechCrunch*. URL: <https://tinyurl.com/mp8zhsy9> [accessed 2023-11-03]
32. de Vere Hunt I, Linos E. Social media for public health: framework for social media-based public health campaigns. *J Med Internet Res* 2022 Dec 14;24(12):e42179 [FREE Full text] [doi: [10.2196/42179](https://doi.org/10.2196/42179)] [Medline: [36515995](https://pubmed.ncbi.nlm.nih.gov/36515995/)]
33. MacKinnon KR, Kia H, Lacombe-Duncan A. Examining TikTok's potential for community-engaged digital knowledge mobilization with equity-seeking groups. *J Med Internet Res* 2021 Dec 09;23(12):e30315 [FREE Full text] [doi: [10.2196/30315](https://doi.org/10.2196/30315)] [Medline: [34889739](https://pubmed.ncbi.nlm.nih.gov/34889739/)]
34. Gordon D. 33% Of Gen Zers trust TikTok more than doctors, new survey shows. *Forbes*. URL: <https://www.forbes.com/sites/debgordon/2022/12/20/33-of-gen-zers-trust-tiktok-more-than-doctors-new-survey-shows/?sh=2f5770e6c7b2> [accessed 2023-04-29]
35. Samulowitz A, Gremyr I, Eriksson E, Hensing G. "Brave men" and "emotional women": a theory-guided literature review on gender bias in health care and gendered norms towards patients with chronic pain. *Pain Res Manag* 2018;2018:6358624-6358614 [FREE Full text] [doi: [10.1155/2018/6358624](https://doi.org/10.1155/2018/6358624)] [Medline: [29682130](https://pubmed.ncbi.nlm.nih.gov/29682130/)]
36. Govender V, Penn-Kekana L. Gender biases and discrimination: a review of health care interpersonal interactions. *Glob Public Health* 2008;3 Suppl 1:90-103. [doi: [10.1080/17441690801892208](https://doi.org/10.1080/17441690801892208)] [Medline: [19288345](https://pubmed.ncbi.nlm.nih.gov/19288345/)]
37. Milner A, Krnjacki L, LaMontagne AD. Age and gender differences in the influence of social support on mental health: a longitudinal fixed-effects analysis using 13 annual waves of the HILDA cohort. *Public Health* 2016 Nov;140:172-178. [doi: [10.1016/j.puhe.2016.06.029](https://doi.org/10.1016/j.puhe.2016.06.029)] [Medline: [27527844](https://pubmed.ncbi.nlm.nih.gov/27527844/)]
38. Kelman HC. Processes of opinion change. *Public Opin Q* 1961;25(1):78 [FREE Full text] [doi: [10.1086/266996](https://doi.org/10.1086/266996)]
39. Andsager JL, Bemker V, Choi HL, Torwel V. Perceived similarity of exemplar traits and behavior: effects on message evaluation. *Commun Res* 2016 Jun 30;33(1):3-18. [doi: [10.1177/0093650205283099](https://doi.org/10.1177/0093650205283099)]
40. Chou WY, Oh A, Klein WM. Addressing health-related misinformation on social media. *JAMA* 2018 Dec 18;320(23):2417-2418. [doi: [10.1001/jama.2018.16865](https://doi.org/10.1001/jama.2018.16865)] [Medline: [30428002](https://pubmed.ncbi.nlm.nih.gov/30428002/)]
41. O'Sullivan NJ, Nason G, Manecksha RP, O'Kelly F. The unintentional spread of misinformation on 'TikTok': a paediatric urological perspective. *J Pediatr Urol* 2022 Jun;18(3):371-375. [doi: [10.1016/j.jpuro.2022.03.001](https://doi.org/10.1016/j.jpuro.2022.03.001)] [Medline: [35331640](https://pubmed.ncbi.nlm.nih.gov/35331640/)]
42. Basch CH, Meleo-Erwin Z, Fera J, Jaime C, Basch CE. A global pandemic in the time of viral memes: COVID-19 vaccine misinformation and disinformation on TikTok. *Hum Vaccin Immunother* 2021 Aug 03;17(8):2373-2377 [FREE Full text] [doi: [10.1080/21645515.2021.1894896](https://doi.org/10.1080/21645515.2021.1894896)] [Medline: [33764283](https://pubmed.ncbi.nlm.nih.gov/33764283/)]
43. Poll: most Americans encounter health misinformation, and most aren't sure whether it's true or false. *KFF*. URL: <https://tinyurl.com/yxc6ms58> [accessed 2023-11-03]
44. Lyons BA, Montgomery JM, Guess AM, Nyhan B, Reifler J. Overconfidence in news judgments is associated with false news susceptibility. *Proc Natl Acad Sci U S A* 2021 Jun 08;118(23):e2019527118 [FREE Full text] [doi: [10.1073/pnas.2019527118](https://doi.org/10.1073/pnas.2019527118)] [Medline: [34074757](https://pubmed.ncbi.nlm.nih.gov/34074757/)]
45. Scheufele DA, Krause NM. Science audiences, misinformation, and fake news. *Proc Natl Acad Sci U S A* 2019 Apr 16;116(16):7662-7669 [FREE Full text] [doi: [10.1073/pnas.1805871115](https://doi.org/10.1073/pnas.1805871115)] [Medline: [30642953](https://pubmed.ncbi.nlm.nih.gov/30642953/)]
46. Metzger MJ, Flanagin AJ. Using Web 2.0 technologies to enhance evidence-based medical information. *J Health Commun* 2011;16 Suppl 1:45-58. [doi: [10.1080/10810730.2011.589881](https://doi.org/10.1080/10810730.2011.589881)] [Medline: [21843095](https://pubmed.ncbi.nlm.nih.gov/21843095/)]
47. Dong Z. How to persuade adolescents to use nutrition labels: effects of health consciousness, argument quality, and source credibility. *Asian J Commun* 2015 Feb 19;25(1):84-101. [doi: [10.1080/01292986.2014.989241](https://doi.org/10.1080/01292986.2014.989241)]
48. Eagly AH, Wood W, Chaiken S. Causal inferences about communicators and their effect on opinion change. *J Pers Soc Psychol* 1978;36(4):424-435. [doi: [10.1037//0022-3514.36.4.424](https://doi.org/10.1037//0022-3514.36.4.424)]
49. Dockter CE, Lee S, Boman CD, Hinnant A, Cameron GT. The impact of retransmission and modality on communicating health research findings via social media. *Health Commun* 2021 Sep;36(10):1231-1241 [FREE Full text] [doi: [10.1080/10410236.2020.1749354](https://doi.org/10.1080/10410236.2020.1749354)] [Medline: [32268798](https://pubmed.ncbi.nlm.nih.gov/32268798/)]
50. Vorias T. Beware of TikTok's medical "experts". *The Michigan Daily*. 2021. URL: <http://www.michigandaily.com/opinion/columns/beware-of-tiktoks-medical-experts/> [accessed 2023-11-12]
51. Sung M. On TikTok, mental health creators are confused for therapists. *Mashable*. URL: <https://mashable.com/article/tiktok-mental-health-therapist-psychology> [accessed 2023-11-12]

52. Hause E, Praska C, Pitt MB, Hendrickson MA, Charpentier V, Allen KA, et al. What's in a name? Laypeople's understanding of medical roles and titles. *J Hosp Med* 2022 Dec 29;17(12):956-960 [FREE Full text] [doi: [10.1002/jhm.12971](https://doi.org/10.1002/jhm.12971)] [Medline: [36173137](https://pubmed.ncbi.nlm.nih.gov/36173137/)]
53. Rhoads KV, Cialdini RB. The business of influence: Principles that lead to success in commercial settings. In: Dillard JP, Pfau M, editors. *The Persuasion Handbook: Developments in Theory and Practice* Sage. Thousand Oaks, CA: SAGE Publications; 2002:513-542.
54. Petty RE, Cacioppo JT. The elaboration likelihood model of persuasion. In: Berkowitz L, editor. *Advances in Experimental Social Psychology*. Washington, DC: Academic Press; 1986:123-205.
55. Kelman HC. Compliance, identification, and internalization three processes of attitude change. *J Conflict Resolut* 2016 Jul 01;2(1):51-60 [FREE Full text] [doi: [10.1177/002200275800200106](https://doi.org/10.1177/002200275800200106)]
56. Singh S. Why short-form video is here to stay (even if TikTok isn't). LinkedIn. 2023. URL: <https://www.linkedin.com/pulse/why-short-form-video-here-stay-even-tiktok-isnt-sunjay-singh/> [accessed 2023-11-02]

Edited by T Mackey; submitted 17.11.23; peer-reviewed by S Chinn, H Song, M Loades; comments to author 10.01.24; revised version received 05.03.24; accepted 11.04.24; published 21.05.24.

Please cite as:

Kirkpatrick CE, Lawrie LL

TikTok as a Source of Health Information and Misinformation for Young Women in the United States: Survey Study

JMIR Infodemiology 2024;4:e54663

URL: <https://infodemiology.jmir.org/2024/1/e54663>

doi: [10.2196/54663](https://doi.org/10.2196/54663)

PMID: [38772020](https://pubmed.ncbi.nlm.nih.gov/38772020/)

©Ciera E Kirkpatrick, LaRissa L Lawrie. Originally published in *JMIR Infodemiology* (<https://infodemiology.jmir.org>), 21.05.2024. 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.

Original Paper

The Use of Social Media to Express and Manage Medical Uncertainty in Dyskeratosis Congenita: Content Analysis

Emily Pearce¹, MPH, PhD; Hannah Raj², BSc; Ngozika Emezienna¹; Melissa B Gilkey³, PhD; Allison J Lazard⁴, PhD; Kurt M Ribisl³, PhD; Sharon A Savage¹, MD, PhD; Paul KJ Han⁵, MD, PhD

¹Division of Cancer Epidemiology and Genetics, Clinical Genetics Branch, National Cancer Institute, National Institutes of Health, Rockville, MD, United States

²Team Telomere, Coeur d'Alene, ID, United States

³Department of Health Behavior, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

⁴Hussman School of Journalism and Media, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

⁵Division of Cancer Control and Population Sciences, Behavioral Research Program, National Cancer Institute, National Institutes of Health, Rockville, MD, United States

Corresponding Author:

Emily Pearce, MPH, PhD

Division of Cancer Epidemiology and Genetics, Clinical Genetics Branch

National Cancer Institute

National Institutes of Health

609 Nelson St

Rockville, MD, 20850

United States

Phone: 1 9196992547

Email: emily.pearce@nih.gov

Abstract

Background: Social media has the potential to provide social support for rare disease communities; however, little is known about the use of social media for the expression of medical uncertainty, a common feature of rare diseases.

Objective: This study aims to evaluate the expression of medical uncertainty on social media in the context of dyskeratosis congenita, a rare cancer-prone inherited bone marrow failure and telomere biology disorder (TBD).

Methods: We performed a content analysis of uncertainty-related posts on Facebook and Twitter managed by Team Telomere, a patient advocacy group for this rare disease. We assessed the frequency of uncertainty-related posts, uncertainty sources, issues, and management and associations between uncertainty and social support.

Results: Across all TBD social media platforms, 45.98% (1269/2760) of posts were uncertainty related. Uncertainty-related posts authored by Team Telomere on Twitter focused on scientific (306/434, 70.5%) or personal (230/434, 53%) issues and reflected uncertainty arising from probability, ambiguity, or complexity. Uncertainty-related posts in conversations among patients and caregivers in the Facebook community group focused on scientific (429/511, 84%), personal (157/511, 30.7%), and practical (114/511, 22.3%) issues, many of which were related to prognostic unknowns. Both platforms suggested uncertainty management strategies that focused on information sharing and community building. Posts reflecting response-focused uncertainty management strategies (eg, emotional regulation) were more frequent on Twitter compared with the Facebook community group ($\chi^2_1=3.9$; $P=.05$), whereas posts reflecting uncertainty-focused management strategies (eg, ordering information) were more frequent in the Facebook community group compared with Twitter ($\chi^2_1=55.1$; $P<.001$). In the Facebook community group, only 36% (184/511) of members created posts during the study period, and those who created posts did so with a low frequency (median 3, IQR 1-7 posts). Analysis of post creator characteristics suggested that most users of TBD social media are White, female, and parents of patients with dyskeratosis congenita.

Conclusions: Although uncertainty is a pervasive and multifactorial issue in TBDs, our findings suggest that the discussion of medical uncertainty on TBD social media is largely limited to brief exchanges about scientific, personal, or practical issues rather than ongoing supportive conversation. The nature of uncertainty-related conversations also varied by user group: patients and caregivers used social media primarily to discuss scientific uncertainties (eg, regarding prognosis), form social connections, or exchange advice on accessing and organizing medical care, whereas Team Telomere used social media to express scientific and

personal issues of uncertainty and to address the emotional impact of uncertainty. The higher involvement of female parents on TBD social media suggests a potentially greater burden of uncertainty management among mothers compared with other groups. Further research is needed to understand the dynamics of social media engagement to manage medical uncertainty in the TBD community.

(*JMIR Infodemiology* 2024;4:e46693) doi:[10.2196/46693](https://doi.org/10.2196/46693)

KEYWORDS

social media; medical uncertainty; telomere biology disorder; dyskeratosis congenita; social support

Introduction

Background

Medical uncertainty is a common experience in rare diseases and may combine with limited scientific knowledge and access to peer groups to impede a patient's ability to seek and adhere to medical treatments [1] and intensify health-related anxiety, decreasing quality of life for patients and their caregivers [2,3]. Dyskeratosis congenita (DC) is a rare telomere biology disorder (TBD) associated with very high risks of bone marrow failure, pulmonary and liver disease, cancer, and other medical conditions. Diagnosis is challenging because of its wide phenotypic spectrum, including the classic DC triad (nail dysplasia, abnormal skin pigmentation, and oral leukoplakia) with pediatric bone marrow failure, middle-age presentation with pulmonary failure or aplastic anemia, abnormally short telomere length, or detection of pathogenic germline variants in >18 different genes [4]. Although age of onset is variable, DC often presents in childhood and adolescence, with most patients experiencing their first symptoms before the age of 20 years [5]. Diagnosis frequently results in a lifetime commitment to screening to detect progressive clinical manifestations of DC, including cancers across multiple organ systems [5]. Owing to the complexity and rarity of DC and related TBDs, patients and their families often have long diagnostic journeys, face complicated health decision-making, and frequently do not have access to medical professionals and supportive peers who are familiar with their condition. This situation likely creates a substantial burden of medical uncertainty for patients with TBDs and their families. Although medical uncertainty has been associated with increased anxiety and difficulty with decision-making in rare diseases and cancer occurrence and recurrence [6-11], to date, no research has addressed the experience or management of medical uncertainty in the TBD context.

As outlined in a previously published taxonomy developed by Han [12], uncertainty in medicine arises from multiple sources (eg, probability, ambiguity, and complexity) and focuses on scientific, personal, and practical issues. These situations activate a variety of management strategies to address uncertainty, which are primarily cognitive, emotional, and relational in nature. Uncertainty management strategies may target ≥ 1 sources or issues of uncertainty and are defined as belonging to ≥ 1 of the following approaches: seeking information to fill knowledge gaps ("ignorance-focused"), reducing or increasing attention to unknowns ("uncertainty-focused"), ameliorating adverse psychological effects of uncertainty ("response-focused"), and fostering

interpersonal relationships to engage with uncertainty as a shared experience ("person-focused"). In situations where uncertainty cannot be reduced, these strategies may mitigate its negative mental health impact and help individuals achieve an adaptive, optimal balance of responses to uncertainty (uncertainty tolerance).

The rarity of TBDs suggests a potential role for internet-based platforms to deliver social support by bridging geographic, knowledge, and community network limitations. Social support, a complex concept encompassing a variety of helping social interactions [13], includes four main types: (1) expression of empathy and care (emotional), (2) provision of tangible assistance (instrumental), (3) provision of knowledge or facts (informational), and (4) evaluative feedback about task performance or personal qualities (appraisal) [14]. Research suggests that social support decreases the experience of stress, anxiety, and depression and improves the overall quality of life in populations experiencing medical uncertainty [8,10,15-17]. The benefit of social support has been demonstrated in patients with Li-Fraumeni syndrome, a rare genetic cancer predisposition, where informational, tangible, spiritual, and emotional support from in-person sources enhanced positive coping capacities [18]. Social media platforms such as Facebook and Twitter have been identified as important resources for social support in rare disease contexts [19-24], and disease-specific social media support has been recommended in oncology [25], rare genetic disease [26-28], and other stigmatized or rare diseases [29-31]. In addition to increasing access to information and social networks, continued participation in socially supportive internet-based communities may also build capacities for uncertainty tolerance [10,17,32-38]. Although social media has the potential to bridge geographic or social boundaries, its use is often concentrated in select populations, limiting its reach and potentially inhibiting its use by some groups [39,40]. In addition, dynamics observed on social media posts may not reflect real-life experiences and are limited in depth and detail, increasing the potential for misinterpretation [39]. Social media can also spread misinformation with damaging consequences, especially in high-uncertainty health contexts [41-43].

Objectives

Although extensive research has investigated the psychosocial benefits of internet-based health forums for patients and their caregivers [23,28,29,44-51], there is still a need to evaluate the use of social media to express or manage medical uncertainty in rare diseases. Specifically, we need to examine social media use for expressing and managing medical uncertainty in TBDs to understand the experience of medical uncertainty in this

context and to build evidence to improve health communication and uncertainty management interventions [52]. This exploratory study aims to review social media posts created by and targeted at patients with TBDs and their caregivers to (1) measure the frequency of uncertainty-related posts; (2) catalog the issues, sources, and types of uncertainty and uncertainty management strategies; (3) measure user engagement with different post types; and (4) explore the relationship between uncertainty and social support. To achieve these aims, we reviewed all publicly available social media sites owned and maintained by Team Telomere (previously DC Outreach, Inc), the oldest and largest patient advocacy organization for individuals, caregivers, and families affected by TBDs worldwide [53]. The social media of Team Telomere constitutes the most expansive and accessible body of internet-based TBD-related content, inclusive of a variety of user perspectives. The variety of posts by users with diverse connections to TBDs (eg, medical providers, patients, caregivers, and health advocacy nonprofits) makes Team Telomere's social media an ideal data source for understanding the range and dynamics of medical uncertainty communication and social support exchange in the TBD context.

Methods

Ethical Considerations

Data collection was undertaken in partnership with Team Telomere following best practices guidelines for social media

Table 1. Data source characteristics at the time of the study.

	Facebook community group	Facebook main page	Twitter
Creation date (y)	2020	2010	2010
Followers, n	187	1637	1933
Posts ^a , n	511	1815	434

^aRepresents posts captured during the study period (June 2019 to December 2021).

Inclusion

All posts made on Team Telomere's social media (Facebook main page: n=1818, Facebook community group: n=518, and Twitter: n=441) between June 2019 and December 2021 were eligible for inclusion. This time frame encompasses the period starting 1 year before the Facebook community group. This group was created in June 2020 as a platform for social connection during the COVID-19 pandemic. Posts were excluded from the analysis if they were (1) removed by the user or Team Telomere (n=5), (2) duplicate posts with identical content from the same day (n=2), or (3) posts without image or text content (n=7). This resulted in a total of 2760 posts, with both primary posts and comments considered unique. The post was used as the unit of analysis and included all content visible to a passive social media user. Additional post content that required clicking links to external sites or embedded audiovisual materials was not included in this study.

Data Extraction and Quality Control

We met with Team Telomere's leadership (eg, executive director and board) before conducting the study and cocreated a community-based research contract outlining parameters.

research [54] and was approved by the National Institutes of Health Institutional Review Board (IRB 000722).

Data Source

The source of data for this study was all publicly available social media owned and maintained by Team Telomere. These sites included the Team Telomere Twitter page [55], the Facebook main page [56], and a public Facebook community group [57] (Table 1). All the sites were open to the public and had no eligibility requirements for membership. Content across all platforms was monitored by Team Telomere to ensure appropriate adherence to community guidelines, and Team Telomere's staff removed posts with offensive or scientifically inaccurate content. The Facebook main page and Twitter accounts were created to promote the work of Team Telomere "supporting families worldwide affected by Dyskeratosis Congenita and Telomere Biology Disorders" [56]. The Facebook community group was created in response to social isolation following the COVID-19 pandemic as "a place to share our everyday lives in the spirit of promoting and maintaining connections among our Team Telomere/Dyskeratosis Congenita/Telomere Biology disorder community" [57].

Although all data were publicly available and Facebook data were manually extracted by the authors, Team Telomere facilitated data extraction from Twitter by sharing downloaded images and text files made available to them as account owners. We used the post (original or responses), rather than post creator, as the unit of measurement and did not collect identifying information of the social media users or interact directly with users.

Data were extracted directly from each social media site manually through (1) screenshots saved as deidentified image files and (2) cut-and-paste of post text into an Excel (Microsoft Corporation) spreadsheet. For the Facebook community group, we assigned unique ID numbers to post creators using public data (usernames) to calculate how many unique users engaged in conversation threads, and we viewed the publicly available profile images to assess observed sex and race. Posts were assigned a unique ID number within Excel, and additional data were manually extracted for each post to capture the post popularity (number of likes, shares, and comments), post type (primary post or comment), and types of emojis present. Demographics of post creators (observed gender and race) were assessed through an independent review of profile images and

profile names by 3 coders (EP, HR, and NE). Quality control for data extraction was performed on a subset of the data (n=100 posts) by NE, and intercoder reliability was assessed during the multiple-reviewer coding process.

Coding and Analysis

We used a combined content analysis mixed methods approach to analyze the social media data [58]. This involved qualitative analysis (coding by multiple independent reviewers) and quantitative analysis (frequency and chi-square testing). Constructs were defined through codebook development using deductive (theory driven) approaches, whereas qualitative themes were identified through inductive (data driven) discussion, as described in greater detail in the *Methods* section. The analysis was performed separately for each social media source, 2 Facebook pages (the Team Telomere main page and a separate community group page established in 2020) and the Team Telomere Twitter feed, resulting in the creation of 3 separate data sets (Facebook main page: n=1815, Facebook community group: n=511, and Twitter: n=434). A subset of Facebook community group posts (n=77; 12 primary posts and 65 comments) was reviewed by 3 coders and used to inform uncertainty inclusion criteria (Multimedia Appendix 1) and the codebook (Multimedia Appendix 2) developed to deductively identify the presence or absence of uncertainty and social support constructs defined in the Han Taxonomy of Medical Uncertainty [12] and the Social Support Framework [14]. Then, all posts were coded for uncertainty and social support by 3 independent coders (EP, HR, and PKJH), with all disagreements in coding resolved through discussion and consensus. Posts identified as uncertainty related in the Facebook community group (n=156) and Twitter (n=210) were then independently subcoded (EP, HR, and PKJH) for uncertainty issues, sources, and management strategies according to the codebook definitions detailed in the *Measures* section. Data were then arranged by subcode and reviewed qualitatively to detect themes that emerged from the data and were refined through discussion between coders.

Measures

Intercoder Reliability

Intercoder reliability among the 3 coders was measured across all social media types for the initial coding of dichotomous social support and uncertainty variables using Cohen κ . The analysis found acceptable reliability of independent coders in assessing the presence or absence of any social support (κ value range across all platforms, $\kappa=0.79-0.95$) and uncertainty (κ value range across all platforms, $\kappa=0.58-0.93$) across all social media platforms. Regardless, all discrepancies were mutually resolved through coder consensus.

Post Creator Characteristics

Post creator characteristics were visible from profile images and usernames that appeared alongside each post. Posts from Team Telomere's organizational account were created by staff members, often identified in the post context (eg, executive director, communications director, or board member). We did not scrutinize user profiles to detect the activity of nonhuman bots; however, in the context of the small population with this

rare disease, most users could be positively identified as human beings from the context of their posts and history of participation in organizational events. Post creator characteristics, including observed gender and race, were assessed by 3 independent coders' perceptions of publicly available usernames and profile images. Disagreements between coders resulted in the characteristic being coded as "unknown."

Uncertainty Issues, Sources, and Management Strategies

Posts were coded as uncertainty related according to 1 of the following types: (1) uncertainty-related primary posts, (2) uncertainty-related comments, and (3) non-uncertainty-related posts captured within a thread where 1 or more other post was uncertainty related. For the Facebook community group and Twitter, posts identified as uncertainty-related primary posts or comments were further analyzed to determine the presence or absence of sources (ambiguity, complexity, and probability), issues (scientific, personal, and practical), and attributes of uncertainty management strategies (ignorance focused, uncertainty focused, response focused, and person focused). We defined sources of uncertainty as insufficient, unreliable, or contradictory information (ambiguity); information features, such as multiple or interacting causes and effects that make a phenomenon difficult to understand (complexity); and fundamental randomness or indeterminacy of a phenomenon that makes outcomes unpredictable (probability). We defined issues of uncertainty as pertaining to the causes, diagnosis, prognosis, or management of disease (scientific); the impact of disease on aspects of personal life (personal); and logistical issues related to health care or disease management (practical). Although the data did not allow assessment of intent to manage uncertainty, we searched posts to identify evidence of management strategies with ≥ 1 of the following attributes: (1) providing or seeking information to fill knowledge gaps (ignorance focused), (2) reducing or increasing attention to unknowns to gain or relinquish a sense of control (uncertainty focused), (3) ameliorating the adverse psychological effects of uncertainty (response focused), and (4) fostering interpersonal relationships to engage with uncertainty as a shared experience (person focused).

Social Support

Posts were categorized as containing social support through qualitative coding by 3 independent reviewers (EP, HR, and PKJH) following definitions developed over decades of research in social support theory [14,59,60]. Dichotomous variables were assigned to indicate the presence or absence of social support and the presence or absence of specific types of support within 4 domains (appraisal, emotional, informational, and instrumental). These domains were defined as (1) giving or receiving evaluative feedback (appraisal); (2) giving or receiving indicators of care, love, appreciation, empathy, or sympathy (emotional); (3) giving or receiving knowledge (informational); and (4) giving or receiving tangible support (instrumental), as recently formulated by Holt-Lunstad and Uchino [14]. Assignment to social support domains was not mutually exclusive.

Relationship Between Social Support and Uncertainty

We examined the relationship between social support and uncertainty by comparing frequencies and chi-square tests. Posts were coded as dichotomous variables for uncertainty (uncertainty related, non-uncertainty related), uncertainty subtypes (presence or absence), and social support subtypes (presence or absence). We examined the frequencies of social support subtypes in uncertainty-related posts overall, by social media platform (Facebook community group and Twitter) and by post type (primary post or comment). We performed chi-square tests to determine the strength of the relationship between uncertainty-related posts and social support across platforms and for uncertainty-related posts by post type (primary post, comment, thread) and issue subtype (scientific, personal, practical).

Popularity and Engagement

Popularity on the Facebook community group, Facebook main page, and Twitter was defined as the sum of comments, likes, and shares. Engagement was defined separately for social media types (Facebook community group and Facebook main page vs Twitter) owing to differences in user tracking approaches between Facebook and Twitter platforms. Facebook engagement was defined as the sum of conversations (number of responses generated by a post or comment), voices (number of unique users responding to a post or comment), and depth (number of back-and-forth responses). Engagement on Twitter was defined as the sum of detail expands (clicks to view more of the post), profile visits, link clicks, and video views. Engagement was also measured for the Facebook community group by examining the proportion of users who contributed posts and post frequencies by author.

Sentiment

Sentiment analysis was performed through manual annotation by 2 independent coders, with differences resolved through consensus. Posts were assigned categorical sentiment variables according to the (1) frequency and (2) presence or absence of keywords and emojis. Unambiguous emotion words (eg, “happy” and “sad”) were chosen as keywords to indicate emotional valence, as described in other studies [61,62]. The emotional valence of emojis was assigned based on the emoji definition in internet-based emoji dictionaries and validated by a coder review of the emoji within the post context ([Multimedia Appendix 3](#)).

Results

Post Characteristics

A total of 2760 posts created on all platforms between June 2019 and December 2021 were included in this study. Across all platforms, most posts were created either by the executive director of Team Telomere or by individual users who were primarily identified as White, female, and parents of children affected by TBDs. Post characteristics differed by platform: on Twitter, most posts (368/434, 84.8%) were primary posts, most of which (384/434, 88.5%) were generated by the executive director of Team Telomere; Facebook main page posts were either primary posts (800/1815, 44.08%) or first-level comments (1014/1815, 55.87%) created by Team Telomere (860/1815, 47.38%) or individual users (955/1815, 52.62%); and on the Facebook community group, most posts (403/511, 78.9%) were comments to primary posts, in sometimes lengthy (up to 8 level) conversation threads created by 67 individual users (502/511, 98.2%). Posts across all platforms were written almost exclusively in English ([Table 2](#)).

Table 2. Characteristics of posts on Team Telomere's social media from June 2019 to December 2021 (N=2760).

	Facebook community group (n=511), n (%)	Facebook main page (n=1815), n (%)	Twitter (n=434), n (%)
Post type			
Primary post	108 (21.1)	800 (44.1)	368 (84.8)
Comment	403 (78.9)	1015 (55.9)	66 (15.2)
Language			
English	487 (95.3)	1807 (99.6)	434 (100)
Other ^a	4 (0.8)	8 (0.4)	0 (0)
Image only	17 (3.3)	0 (0)	0 (0)
Creator type			
Team telomere	8 (1.6)	861 (47.4)	385 (88.7)
Individual	503 (98.4)	954 (52.6)	49 (11)
Observed creator sex^b			
Male	25 (5)	69 (7.2)	5 (10)
Female	478 (95)	885 (92.8)	41 (83.7)
Unknown	0 (0)	1 (0.1)	3 (6.1)
Observed creator race^b			
White	443 (88.1)	766 (80.3)	40 (81.6)
Other ^c	46 (9.1)	30 (3.1)	6 (12.2)
Unknown	14 (2.8)	158 (16.6)	3 (6.1)
Observed creator telomere biology disorder relationship^{b,d}			
Patient	65 (12.9)	42 (4.4)	1 (2)
Parent	428 (85.1)	384 (40.3)	14 (28.6)
Medical provider	3 (0.6)	31 (3.2)	10 (20.4)
Other ^e	5 (1)	59 (6.2)	22 (44.9)
Unknown	40 (8)	495 (51.9)	2 (4.1)
Multiple	129 (25.6)	126 (13.2)	0 (0)

^aRespectively by platform (Facebook community group, Facebook main page, and Twitter), "other" language included Spanish (0.2%, 0.2%, and 0%), French (0.4%, 0.1%, and 0%). In the Facebook community group the following languages also appeared: Hebrew (0.1%), Italian (0.1%), Swedish (0.1%), and Māori (0.2%).

^bIncludes frequencies for individual creator types only; does not include Team Telomere organization (Facebook community group: n=503, Facebook main page: n=954, and Twitter: n=49).

^cRespectively by platform (Facebook community group, Facebook main page, and Twitter), "other" identified creator race and ethnicity included Latinx (7.7%, 1.5%, and 1.4%) and Arab or Middle Eastern (1.4%, 11%, and 0%).

^dFrequency does not total to 100% because of some individuals occupying multiple categories.

^eRespectively by platform (Facebook community group, Facebook main page, and Twitter), "other" creator telomere biology disorder relationship included grandparent (0%, 0.2%, and 0%), sibling (0.4%, 0.9%, and 0%), spouse (0%, 0.2%, and 0%), other advocacy organization representative (not Team Telomere; 0%, 0%, and 40.8%), and clinical or pharmaceutical industry representative (0%, 0.1%, and 4.1%).

Qualitative Findings

Qualitative analysis of posts revealed multiple uncertainty issues, sources, and management indicators. Issues included diagnostic, prognostic, therapeutic, and causal uncertainties (scientific); assembly of medical care teams, geographic or financial constraints, and limitations to research funding and dissemination (practical); and building "rare" identity, communicating complex health information to children, and reframing educational or developmental goals (personal).

Sources of uncertainty included confusing symptoms and lack of clarity in medical advice (ambiguity); the TBD impact of TBD on multiple organ systems, managing medications or screening regimens, emotional confusion, and achieving scientific literacy across different medical specialties (complexity); and prognostic outcomes, behavioral health risks, or genetic inheritance (probability). Attributes of uncertainty management strategies included (1) information seeking, participation in research, and connection to trusted information sources and care providers (ignorance focused); (2) ordering

multiple uncertainties through categorization, prioritization, and sequential narratives, including counting of survival days since transplant (uncertainty focused); (3) sharing positive emotions, portraying TBD experience as a source of strength, and encouraging relaxation (response focused); and (4) promoting a TBD community identity by creating a community

mascot (a unicorn named “Tillymere”), recognizing community-specific celebrations (TBD month and transplant anniversaries), providing TBD-pride identifiers (T-shirts and swag), and making reference to Team Telomere as a “family” (person focused; [Table 3](#)).

Table 3. Uncertainty in telomere biology disorder (TBD) social media.

	Post text
Sources of uncertainty	
Ambiguity	<ul style="list-style-type: none"> “This is a tough one! One of those maybe/maybe not symptoms...I often ask myself the same questions about my daughter’s more obscure symptoms.” [FBCG218304.21.07.30] “Pre-lung # transplantation patients with # pulmonary # fibrosis who have short # telomeres may need different # clinical care...” [TWT180100.19.06.11]
Complexity	<ul style="list-style-type: none"> “[Name] is having kidney, heart, and lung problems. Oh, and who can forget the liver? This week has been too long at the hospital!” [FBCG2110000.21.11.23] “# DYK Those with # telomere biology disorders may be especially vulnerable to the effects of taking multiple medicines at the same time and may respond to medications differently.” [TWT186700.19.11.14]
Probability	<ul style="list-style-type: none"> “80% of patients diagnosed with dyskeratosis congenita will experience bone marrow failure.” [TWT185500.19.11.04] “5 out of 6 of the cell lines tested were less than 1%. And when that’s the case, patients have a 10-20% chance of getting cancer...” [FBCG203500.20.09.08] “A recent publication advises against an elective eye surgery in patients with DC due to higher long-term risks caused by delayed healing...” [TWT182100.19.08.25]
Issues of uncertainty	
Scientific	<ul style="list-style-type: none"> “Has anybody experienced hearing loss with connection to short telomere length?” [FBCG218300.21.07.30] “Has anyone had kidney problems outside of BMT? Are there any articles anyone has seen on kidneys and short telomeres?” [FBCG2110000.21.11.23]
Practical	<ul style="list-style-type: none"> “At the moment [Name] has 1-2 appointments each week. Add to that emails to/from paediatrician, calls from hospital to change/confirm appointments...It’s overwhelming some weeks. And I’m usually doing all this from work. We are also applying for different supports...so lots of forms, phone calls and emails!” [FBCG204305-8.20.10.13]
Personal	<ul style="list-style-type: none"> “It’s # PFMonth, and we want you to know you have a team surrounding you...” [TWT1816000.20.09.04] “TBDs are not just a pediatric disease. Affected adults with a # rare disease, you are NOT ALONE!” [TWT183100.19.09.21] “Another milestone reached. This time five years ago as we celebrated [Name]’s 5th birthday we were also getting ready to go to transplant two weeks later. Yesterday we celebrated the big 10...” [FBCG201300.20.06.27]
Focus of uncertainty management	
Ignorance	<ul style="list-style-type: none"> “Wondering if anyone with DC had a dental implant post-transplant...? And did your medical team have any concerns or recommendations?” [FBCG215500.21.01.05] “Hello—any contraindications to getting COVID 19 vaccine if you have DC?” [FBCG217100.21.04.04] “Do you have a copy of the clinical guidelines?” [FBCG203509.20.09.08] “Take time to learn more about #Telomere Biology Disorders through our informational video!” [TWT1822100.21.11.04]
Uncertainty	<ul style="list-style-type: none"> “Each Family Story is set up so you can find a connection via gene or experience.” [FBCG204400.20.10.29] “My daughter has yearly bone marrow biopsies, lung and liver screenings. ENT and skin checks for cancer.” [FBCG203513.20.09.08] “I’ve been preparing something for the new school trying to give them what her medical challenges are.” [FBCG219900.21.11.16]
Response	<ul style="list-style-type: none"> “Our family is celebrating today! [Name]’s Happy 8th bone marrow transplant anniversary!” [FBCG203300.20.08.24] “Fitting for us all: it wasn’t the trauma that made you strong, kinder, and more compassionate. It’s how you handled it. That credit is yours.” [FBCG216200.21.02.28] “Join@sixnwstevies as she teaches yoga for research...” [TWT1822600.21.03.16]
Person	<ul style="list-style-type: none"> “Thank goodness for social media otherwise it would be very isolating.” [FBCG203821.20.09.25] “Don’t forget to register for our Young Adult Meetup...” [TWT1814300.20.06.23] “[Name] it’s never ending, I hope you find a way to take care of you” [FBCG204307.20.10.13] “You are in great hands but always happy to connect with [Provider Name]” [FBCG203504.20.09.08] “Check out # tillymere! All # sparkly and ready for # TBDmonth!” [TWT185400.19.11.04] “We have all known the long loneliness and we have learned that the only solution is love and that love comes with community. – Dorothy Day” [TWT1816300.20.09.12]

Uncertainty Issues, Sources, and Management Strategies

Content analysis revealed that 45.98% (1269/2760) of posts overall were uncertainty related, although the frequency differed by platform (Facebook main page: 691/1715, 40.29%; Facebook community group: 155/511, 30.3%; and Twitter: 210/434, 48.4%). Most uncertainty-related posts on Facebook community group and Twitter were generated by Team Telomere's organizational profile (332/511, 65% and 353/434, 81.3%, respectively) and were often similar in topic, wording, and image content. In the Facebook community group, all uncertainty-related posts were generated by individual users, including a portion (119/511, 23.3%) posted by Team Telomere-affiliated volunteer group moderators.

Owing to low frequency of community-generated uncertainty content on the Facebook community group and Twitter, compared with the Facebook community group, we decided to code uncertainty subtypes only within the Facebook community group and Twitter to compare how medical uncertainty was expressed on social media by 2 contrasting content creator

groups (community members vs advocacy organization). Scientific uncertainty was the most common issue on both platforms (305/434, 70.3% to 429/511, 84%). On Twitter, personal uncertainty was more frequently discussed, whereas in the Facebook community group, practical uncertainty was more frequent. Across platforms, most posts (1713/2760, 62.07%) had multiple sources of uncertainty, and a substantial number of posts (1126/2760, 40.8%) were coded as emerging from the combined information features of probability, complexity, and ambiguity. The most common attributes of uncertainty management styles detected on both platforms were requests or offers of information to fill knowledge gaps (ignorance focused) and offers of emotional support or community building (person focused). Response-focused management style attributes (eg, yoga and meditation classes) were marginally more frequent on Twitter compared with the Facebook community group ($\chi^2_1=3.9$; $P=.05$), but on the Facebook community group, indicators of uncertainty-focused management (eg, strategies for organization of care logistics) were more frequent compared with Twitter ($\chi^2_1=55.1$; $P<.001$; Table 4).

Table 4. Characteristics and frequency of uncertainty-related posts on Team Telomere's Facebook community group and Twitter (N=2760).

	Facebook community group (n=156), n (%)	Twitter (n=210), n (%)	Chi-square (df) ^a	P value
Issue				
Personal	48 (30.8)	111 (52.9)	16.6 (1)	<.001
Practical	35 (22.4)	23 (11)	9.2 (1)	.002
Scientific	131 (84)	148 (70.5)	11.4 (1)	.007
Multiple	53 (34)	59 (28.1)	— ^a	—
Source				
Ambiguity	81 (51.9)	80 (38.1)	17.6 (1)	<.001
Complexity	81 (51.9)	75 (35.7)	20.8 (1)	<.001
Probability	112 (71.8)	81 (38.6)	71.3 (1)	<.001
Multiple	88 (56.4)	77 (36.7)	—	—
Management attributes				
Ignorance focused	124 (79.5)	156 (74.3)	1.9 (1)	.16
Person focused	106 (67.9)	125 (59.5)	3.6 (1)	.06
Response focused	57 (36.5)	100 (47.6)	3.9 (1)	.05
Uncertainty focused ^b	53 (34)	10 (4.8)	55.1 (1)	<.001
Multiple	106 (67.9)	131 (62.4)	—	—

^aChi-square tests were not performed for issues, sources, or management attributes assigned to multiple categories.

^bUncertainty thread includes non-uncertainty-related posts captured in a thread where ≥ 1 other posts were uncertainty related.

Facebook Social Support and Uncertainty

Frequent overlap of social support and uncertainty was found across all platforms, with uncertainty-related posts being more likely to contain social support compared with non-uncertainty-related posts ($\chi^2_1=70.7$; $P<.001$). However, within social support subtypes, only informational support remained significantly more frequent within uncertainty-related

posts ($\chi^2_1=486.0$; $P<.001$), whereas emotional support was significantly less frequent in uncertainty-related posts ($\chi^2_1=66.5$; $P<.001$) compared with non-uncertainty-related posts. The relationship between informational support and uncertainty remained significant for all social media types, but the relationship between emotional support and uncertainty differed by platform (Multimedia Appendix 4). Emotional support was significantly more frequent in uncertainty-related posts for the

Facebook community group ($\chi^2_1=7.8$; $P=.005$), was significantly less frequent in uncertainty-related posts on the Facebook main page ($\chi^2_1=79.5$; $P<.001$), and had no relationship with uncertainty-related posts on Twitter ($\chi^2_1=0.5$; $P=.47$).

On all platforms, uncertainty-related posts were more frequently offers of support than requests. When requests occurred, they were more likely to appear on the Facebook community group compared with Twitter ($\chi^2_1=12.7$; $P<.001$). Posts that were not uncertainty related but appeared in an uncertainty-related thread frequently contained offers of emotional support.

Given the greater variation in types and direction (offer vs request) of social support in the Facebook community group, we decided to focus on subsequent analyses of the relationship between social support and uncertainty subtypes on this platform. Analysis of social support in the Facebook community group posts by uncertainty issue found that informational support was offered more frequently in response to scientific and practical uncertainty posts compared with personal uncertainty posts. Informational support was also the most frequent type of support requested and offered across uncertainty source types in the Facebook community group; however, uncertainty posts emerging from probability concerns had similar frequencies of emotional and informational support (320/511, 62.6% and 511/836, 61.1%, respectively). This was particularly true in the case where a post had multiple uncertainty sources, which were more likely to be coded as informational support offers or requests compared with posts with only a single uncertainty source ($\chi^2_1=90.4$; $P\leq.001$).

Popularity and Engagement

Popularity and engagement were positively skewed toward lower values across all social media types. Popularity was highest for posts on Twitter (Facebook community group: median 1, range 0-55, mean 4, SD 7.5; Facebook main page: range 0-151, median 1, mean 5.9, SD 13.3; and Twitter: range 0-1147, median 13, mean 28.8, SD 76.6). However, engagement was higher in the Facebook community group than on the Facebook main page or Twitter (Facebook community group: range 0-29.6, median 0.54, mean 2.15, SD 4.0; Facebook main page: median 0.0006, range 0-0.09, mean 0.004, SD 0.008; and Twitter: median 0.007, range 0-0.56, mean 0.02, SD 0.04). Most uncertainty-related posts were categorized as having below-median popularity and engagement. The uncertainty-related post with the highest engagement was a question about kidney issues and telomere length posted on Facebook community group by a parent of a child with TBDs, which generated 12 comments from 6 unique users, including a self-identified medical expert. The nonnormal distribution combined with low (<20) frequency in cross-tabulation groups made it ineffective to analyze the relationships between the presence of social support and popularity or engagement (Multimedia Appendix 5).

In the Facebook community group, posts were created by 67 unique individuals, representing 35.8% (183/511) of all group members. Frequency per user was positively skewed toward lower numbers (range 1-94 posts and median 3 posts), and the

majority of post creators (343/511, 67.1%) generated ≤ 5 posts. Although Team Telomere rarely posted directly on the Facebook community group (8/511, 1.6% posts), the top 2 post creators (156/511, 30.5% posts) were identified as White, female, parents of children affected by DC who were also group moderators for Team Telomere. After removing the moderators, the remaining median post frequency was 3 posts per user, with 22.3% (114/511) of the users creating only a single post.

Sentiment

The majority of posts (2208/2760, 80%) on all social media types were categorized as positive sentiment. Negative sentiment was rarely expressed and was more likely to be expressed on Facebook compared with Twitter ($\chi^2_1=45.4$; $P<.001$). Uncertainty-related posts demonstrated a similarly high frequency of positive sentiment across all social media types (Facebook community group: 433/511, 84.7%; Facebook main page: 1495/1815, 82.37%; and Twitter: 328/434, 75.6%; Multimedia Appendix 6).

Discussion

Principal Findings

In this study, we explored the use of TBD social media to express health-related uncertainty. We found that uncertainty was a frequent focus of TBD social media across platforms but was primarily limited to scientific issues, requests for informational support, and offers of emotional support, with most posts generated by White, female, English-speaking parents of children with TBDs. These findings are in keeping with other research on rare disease internet-based communities, which found that post content focused on biomedical questions and emotional support provision [63] and was frequently created by White, female users [40,63-65].

The high frequency of uncertainty-related posts on TBD social media created by female caregivers suggests a potentially higher burden of uncertainty management among mothers, which is in agreement with the extensive literature documenting the psychosocial burden of childhood illness on female caregivers [66-68]. However, the observed demographics of TBD social media users may also be an artifact of greater social media engagement among this group, as previous research suggests that female users frequently rely on internet-based communities for navigating uncertainty related to motherhood and other sex-specific health topics [69,70]. Additional research is needed to investigate the relative burden of medical uncertainty among female care providers and to understand the potential barriers to internet-based community formation for users outside this identity group.

Despite the multiplicity of identified uncertainty sources, issues, management, and social support strategies, we found that scientific uncertainty, informational support, and emotional support were the predominant features of uncertainty-related posts on TBD social media. The high frequency of scientific uncertainty issues across platforms suggests that limited scientific and medical knowledge is a salient concern for the TBD community. Gaps in scientific knowledge likely contribute to the focus on probability as a source of uncertainty in TBD

social media posts, especially concerning matters such as prognosis, diagnosis, and symptom experiences. Informational support was the most common form of social support in uncertainty-related posts overall, which is in line with other studies showing information seeking as the principal motivator for participation in disease-specific social media [24,26,71-73]. The high frequency of emotional support suggests the potential for TBD social media to enable uncertainty management through person-focused strategies, such as community building, networking, and relationship formation, as seen in other rare disease contexts [24,72]. In addition, evidence of positive asynchronous internet-based communication as a form of “cybertherapy” [32,44] suggests that the emotionally supportive culture of TBD social media may provide psychological benefits for peers, even without explicit conversations about the personal burden of uncertainty. In addition, items coded as emotional support (eg, emoji hearts) that appeared in response to a variety of uncertainty-related content may have communicated multiple forms of support (eg, care, approval, agreement, or affinity) and may be a common reaction to intractable sources of uncertainty, such as probabilistic and scientific unknowns surrounding TBDs. Further exploration of the complex, dynamic, and potentially interactive relationships between social support and uncertainty on social media may be a fruitful area of investigation for future studies.

Given the evidence of the high psychosocial burden of personal uncertainty in similar rare disease contexts [18,36,74,75], it is surprising that the mental and emotional impacts of uncertainty appeared infrequently in TBD social media discussions. When these topics did arise, they were more likely to appear on Twitter content generated by Team Telomere, as opposed to within the conversations of individual users. In the Facebook community group, the impact of uncertainty on personal life was commonly presented in terms of practical issues and focused on ordering uncertainty, such as providing lists of symptoms, organizing information and screening schedules, and triaging problems. This suggests that despite the frequent focus on personal uncertainty issues by Team Telomere, most individual users engaged with TBD social media to troubleshoot and strategize practical issues, rather than to discuss the impact of uncertainty on areas of psychosocial well-being, such as personal identity, goals, or values. This is also reflected in the positive sentiment valence and rare expression of negative emotion on TBD social media, which suggest that social media may not be perceived as a “safe space” for exploring personal topics beyond surface-level stressors [23]. Future research is needed to investigate the shortcomings of social media for expressing personal uncertainty and painful emotions and may highlight a need for psychosocial support to fill this gap in TBD community resources.

Our finding that uncertainty-related support varied by platform could be explained by differences in the structure and expectations of engagement inherent to Twitter compared with the Facebook community group. The predominance of emotional support and greater overall user engagement in the Facebook community group suggests that internet-based platforms structured for mutual conversational exchange may have the most utility for psychosocial support delivery. In addition, the

Facebook community group may have encouraged more community participation owing to user familiarity with the platform and its explicit creation for supportive internet-based connection in the context of COVID-19 isolation. Similarly, the nature of the Twitter platform, which is limited to one-way communication streams, suggests that uncertainty management and social support on Twitter would be limited to information provision. However, recent research indicates that Twitter retweets and endorsements may be effective methods for receiving and providing emotional support [76]. The formation of the Facebook community group and the use of Twitter to encourage community activities (eg, webinars and internet-based meetups) underscores the potential of these platforms in person-focused uncertainty management, but additional research is required to evaluate the capacity of TBD social media to build health-promoting personal relationships.

Although we found substantial potential for social media to deliver support for uncertainty management, analysis of engagement rates demonstrated that the primary function of TBD social media was a “drop-in” source of information. Although the Facebook community group included some multilevel, ongoing conversations, an analysis of posts within this group revealed that most user engagement was limited to single posts, suggesting quick check-ins or requests for answers to targeted questions, not ongoing social connection. Although low engagement may suggest limited supportive utility of TBD social media, findings from previous research with young adults with cancer showed that support delivered via social media benefited a variety of users, including those actively seeking deep connections, those seeking information only, and those who do not actively participate but frequently observe the conversation of others (eg, “lurkers”) [77]. As suggested by other research, any benefit from engagement with social media likely varies over time and may be most pronounced during experiences of novelty or discrepancy in diagnosis, treatment, or prognosis [28,48,63]. The uncertainty-related post that generated the highest engagement involved the participation of a medical expert, suggesting a desire among TBD social media users to engage with clinicians on internet-based platforms that facilitate reciprocal information exchange, including both synchronous (eg, internet-based group meetings) and asynchronous (eg, post exchanges) formats. Further research is needed to understand the motivations, perceived benefits, and perceived barriers to participation in TBD internet-based support platforms, including the perspectives of patients, caregivers, and medical providers.

Limitations

The limitations of our study include the use of social media data, which biases our sample toward active social media users who may have higher levels of distress [64], greater disenchantment with medical care [78], or lower perceived social support [79] compared with patients with TBDs and their families who do not actively use social media. Demographic analysis revealed that our sample of posts was generated primarily by White females, parents of patients with TBDs, or representatives of Team Telomere. This limited the generalizability of our findings. In addition, our use of social media posts, rather than content creators, as the unit of analysis

precludes the observation of the longitudinal impacts of social media participation on uncertainty management. Furthermore, our findings allow us to infer the presence of uncertainty management strategies on social media but not the motivations for or effects of these activities.

In addition, our data were limited to social media that was actively moderated by Team Telomere. This moderation activity, which included removing posts that were inappropriate or scientifically inaccurate, likely decreased the presence of medical misinformation compared with unmoderated social media content. The moderation of posts by Team Telomere could also have impacted the range and authenticity of social and emotional expression owing to social desirability bias. This is in keeping with recent research challenging the assumption that the privacy and anonymity of internet-based environments decreases the likelihood of social desirability compared with in-person interactions [80,81]. In addition, we did not access the private Facebook community group maintained by Team Telomere described as “where we share detailed and private medical information” [57], which may contain additional uncertainty-related posts and a wider range of social and emotional expression. Limiting ourselves to social media owned and maintained by Team Telomere also prevented us from discerning the perspectives of individuals affected by TBD who lacked knowledge of or who chose not to engage with Team Telomere.

Finally, our study was limited by the occurrence of the COVID-19 pandemic, first mentioned in Team Telomere social media on February 28, 2020, which may have changed the nature of uncertainty-related conversations or social support in that portion of our data timeline (June 6, 2019, to December 7, 2021). To test the impact of this, we included available posts

(Twitter and Facebook main page) from 1 year before the pandemic and tested the difference. Greater frequencies of uncertainty-related posts after COVID-19 suggest that the pandemic may have increased the expression of uncertainty on TBD-related social media, thus limiting the applicability of our findings to other time points ([Multimedia Appendix 7](#)).

Conclusions

This study found the frequent use of disease-specific social media for the discussion and management of uncertainty in TBDs. Uncertainty-related posts appeared across all TBD social media platforms and communicated a burden of multiple, often interacting sources and issues of uncertainty, particularly focused on scientific knowledge gaps and the desire to predict health outcomes. Posts also indicated multiple uncertainty management attributes, with a focus on information-seeking and community-building approaches. Uncertainty-related posts frequently co-occurred with social support, primarily emotional and informational. Female parents were most often the creators of uncertainty-related posts on TBD social media, suggesting a potentially higher burden of uncertainty management in this population. Overall, social media provided access to a positive emotional environment and frequent information exchange but was limited in the type and depth of uncertainty-related discussions. Despite these limitations, our findings suggest that social media is a useful lens for researching and understanding the experience of uncertainty in TBDs and holds potential as a tool for uncertainty management. Future research is needed to further explore the experience of medical uncertainty in TBDs and to determine the usefulness of TBD-related social media as a tool for improving mental health and quality of life outcomes in this context.

Acknowledgments

This study was supported by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics, National Cancer Institute. Katherine Stevens from Team Telomere facilitated our access to the publicly available social media data.

Data Availability

In compliance with the National Institutes of Health data management and sharing policy, data, analysis code, and research materials are available upon reasonable request from the corresponding author.

Authors' Contributions

EP contributed to the study design, formative research, data collection, data analysis, codebook development, coding, and manuscript preparation; HR contributed to data collection, codebook development, and coding; PKJH contributed to the study design, codebook development, coding, and manuscript preparation; MBG, KMR, and AJL contributed to the study design, codebook development, and manuscript preparation; SAS contributed to manuscript preparation, study primary investigator, and National Institutes of Health; and NE contributed to data quality control and coding.

Conflicts of Interest

SAS and HR are members of Team Telomere Advisory Boards.

Multimedia Appendix 1

Criteria for identification of posts for inclusion in qualitative uncertainty analysis.

[[DOCX File, 21 KB - infodemiology_v4i1e46693_app1.docx](#)]

Multimedia Appendix 2

Social media study codebook.

[\[DOCX File, 20 KB - infodemiology_v4i1e46693_app2.docx\]](#)

Multimedia Appendix 3

Emoji dictionary.

[\[DOCX File, 26 KB - infodemiology_v4i1e46693_app3.docx\]](#)

Multimedia Appendix 4

Frequency of social support by support type and direction.

[\[DOCX File, 262 KB - infodemiology_v4i1e46693_app4.docx\]](#)

Multimedia Appendix 5

Engagement and popularity by platform.

[\[DOCX File, 14 KB - infodemiology_v4i1e46693_app5.docx\]](#)

Multimedia Appendix 6

Sentiment by post type.

[\[DOCX File, 16 KB - infodemiology_v4i1e46693_app6.docx\]](#)

Multimedia Appendix 7

COVID-19 impact summary.

[\[DOCX File, 13 KB - infodemiology_v4i1e46693_app7.docx\]](#)**References**

1. Carpenter DM, DeVellis RF, Fisher EB, DeVellis BM, Hogan SL, Jordan JM. The effect of conflicting medication information and physician support on medication adherence for chronically ill patients. *Patient Educ Couns* 2010 Nov;81(2):169-176 [FREE Full text] [doi: [10.1016/j.pec.2009.11.006](#)] [Medline: [20044230](#)]
2. Han PK, Moser RP, Klein WM. Perceived ambiguity about cancer prevention recommendations: relationship to perceptions of cancer preventability, risk, and worry. *J Health Commun* 2006;11 Suppl 1(0 1):51-69 [FREE Full text] [doi: [10.1080/10810730600637541](#)] [Medline: [16641074](#)]
3. Santacroce S. Uncertainty, anxiety, and symptoms of posttraumatic stress in parents of children recently diagnosed with cancer. *J Pediatr Oncol Nurs* 2002;19(3):104-111. [doi: [10.1177/104345420201900305](#)] [Medline: [12066262](#)]
4. Savage SA. Dyskeratosis congenita and telomere biology disorders. *Hematology Am Soc Hematol Educ Program* 2022 Dec 09;2022(1):637-648 [FREE Full text] [doi: [10.1182/hematology.2022000394](#)] [Medline: [36485133](#)]
5. Niewisch MR, Giri N, McReynolds LJ, Alsaggaf R, Bhala S, Alter BP, et al. Disease progression and clinical outcomes in telomere biology disorders. *Blood* 2022 Mar 24;139(12):1807-1819 [FREE Full text] [doi: [10.1182/blood.2021013523](#)] [Medline: [34852175](#)]
6. Fortier MA, Batista ML, Wahi A, Kain A, Strom S, Sender LS. Illness uncertainty and quality of life in children with cancer. *J Pediatr Hematol Oncol* 2013 Jul;35(5):366-370. [doi: [10.1097/MPH.0b013e318290cfdb](#)] [Medline: [23669725](#)]
7. Hamilton JG, Hutson SP, Moser RP, Kobrin SC, Frohnmayer AE, Alter BP, et al. Sources of uncertainty and their association with medical decision making: exploring mechanisms in Fanconi anemia. *Ann Behav Med* 2013 Oct;46(2):204-216. [doi: [10.1007/s12160-013-9507-5](#)] [Medline: [23637072](#)]
8. Hill EM, Hamm A. Intolerance of uncertainty, social support, and loneliness in relation to anxiety and depressive symptoms among women diagnosed with ovarian cancer. *Psychooncology* 2019 Mar;28(3):553-560. [doi: [10.1002/pon.4975](#)] [Medline: [30614141](#)]
9. Kim B, Kim J. Influence of uncertainty, depression, and social support on self-care compliance in hemodialysis patients. *Ther Clin Risk Manag* 2019 Oct 22;15:1243-1251 [FREE Full text] [doi: [10.2147/TCRM.S218934](#)] [Medline: [31695397](#)]
10. Neville K. The relationships among uncertainty, social support, and psychological distress in adolescents recently diagnosed with cancer. *J Pediatr Oncol Nurs* 1998 Jan;15(1):37-46. [doi: [10.1177/104345429801500106](#)] [Medline: [9473892](#)]
11. Cohen MH. The triggers of heightened parental uncertainty in chronic, life-threatening childhood illness. *Qual Health Res* 1995;5(1):63-77. [doi: [10.1177/104973239500500105](#)]
12. Han PK. *Uncertainty in Medicine: A Framework for Tolerance*. Oxford, UK: Oxford University Press; 2021.
13. Hupcey JE. Clarifying the social support theory-research linkage. *J Adv Nurs* 1998 Jun;27(6):1231-1241. [doi: [10.1046/j.1365-2648.1998.01231.x](#)] [Medline: [9663875](#)]
14. Holt-Lunstad J, Uchino BN. Social support and health. In: Glanz K, Rimer BK, Viswanath KV, editors. *Health Behavior: Theory, Research, and Practice*. San Francisco, CA: Jossey-Bass Publishers; 2015:183-204.

15. Zhuo L, Wu Q, Le H, Li H, Zheng L, Ma G, et al. COVID-19-related intolerance of uncertainty and mental health among back-to-school students in Wuhan: the moderation effect of social support. *Int J Environ Res Public Health* 2021 Jan 22;18(3):981 [FREE Full text] [doi: [10.3390/ijerph18030981](https://doi.org/10.3390/ijerph18030981)] [Medline: [33499409](https://pubmed.ncbi.nlm.nih.gov/33499409/)]
16. He X, Zhang Y, Chen M, Zhang J, Zou W, Luo Y. Media exposure to COVID-19 predicted acute stress: a moderated mediation model of intolerance of uncertainty and perceived social support. *Front Psychiatry* 2021 Feb 10;11:613368. [doi: [10.3389/fpsy.2020.613368](https://doi.org/10.3389/fpsy.2020.613368)] [Medline: [33643082](https://pubmed.ncbi.nlm.nih.gov/33643082/)]
17. Lee I, Park C. The mediating effect of social support on uncertainty in illness and quality of life of female cancer survivors: a cross-sectional study. *Health Qual Life Outcomes* 2020 May 19;18(1):143 [FREE Full text] [doi: [10.1186/s12955-020-01392-2](https://doi.org/10.1186/s12955-020-01392-2)] [Medline: [32429954](https://pubmed.ncbi.nlm.nih.gov/32429954/)]
18. Peters JA, Kenen R, Bremer R, Givens S, Savage SA, Mai PL. Easing the burden: describing the role of social, emotional and spiritual support in research families with Li-Fraumeni syndrome. *J Genet Couns* 2016 Jun;25(3):529-542. [doi: [10.1007/s10897-015-9905-x](https://doi.org/10.1007/s10897-015-9905-x)] [Medline: [26621765](https://pubmed.ncbi.nlm.nih.gov/26621765/)]
19. El Hussein S, Lyapichev KA, Crane GM, Mirza KM, Pemmaraju N, Medeiros LJ, et al. Social media for hematopathologists: medical practice reinvented-#Hemepath. *Curr Hematol Malig Rep* 2020 Oct;15(5):383-390. [doi: [10.1007/s11899-020-00600-6](https://doi.org/10.1007/s11899-020-00600-6)] [Medline: [33128122](https://pubmed.ncbi.nlm.nih.gov/33128122/)]
20. Savage N. Scientists in the Twittersverse. *Cell* 2015 Jul 16;162(2):233-234 [FREE Full text] [doi: [10.1016/j.cell.2015.06.062](https://doi.org/10.1016/j.cell.2015.06.062)] [Medline: [26186181](https://pubmed.ncbi.nlm.nih.gov/26186181/)]
21. Pemmaraju N, Utengen A, Gupta V, Thompson MA, Lane AA. Analysis of first-year Twitter metrics of a rare disease community for blastic plasmacytoid dendritic cell neoplasm (BPDCN) on social media: #BPDCN. *Curr Hematol Malig Rep* 2017 Dec;12(6):592-597. [doi: [10.1007/s11899-017-0422-x](https://doi.org/10.1007/s11899-017-0422-x)] [Medline: [29064025](https://pubmed.ncbi.nlm.nih.gov/29064025/)]
22. Brown SA, Daly RP, Duma N, Yang EH, Pemmaraju N, Parwani P, et al. Leveraging social media for cardio-oncology. *Curr Treat Options Oncol* 2020 Aug 13;21(10):83. [doi: [10.1007/s11864-020-00775-3](https://doi.org/10.1007/s11864-020-00775-3)] [Medline: [32789716](https://pubmed.ncbi.nlm.nih.gov/32789716/)]
23. Benson JJ, Oliver DP, Washington KT, Rolbiecki AJ, Lombardo CB, Garza JE, et al. Online social support groups for informal caregivers of hospice patients with cancer. *Eur J Oncol Nurs* 2020 Feb;44:101698 [FREE Full text] [doi: [10.1016/j.ejon.2019.101698](https://doi.org/10.1016/j.ejon.2019.101698)] [Medline: [31816508](https://pubmed.ncbi.nlm.nih.gov/31816508/)]
24. Barton KS, Wingerson A, Barzilay JR, Tabor HK. "Before Facebook and before social media...we did not know anybody else that had this": parent perspectives on internet and social media use during the pediatric clinical genetic testing process. *J Community Genet* 2019 Jul;10(3):375-383 [FREE Full text] [doi: [10.1007/s12687-018-0400-6](https://doi.org/10.1007/s12687-018-0400-6)] [Medline: [30569339](https://pubmed.ncbi.nlm.nih.gov/30569339/)]
25. Dizon DS, Graham D, Thompson MA, Johnson LJ, Johnston C, Fisch MJ, et al. Practical guidance: the use of social media in oncology practice. *J Oncol Pract* 2012 Sep;8(5):e114-e124 [FREE Full text] [doi: [10.1200/JOP.2012.000610](https://doi.org/10.1200/JOP.2012.000610)] [Medline: [23277774](https://pubmed.ncbi.nlm.nih.gov/23277774/)]
26. Iyer AA, Barzilay JR, Tabor HK. Patient and family social media use surrounding a novel treatment for a rare genetic disease: a qualitative interview study. *Genet Med* 2020 Nov;22(11):1830-1837 [FREE Full text] [doi: [10.1038/s41436-020-0890-6](https://doi.org/10.1038/s41436-020-0890-6)] [Medline: [32601388](https://pubmed.ncbi.nlm.nih.gov/32601388/)]
27. Gallagher L, McCuaig J, Benoit L, Davies C. It's time for the genetic counseling profession to embrace social media. *J Genet Couns* 2016 Dec;25(6):1338-1341. [doi: [10.1007/s10897-016-9950-0](https://doi.org/10.1007/s10897-016-9950-0)] [Medline: [27026235](https://pubmed.ncbi.nlm.nih.gov/27026235/)]
28. Smedley RM, Coulson NS. Genetic testing for Huntington's disease: a thematic analysis of online support community messages. *J Health Psychol* 2021 Mar;26(4):580-594. [doi: [10.1177/1359105319826340](https://doi.org/10.1177/1359105319826340)] [Medline: [30696276](https://pubmed.ncbi.nlm.nih.gov/30696276/)]
29. Westmaas JL, Fallon E, McDonald BR, Driscoll D, Richardson K, Portier K, et al. Investigating relationships among cancer survivors' engagement in an online support community, social support perceptions, well-being, and moderating effects of existing (offline) social support. *Support Care Cancer* 2020 Aug;28(8):3791-3799. [doi: [10.1007/s00520-019-05193-2](https://doi.org/10.1007/s00520-019-05193-2)] [Medline: [31828494](https://pubmed.ncbi.nlm.nih.gov/31828494/)]
30. Rising CJ, Bol N, Burke-Garcia A, Rains S, Wright KB. Perceived stress in online prostate cancer community participants: examining relationships with stigmatization, social support network preference, and social support seeking. *J Health Commun* 2017 Jun;22(6):469-476. [doi: [10.1080/10810730.2017.1304471](https://doi.org/10.1080/10810730.2017.1304471)] [Medline: [28414570](https://pubmed.ncbi.nlm.nih.gov/28414570/)]
31. Rimer BK, Lyons EJ, Ribisl KM, Bowling JM, Golin CE, Forlenza MJ, et al. How new subscribers use cancer-related online mailing lists. *J Med Internet Res* 2005 Jul 01;7(3):e32 [FREE Full text] [doi: [10.2196/jmir.7.3.e32](https://doi.org/10.2196/jmir.7.3.e32)] [Medline: [15998623](https://pubmed.ncbi.nlm.nih.gov/15998623/)]
32. Lorence D. Examining online chat within a domain of uncertainty: the case of Asperger's syndrome. *Health Info Libr J* 2007 Jun;24(2):128-136 [FREE Full text] [doi: [10.1111/j.1471-1842.2007.00715.x](https://doi.org/10.1111/j.1471-1842.2007.00715.x)] [Medline: [17584216](https://pubmed.ncbi.nlm.nih.gov/17584216/)]
33. Scott AM, Martin SC, Stone AM, Brashers DE. Managing multiple goals in supportive interactions: using a normative theoretical approach to explain social support as uncertainty management for organ transplant patients. *Health Commun* 2011;26(5):393-403. [doi: [10.1080/10410236.2011.552479](https://doi.org/10.1080/10410236.2011.552479)] [Medline: [21409670](https://pubmed.ncbi.nlm.nih.gov/21409670/)]
34. Sammarco A. Perceived social support, uncertainty, and quality of life of younger breast cancer survivors. *Cancer Nurs* 2001 Jun;24(3):212-219. [Medline: [11409065](https://pubmed.ncbi.nlm.nih.gov/11409065/)]
35. Li X, He L, Wang J, Wang M. Illness uncertainty, social support, and coping mode in hospitalized patients with systemic lupus erythematosus in a hospital in Shaanxi, China. *PLoS One* 2019 Feb 21;14(2):e0211313 [FREE Full text] [doi: [10.1371/journal.pone.0211313](https://doi.org/10.1371/journal.pone.0211313)] [Medline: [30789919](https://pubmed.ncbi.nlm.nih.gov/30789919/)]

36. Forbes Shepherd R, Werner-Lin A, Keogh LA, Delatycki MB, Forrest LE. "I need to know if I'm going to die young": adolescent and young adult experiences of genetic testing for Li-Fraumeni syndrome. *J Psychosoc Oncol* 2021;39(1):54-73. [doi: [10.1080/07347332.2020.1768199](https://doi.org/10.1080/07347332.2020.1768199)] [Medline: [32449501](#)]
37. Eche JJ, Aronowitz T. Factors that influence parental uncertainty and health-related quality of life in children with cancer: a framework. *Nurs Sci Q* 2018 Oct;31(4):362-368. [doi: [10.1177/0894318418792896](https://doi.org/10.1177/0894318418792896)] [Medline: [30223741](#)]
38. Oprea F, Campo S, Lowe J, Andsager J, Morcuende JA. Managing uncertainty in the context of clubfoot care: exploring the value of uncertainty management theory and the sense of virtual community. *Iowa Orthop J* 2013;33:142-148 [FREE Full text] [Medline: [24027474](#)]
39. Miller EG, Woodward AL, Flinchum G, Young JL, Tabor HK, Halley MC. Opportunities and pitfalls of social media research in rare genetic diseases: a systematic review. *Genet Med* 2021 Dec;23(12):2250-2259 [FREE Full text] [doi: [10.1038/s41436-021-01273-z](https://doi.org/10.1038/s41436-021-01273-z)] [Medline: [34282302](#)]
40. Fogel J, Ribisl KM, Morgan PD, Humphreys K, Lyons EJ. Underrepresentation of African Americans in online cancer support groups. *J Natl Med Assoc* 2008 Jun;100(6):705-712. [doi: [10.1016/s0027-9684\(15\)31346-8](https://doi.org/10.1016/s0027-9684(15)31346-8)] [Medline: [18595573](#)]
41. Johnson SB, Parsons M, Dorff T, Moran MS, Ward JH, Cohen SA, et al. Cancer misinformation and harmful information on Facebook and other social media: a brief report. *J Natl Cancer Inst* 2022 Jul 11;114(7):1036-1039 [FREE Full text] [doi: [10.1093/jnci/djab141](https://doi.org/10.1093/jnci/djab141)] [Medline: [34291289](#)]
42. Chrousos GP, Mentis AF. Medical misinformation in mass and social media: an urgent call for action, especially during epidemics. *Eur J Clin Invest* 2020 May;50(5):e13227. [doi: [10.1111/eci.13227](https://doi.org/10.1111/eci.13227)] [Medline: [32294232](#)]
43. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. *J Med Internet Res* 2021 Jan 20;23(1):e17187 [FREE Full text] [doi: [10.2196/17187](https://doi.org/10.2196/17187)] [Medline: [33470931](#)]
44. Beaudoin CE, Tao CC. Benefiting from social capital in online support groups: an empirical study of cancer patients. *Cyberpsychol Behav* 2007 Aug;10(4):587-590 [FREE Full text] [doi: [10.1089/cpb.2007.9986](https://doi.org/10.1089/cpb.2007.9986)] [Medline: [17711369](#)]
45. Bloom RD, Reblin M, Chou WY, Beck SL, Wilson A, Ellington L. Online social support for cancer caregivers: alignment between requests and offers on CaringBridge. *J Psychosoc Oncol* 2021;39(1):118-134. [doi: [10.1080/07347332.2020.1806174](https://doi.org/10.1080/07347332.2020.1806174)] [Medline: [32835649](#)]
46. Coulson NS, Greenwood N. Families affected by childhood cancer: an analysis of the provision of social support within online support groups. *Child Care Health Dev* 2012 Nov;38(6):870-877. [doi: [10.1111/j.1365-2214.2011.01316.x](https://doi.org/10.1111/j.1365-2214.2011.01316.x)] [Medline: [21916931](#)]
47. LaCoursiere SP. A theory of online social support. *ANS Adv Nurs Sci* 2001 Sep;24(1):60-77. [doi: [10.1097/00012272-200109000-00008](https://doi.org/10.1097/00012272-200109000-00008)] [Medline: [11554534](#)]
48. Mikal JP, Beckstrand MJ, Parks E, Oyenuga M, Odeunmi T, Okedele O, et al. Online social support among breast cancer patients: longitudinal changes to Facebook use following breast cancer diagnosis and transition off therapy. *J Cancer Surviv* 2020 Jun;14(3):322-330. [doi: [10.1007/s11764-019-00847-w](https://doi.org/10.1007/s11764-019-00847-w)] [Medline: [31897878](#)]
49. DeHoff BA, Staten LK, Rodgers RC, Denne SC. The role of online social support in supporting and educating parents of young children with special health care needs in the United States: a scoping review. *J Med Internet Res* 2016 Dec 22;18(12):e333 [FREE Full text] [doi: [10.2196/jmir.6722](https://doi.org/10.2196/jmir.6722)] [Medline: [28007689](#)]
50. Ye Z, Li W, Zhu R. Online psychosocial interventions for improving mental health in people during the COVID-19 pandemic: a systematic review and meta-analysis. *J Affect Disord* 2022 Nov 01;316:120-131 [FREE Full text] [doi: [10.1016/j.jad.2022.08.023](https://doi.org/10.1016/j.jad.2022.08.023)] [Medline: [35970325](#)]
51. Hong Y, Peña-Purcell NC, Ory MG. Outcomes of online support and resources for cancer survivors: a systematic literature review. *Patient Educ Couns* 2012 Mar;86(3):288-296. [doi: [10.1016/j.pec.2011.06.014](https://doi.org/10.1016/j.pec.2011.06.014)] [Medline: [21798685](#)]
52. Politi MC, Han PK, Col NF. Communicating the uncertainty of harms and benefits of medical interventions. *Med Decis Making* 2007;27(5):681-695. [doi: [10.1177/0272989X07307270](https://doi.org/10.1177/0272989X07307270)] [Medline: [17873256](#)]
53. Team Telomere homepage. Team Telomere. URL: <https://teamtelomere.org/> [accessed 2023-12-20]
54. Facca D, Smith MJ, Shelley J, Lizotte D, Donelle L. Exploring the ethical issues in research using digital data collection strategies with minors: a scoping review. *PLoS One* 2020 Aug 27;15(8):e0237875 [FREE Full text] [doi: [10.1371/journal.pone.0237875](https://doi.org/10.1371/journal.pone.0237875)] [Medline: [32853218](#)]
55. Team Telomere Twitter. Twitter. URL: <https://twitter.com/teamtelomereinc?lang=en> [accessed 2023-12-20]
56. Team Telomere. Facebook. URL: <https://www.facebook.com/teamtelomere/> [accessed 2023-12-20]
57. Team telomere community. Facebook. URL: <https://www.facebook.com/groups/244926586962809/> [accessed 2023-12-22]
58. Hamad EO, Savundranayagam MY, Holmes JD, Kinsella EA, Johnson AM. Toward a mixed-methods research approach to content analysis in the digital age: the combined content-analysis model and its applications to health care Twitter feeds. *J Med Internet Res* 2016 Mar 08;18(3):e60 [FREE Full text] [doi: [10.2196/jmir.5391](https://doi.org/10.2196/jmir.5391)] [Medline: [26957477](#)]
59. Cohen LH, McGowan J, Fooskas S, Rose S. Positive life events and social support and the relationship between life stress and psychological disorder. *Am J Community Psychol* 1984 Oct;12(5):567-587. [doi: [10.1007/BF00897213](https://doi.org/10.1007/BF00897213)] [Medline: [6496413](#)]
60. Cutrona C, Russell D, Rose J. Social support and adaptation to stress by the elderly. *Psychol Aging* 1986;1(1):47-54 [FREE Full text] [doi: [10.1037/0882-7974.1.1.47](https://doi.org/10.1037/0882-7974.1.1.47)]

61. Gendron M, Lindquist KA, Barsalou L, Barrett LF. Emotion words shape emotion percepts. *Emotion* 2012 Apr;12(2):314-325 [FREE Full text] [doi: [10.1037/a0026007](https://doi.org/10.1037/a0026007)] [Medline: [22309717](https://pubmed.ncbi.nlm.nih.gov/22309717/)]
62. Barrett LF, Lindquist KA, Gendron M. Language as context for the perception of emotion. *Trends Cogn Sci* 2007 Aug;11(8):327-332 [FREE Full text] [doi: [10.1016/j.tics.2007.06.003](https://doi.org/10.1016/j.tics.2007.06.003)] [Medline: [17625952](https://pubmed.ncbi.nlm.nih.gov/17625952/)]
63. Lasker JN, Sogolow ED, Sharim RR. The role of an online community for people with a rare disease: content analysis of messages posted on a primary biliary cirrhosis mailinglist. *J Med Internet Res* 2005 Mar 31;7(1):e10 [FREE Full text] [doi: [10.2196/jmir.7.1.e10](https://doi.org/10.2196/jmir.7.1.e10)] [Medline: [15829472](https://pubmed.ncbi.nlm.nih.gov/15829472/)]
64. Bender JL, Hueniken K, Eng L, Brown MC, Kassirian S, Geist I, et al. Internet and social media use in cancer patients: association with distress and perceived benefits and limitations. *Support Care Cancer* 2021 Sep;29(9):5273-5281. [doi: [10.1007/s00520-021-06077-0](https://doi.org/10.1007/s00520-021-06077-0)] [Medline: [33651181](https://pubmed.ncbi.nlm.nih.gov/33651181/)]
65. Claus EB, Feliciano J, Benz LS, Calvocoressi L. Social media partnerships with patient organizations for neuro-oncology patient recruitment. *Neurooncol Pract* 2020 Mar;7(2):143-151 [FREE Full text] [doi: [10.1093/nop/npz049](https://doi.org/10.1093/nop/npz049)] [Medline: [32626583](https://pubmed.ncbi.nlm.nih.gov/32626583/)]
66. Thompson RJ, Gil KM, Gustafson KE, George LK, Keith BR, Spock A, et al. Stability and change in the psychological adjustment of mothers of children and adolescents with cystic fibrosis and sickle cell disease. *J Pediatr Psychol* 1994 Apr;19(2):171-188. [doi: [10.1093/jpepsy/19.2.171](https://doi.org/10.1093/jpepsy/19.2.171)] [Medline: [8051601](https://pubmed.ncbi.nlm.nih.gov/8051601/)]
67. Cai RY, Uljarević M, Leekam SR. Predicting mental health and psychological wellbeing in mothers of children with autism spectrum disorder: roles of intolerance of uncertainty and coping. *Autism Res* 2020 Oct;13(10):1797-1801. [doi: [10.1002/aur.2341](https://doi.org/10.1002/aur.2341)] [Medline: [32618140](https://pubmed.ncbi.nlm.nih.gov/32618140/)]
68. Roorda D, van der Steeg AF, van Dijk M, Derikx JP, Gorter RR, Rotteveel J, et al. Distress and post-traumatic stress in parents of patients with congenital gastrointestinal malformations: a cross-sectional cohort study. *Orphanet J Rare Dis* 2022 Sep 11;17(1):353 [FREE Full text] [doi: [10.1186/s13023-022-02502-7](https://doi.org/10.1186/s13023-022-02502-7)] [Medline: [36089585](https://pubmed.ncbi.nlm.nih.gov/36089585/)]
69. Wright KB, Cai X, Fisher C, Rising CJ, Burke-Garcia A, Afanaseva D. A content analysis of social support messages about environmental breast cancer risk within blogs for mothers. *Health Commun* 2021 Nov;36(13):1796-1804 [FREE Full text] [doi: [10.1080/10410236.2020.1800241](https://doi.org/10.1080/10410236.2020.1800241)] [Medline: [32744079](https://pubmed.ncbi.nlm.nih.gov/32744079/)]
70. Wagg AJ, Callanan MM, Hassett A. Online social support group use by breastfeeding mothers: a content analysis. *Heliyon* 2019 Mar 09;5(3):e01245 [FREE Full text] [doi: [10.1016/j.heliyon.2019.e01245](https://doi.org/10.1016/j.heliyon.2019.e01245)] [Medline: [30906891](https://pubmed.ncbi.nlm.nih.gov/30906891/)]
71. Jacobs R, Boyd L, Brennan K, Sinha CK, Giuliani S. The importance of social media for patients and families affected by congenital anomalies: a Facebook cross-sectional analysis and user survey. *J Pediatr Surg* 2016 Nov;51(11):1766-1771. [doi: [10.1016/j.jpedsurg.2016.07.008](https://doi.org/10.1016/j.jpedsurg.2016.07.008)] [Medline: [27522307](https://pubmed.ncbi.nlm.nih.gov/27522307/)]
72. Khouri JS, McCheyne MJ, Morrison CS. #Cleft: the use of social media amongst parents of infants with clefts. *Cleft Palate Craniofac J* 2018 Aug;55(7):974-976. [doi: [10.1597/16-156](https://doi.org/10.1597/16-156)] [Medline: [28085513](https://pubmed.ncbi.nlm.nih.gov/28085513/)]
73. Pemmaraju N, Gupta V, Mesa R, Thompson MA. Social media and myeloproliferative neoplasms (MPN)--focus on Twitter and the development of a disease-specific community: #MPNSM. *Curr Hematol Malig Rep* 2015 Dec;10(4):413-420 [FREE Full text] [doi: [10.1007/s11899-015-0287-9](https://doi.org/10.1007/s11899-015-0287-9)] [Medline: [26411990](https://pubmed.ncbi.nlm.nih.gov/26411990/)]
74. Ross J, Bojadzieva J, Peterson S, Noblin SJ, Yzquierdo R, Askins M, et al. The psychosocial effects of the Li-Fraumeni education and early detection (LEAD) program on individuals with Li-Fraumeni syndrome. *Genet Med* 2017 Sep;19(9):1064-1070 [FREE Full text] [doi: [10.1038/gim.2017.8](https://doi.org/10.1038/gim.2017.8)] [Medline: [28301458](https://pubmed.ncbi.nlm.nih.gov/28301458/)]
75. Young JL, Pantaleao A, Zaspel L, Bayer J, Peters JA, Khincha PP, et al. Couples coping with screening burden and diagnostic uncertainty in Li-Fraumeni syndrome: connection versus independence. *J Psychosoc Oncol* 2019;37(2):178-193 [FREE Full text] [doi: [10.1080/07347332.2018.1543376](https://doi.org/10.1080/07347332.2018.1543376)] [Medline: [30591002](https://pubmed.ncbi.nlm.nih.gov/30591002/)]
76. Esener Y, McCall T, Lakdawala A, Kim H. Seeking and providing social support on Twitter for trauma and distress during the COVID-19 pandemic: content and sentiment analysis. *J Med Internet Res* 2023 Aug 31;25:e46343 [FREE Full text] [doi: [10.2196/46343](https://doi.org/10.2196/46343)] [Medline: [37651178](https://pubmed.ncbi.nlm.nih.gov/37651178/)]
77. Lazard AJ, Collins MK, Hedrick A, Horrell LN, Varma T, Love B, et al. Initiation and changes in use of social media for peer support among young adult cancer patients and survivors. *Psychooncology* 2021 Nov;30(11):1859-1865. [doi: [10.1002/pon.5758](https://doi.org/10.1002/pon.5758)] [Medline: [34165848](https://pubmed.ncbi.nlm.nih.gov/34165848/)]
78. Dolce MC. The Internet as a source of health information: experiences of cancer survivors and caregivers with healthcare providers. *Oncol Nurs Forum* 2011 May;38(3):353-359. [doi: [10.1188/11.ONF.353-359](https://doi.org/10.1188/11.ONF.353-359)] [Medline: [21531685](https://pubmed.ncbi.nlm.nih.gov/21531685/)]
79. Kim SC, Shah DV, Namkoong K, McTavish FM, Gustafson DH. Predictors of online health information seeking among women with breast cancer: the role of social support perception and emotional well-being. *J Comput Mediat Commun* 2013 Jan;18(2):98-118 [FREE Full text] [doi: [10.1111/jcc4.12002](https://doi.org/10.1111/jcc4.12002)] [Medline: [24634575](https://pubmed.ncbi.nlm.nih.gov/24634575/)]
80. Dodou D, de Winter JC. Social desirability is the same in offline, online, and paper surveys: a meta-analysis. *Comput Hum Behav* 2014 Jul;36:487-495. [doi: [10.1016/j.chb.2014.04.005](https://doi.org/10.1016/j.chb.2014.04.005)]
81. Massara F, Ancarani F, Costabile M, Ricotta F. Social desirability in virtual communities. *Int J Bus Adm* 2012 Oct;3(6):93-100. [doi: [10.5430/ijba.v3n6p93](https://doi.org/10.5430/ijba.v3n6p93)]

Abbreviations**DC:** dyskeratosis congenita**TBD:** telomere biology disorder

Edited by T Mackey; submitted 21.02.23; peer-reviewed by M Beckstrand, R Parker, J Chen, L Zheng; comments to author 30.09.23; revised version received 09.11.23; accepted 28.11.23; published 15.01.24.

Please cite as:

Pearce E, Raj H, Emezienna N, Gilkey MB, Lazard AJ, Ribisl KM, Savage SA, Han PKJ

The Use of Social Media to Express and Manage Medical Uncertainty in Dyskeratosis Congenita: Content Analysis

JMIR Infodemiology 2024;4:e46693

URL: <https://infodemiology.jmir.org/2024/1/e46693>

doi: [10.2196/46693](https://doi.org/10.2196/46693)

PMID: [38224480](https://pubmed.ncbi.nlm.nih.gov/38224480/)

©Emily Pearce, Hannah Raj, Ngozika Emezienna, Melissa B Gilkey, Allison J Lazard, Kurt M Ribisl, Sharon A Savage, Paul KJ Han. Originally published in JMIR Infodemiology (<https://infodemiology.jmir.org>), 15.01.2024. 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.

Original Paper

The Role of Social Media in Knowledge, Perceptions, and Self-Reported Adherence Toward COVID-19 Prevention Guidelines: Cross-Sectional Study

Camryn Garrett¹, MPH; Shan Qiao¹, MA, PhD; Xiaoming Li¹, PhD

Department of Health Promotion, Education, and Behavior, Arnold School of Public Health, The University of South Carolina, Columbia, SC, United States

Corresponding Author:

Camryn Garrett, MPH

Department of Health Promotion, Education, and Behavior

Arnold School of Public Health

The University of South Carolina

Discovery Building I

915 Greene Street

Columbia, SC, 29208

United States

Phone: 1 803 777 6844

Fax: 1 803 777 6290

Email: camrynmg@email.sc.edu

Abstract

Background: Throughout the COVID-19 pandemic, social media has served as a channel of communication, a venue for entertainment, and a mechanism for information dissemination.

Objective: This study aims to assess the associations between social media use patterns; demographics; and knowledge, perceptions, and self-reported adherence toward COVID-19 prevention guidelines, due to growing and evolving social media use.

Methods: Quota-sampled data were collected through a web-based survey of US adults through the Qualtrics platform, from March 15, 2022, to March 23, 2022, to assess covariates (eg, demographics, vaccination, and political affiliation), frequency of social media use, social media sources of COVID-19 information, as well as knowledge, perceptions, and self-reported adherence toward COVID-19 prevention guidelines. Three linear regression models were used for data analysis.

Results: A total of 1043 participants responded to the survey, with an average age of 45.3 years, among which 49.61% (n=515) of participants were men, 66.79% (n=696) were White, 11.61% (n=121) were Black or African American, 13.15% (n=137) were Hispanic or Latino, 37.71% (n=382) were Democrat, 30.21% (n=306) were Republican, and 25% (n=260) were not vaccinated. After controlling for covariates, users of TikTok ($\beta=-.29$, 95% CI -0.58 to -0.004 ; $P=.047$) were associated with lower knowledge of COVID-19 guidelines, users of Instagram ($\beta=-.40$, 95% CI -0.68 to -0.12 ; $P=.005$) and Twitter ($\beta=-.33$, 95% CI -0.58 to -0.08 ; $P=.01$) were associated with perceiving guidelines as strict, and users of Facebook ($\beta=-.23$, 95% CI -0.42 to -0.043 ; $P=.02$) and TikTok ($\beta=-.25$, 95% CI -0.5 to -0.009 ; $P=.04$) were associated with lower adherence to the guidelines (R^2 0.06-0.23).

Conclusions: These results allude to the complex interactions between online and physical environments. Future interventions should be tailored to subpopulations based on their demographics and social media site use. Efforts to mitigate misinformation and implement digital public health policy must account for the impact of the digital landscape on knowledge, perceptions, and level of adherence toward prevention guidelines for effective pandemic control.

(*JMIR Infodemiology* 2024;4:e44395) doi:[10.2196/44395](https://doi.org/10.2196/44395)

KEYWORDS

COVID-19; digital media; social media; TikTok; Instagram; Twitter; Facebook; prevention guidelines

Introduction

In March 2020, the infectious disease SARS-CoV-2, more commonly known as COVID-19, was classified as a pandemic [1,2]. As the virus is transmitted through the respiratory systems of individuals in close contact, preventative measures include wearing a facial mask, social distancing, and receiving recommended COVID-19 vaccinations [3]. Over the course of the pandemic, prevention recommendations changed in response to emerging scientific evidence. Initially, a 14-day quarantine and isolation were recommended, which was then shortened to 10 days, and was once more shortened to 5 days [3]. As of March 2022, masks were still recommended in indoor spaces, COVID-19 vaccinations and boosters were widely available, and rapid self-testing was advised in response to exposure or symptom onset [3]. In the United States, as of November 2, 2022, there have been over 97 million confirmed cases and over 1 million total deaths due to COVID-19 [4]. Despite these prevention recommendations, case numbers continued to rise, necessitating research into prevention efforts.

In response to social distancing recommendations, many aspects of life shifted from physical to online environments. Adapting to this change, most US adults (ie, 90%) indicated that digital media was either essential or important for them throughout the pandemic [5]. Digital media encapsulates social media as the platforms that enable human connection in the online environment, with varying degrees of privacy [6]. On social media, individuals encounter and consume information, government announcements, and reactions from other users as they work, learn, connect, and are entertained online [7]. Popular social media sites include Facebook, Twitter, Instagram, Snapchat, TikTok, Pinterest, Reddit, and LinkedIn, among others. As of 2021, a total of 72% of adults in the United States report using at least 1 social media site, representing a 3% increase since 2018 [8]. When stratified by age, 84% of US adults aged 18-29 years indicate using at least 1 social media site [8]. Of those who use Facebook, Snapchat, and Instagram, a majority indicate visiting the platform at least once a day [9]. In considering news consumption on social media, when stratified by age, 42% of users aged 18-29 years indicate social media as their primary source of news [9].

With an increasing proportion of individuals active on social media, thereby encountering COVID-19 news and information online, there are concerns about information accuracy, where unsourced or false information that is widely distributed threatens the dissemination of scientifically accurate information [7,10]. The modalities of social media (eg, concise, organized content formats, and sharing capabilities) allow information to quickly trend as a result of high engagement. The visibility of trending content on social media is determined by engagement and is often based on sensationalism rather than factual accuracy [7]. Sensational misinformation risks reducing the visibility and reach of reputable information [7]. Due to the saturation of misinformation online, the United States is understood to be in a syndemic, denoting the interactions between the COVID-19 pandemic and the infodemic. Social media, therefore, has the capacity to serve both as a tool and a hindrance to health communication.

Despite motivations for use, social media users are subject to unintentionally overconsuming content related to COVID-19 due to the saturation of pandemic information online. Social media has been preliminarily found to negatively contribute to COVID-19 prevention guideline adherence [11]. Among US adults, 53.3% indicate that the amount of information on COVID-19 is overwhelming to the effect that 54.7% indicate that it has led to their avoidance of consuming information about COVID-19 [12]. Resembling emerging trends in the United States, a study in Turkey indicated that 34.4% of respondents follow COVID-19 guidelines less in the present than at the beginning of the pandemic [13]. Fluctuations in pandemic prevention perceptions and adherence over time can be expected, but negative trends, regardless of their cause, necessitate investigation and intervention to bolster commitment to prevention guidelines to limit further pandemic-related exposures [13]. Although a complicated mechanism with additionally probable explanations (eg, milder virus mutations, vaccination availability, mental health burdens, and pandemic fatigue), these downward patterns of adherence are thought to be partially explained by social media use (eg, misinformation and overconsumption). The effective dissemination of scientific, evidence-based health communication must be prioritized in stark opposition to skepticism and disbelief, as sustained by misinformation.

There exists a limited understanding of the associations between demographics and frequency of social media site use and engagement with pandemic prevention behaviors, despite the significant risks to public health. Therefore, there is a present and pressing need to address the field's limited understanding of pandemic-related knowledge, perceptions, and adherence, as impacted online, to design effective health behavior and communication interventions. As the emerging literature demonstrates that content consumption impacts perceptions and, subsequently, health behaviors, the field of health communication must understand the compounding effects of the online environment on COVID-19 prevention efforts [7]. This study therefore aims to investigate the associations between the social media platforms from which individuals consume pandemic-related information as well as their frequency of use and their knowledge of, perceptions of, and adherence to COVID-19 prevention guidelines.

Methods

Survey Development and Data Collection

Preliminary development of the survey involved compiling constructs related to the topics of interest. Survey items were then drafted to measure participant knowledge, perceptions, and adherence toward COVID-19 prevention guidelines. The items were then reviewed by an expert to evaluate and ensure readability, applicability, and response options. The data were obtained using a web-based survey fielded using Qualtrics paid, opt-in distribution services. The data were collected from March 15, 2022, to March 23, 2022.

Ethical Considerations

The University of South Carolina's Institutional Review Board exempted the study (Pro00119512) from Human Research

Subject Regulations based on its minimal risk to participants in providing web-based survey responses. Informed consent was obtained from all participants prior to survey completion. All participants were compensated for their time and efforts in completing the survey (ie, US \$6).

Sample

All adults in the United States were eligible for participation, given that they were 18 years or older at the time of survey response. Responses that were deemed low quality based on response speed, lack of variability in selection, or repetitive attempts were removed before analysis to ensure data quality. Qualtrics used quota sampling methods to ensure the collection of a sample proportionate to that of the United States by way of gender, age, income, race, ethnicity, and education level. The final sample size included 1043 viable responses.

Measures

Demographics

Participant demographics collected included age, gender identity, race or ethnicity, education, employment, income, political affiliation, and COVID-19 vaccination status. Due to limited representation, the American Indian or Alaska Native and Native Hawaiian or Pacific Islander categories were collapsed into 1 category. Age, education, employment, and income were used as continuous variables in the regression models. Gender identity, race or ethnicity, political affiliation, and COVID-19 vaccination status were used as categorical variables in the regression models.

Frequency of Social Media Use

Participants' frequency of any social media use was measured through the item: "About how often do you use social media sites?" Response options ranged from several times a day, once per day, a few times per week, once per week, less than once per week, to never.

Social Media Sources of COVID-19 Information

Participants were asked to check all that apply to the question, "Which of these social media sites have you used to get information about COVID-19?" with the possible response options of Facebook, Twitter, Instagram, Snapchat, Pinterest, TikTok, Reddit, LinkedIn, and another social media site. The social media sites available as response options were chosen due to their popularity and presentation of short-form, user-generated content. Although there exist additional social media platforms (eg, YouTube), those chosen to be included here have active engagement and content sharing capabilities. Demographic profiles of the included social media sites were not accounted for in participant sampling procedures, as it is assumed that user bases may have fluctuated during the pandemic. The selections of these sites were operationalized as categorical predictors in the regression models.

Knowledge of COVID-19 Guidelines

Set forth by the Centers for Disease Control and Prevention, as of March 2022, relevant COVID-19 guidelines were used in crafting 4 items to assess participant pandemic-related knowledge. The assessment evaluated respondents' knowledge

of calculating exposure date, the minimum length of isolation after an exposure or positive test, the percentage of alcohol in hand sanitizer required to kill COVID-19, and what a negative rapid test result indicates. Participants were asked to indicate what they believe the current, official recommendations to be, at the time of survey administration, rather than what they may prefer them to be. These 4 items were then compiled for a final score out of 100%. Knowledge scores of the COVID-19 prevention guidelines were used continuously in the regression models.

Perceptions of COVID-19 Guidelines

Participants were asked to indicate the degree to which they perceived COVID-19 prevention guidelines to be relaxed or strict. The terminology "strict" was operationalized through concurrent dimensions that encapsulate participant responses to legal and scientific guidelines as well as enforcement. As perceptions of COVID-19 guidelines were assessed after the knowledge assessment, the guidelines were not explicitly defined but rather assumed to encapsulate mask-wearing, gathering size limitations, hygiene measures, as well as quarantine and isolation timelines. This ordering provided participants with context as to what the term "guidelines" referred to. Participants were asked: "Do you consider the current COVID-19 guidelines as:" with the response options ranging from too strict, a little too strict, about right, a little too relaxed, to too relaxed.

Adherence to COVID-19 Guidelines

Adherence to COVID-19 guidelines was evaluated by asking participants if they generally follow the official COVID-19 prevention guidelines, with the available response options of strongly, sometimes, rarely, and never follow the guidelines. This item provided an average, typical measure of self-reported participant adherence to COVID-19 guidelines, broadly. Given the state of the pandemic, this item was reliant upon participant understanding of guidelines in the organizations and institutions to which they belong (ie, schools and workplaces).

Statistical Analysis

All statistical analyses were conducted using the statistical analysis software, SAS (version 9.4; SAS Institute). Descriptive analyses were conducted for key predictors. All data were screened for outliers, missing data, and normality. As all data used in this study was collected through discrete response options, excluding age, their distributions were considered to assess the presence of outliers. This was done by considering the frequency of responses within available options through histograms and box plots, as applicable. Those categories that were lower in response volume were collapsed (eg, race or ethnicity response of American Indian or Alaska Native and Native Hawaiian or Pacific Islander) or excluded from the analysis before modeling (eg, gender identity response option of nonbinary). Data quality was ensured as Qualtrics excluded participants who did not complete the survey in a single session, who were not continuously and carefully responding, who missed embedded attention checks, or who completed the survey in less than a third or more than 3 times the median time it took other participants to complete the survey. Due to the use of

these features, respondents who did not complete the survey were not tracked. No systematic patterns of missing data within the data collected, or between variables, were observed. There is limited item nonresponse. Bivariate associations were assessed through ANOVA and Pearson correlation tests, as appropriate. Three generalized linear regressions, using a maximum likelihood estimation procedure, were conducted, independently, to explore associations between social media use and demographics and knowledge, perceptions, and self-reported adherence toward prevention guidelines, respectively. Although the 3 outcomes of knowledge, perceptions, and self-reported adherence were run independently, their theoretically dependent nature led us to consider implementing a correction (ie, Bonferroni), but as it resulted in a minimal impact on our findings, the traditional α level of .05 was here used to evaluate our findings.

Results

Overview

Of the 1043 participants, the median age of participants was 45.3 years (Table 1). The distribution of the gender identity of the participants was split approximately equally between men (515/1032, 49.9%) and women (513/1032, 49.71%), with few participants indicating being nonbinary or transgender. The race or ethnicity of participants was primarily White (696/1042, 66.79%), followed by Latino or Hispanic (137/1042, 13.15%) and Black or African American (121/1042, 11.61%). A quarter (253/1042, 24.28%) of participants held a bachelor's degree and approximately a quarter (269/1042, 25.82%) of participants indicated earning US \$50,000-US \$79,999 annually. Finally, almost half (498/1040, 47.88%) of the participants had received a full vaccination series and booster against COVID-19.

Table 1. Demographic characteristics of study participants (N=1043).

Variables	Values, n (%)
Age (years; 1 participant's data are missing), mean (SD)	45.3 (16.94)
Gender (11 participants' data are missing)	
Men	515 (49.9)
Women	513 (49.71)
Nonbinary or other	4 (0.39)
Race or ethnicity (check all that apply; 1 participant's data are missing)	
Black or African American	121 (11.61)
Latino or Hispanic	137 (13.15)
American Indian or Alaska Native and Native Hawaiian or Pacific Islander	22 (2.11)
White	696 (66.79)
Other	66 (6.33)
Education (1 participant's data are missing)	
Less than high school degree	25 (2.4)
High school graduate or equivalent	248 (23.8)
Some college but no degree	248 (23.8)
Associate degree	123 (11.8)
Bachelor's degree	253 (24.28)
Master's degree	112 (10.75)
Doctoral or professional degree (JD, MD, or PhD)	33 (3.17)
Employment status over the last 3 months (6 participant's data are missing)	
Working full-time	499 (48.12)
Working part-time	132 (12.73)
Unemployed and looking for work	74 (7.14)
Homemaker or stay-at-home parent	70 (6.75)
Student	35 (3.38)
Retired	200 (19.29)
Other	27 (2.6)
Previous year income (US \$; 1 participant's data are missing)	
Less than 10,000	56 (5.37)
10,000-19,999	58 (5.57)
20,000-29,999	96 (9.21)
30,000-39,999	87 (8.35)
40,000-49,999	70 (6.72)
50,000-59,000	117 (11.23)
60,000-69,999	70 (6.72)
70,000-79,999	82 (7.87)
80,000-89,999	47 (4.51)
90,000-99,999	51 (4.89)
100,000-149,999	215 (20.63)
150,000 or more	93 (8.93)
Political affiliation (30 participants' data are missing)	
Republican	306 (30.21)

Variables	Values, n (%)
Democrat	382 (37.71)
Independent	325 (32.08)
COVID-19 vaccination status (3 participant's data are missing)	
No	260 (25)
Yes, but no booster	282 (27.12)
Yes, including booster	498 (47.88)

Social Media Site Use

Participants reported using, generally or for any reason, the social media sites Facebook (835/1042, 80.13%), Twitter (396/1042, 38%), Instagram (586/1042, 56.24%), Snapchat (329/1042, 31.57%), Pinterest (320/1042, 30.71%), TikTok (401/1042, 38.48%), Reddit (208/1042, 19.96%), LinkedIn (254/1042, 24.38%), or another social media site (69/1042, 6.62%). Further, participants reported accessing COVID-19 information using the social media sites Facebook (604/1042, 57.97%), Twitter (220/1042, 21.11%), Instagram (258/1042, 24.76%), Snapchat (85/1042, 8.16%), Pinterest (59/1042, 5.66%), TikTok (129/1042, 12.38%), Reddit (84/1042, 8.06%),

LinkedIn (72/1042, 6.91%), and another social media site (42/1042, 4.03%).

Table 2 presents the results of the bivariate analyses. Pearson correlations suggest that the demographic variables of age, education, and income were correlated with the prevention mitigation outcomes of guideline knowledge, perceptions, and self-reported adherence. The ANOVA suggests that political affiliation was correlated with all 3 outcomes while gender, race or ethnicity, and COVID-19 vaccination status were correlated with prevention guideline perceptions and self-reported adherence. Social media sites used to consume COVID-19 news were correlated with self-reported adherence. Employment and regularity of social media use were not correlated with the outcomes of interest.

Table 2. Bivariate analysis results.

Variable	Outcomes		
	Knowledge	Perceptions	Self-reported adherence
Age			
<i>r</i>	0.09	-0.14	0.08
<i>P</i> value	.006	<.001	.01
Education			
<i>r</i>	0.11	0.001	0.11
<i>P</i> value	<.001	.97	<.001
Employment			
<i>r</i>	0.02	-0.03	0.04
<i>P</i> value	.48	.27	.21
Income			
<i>r</i>	0.15	-0.08	0.04
<i>P</i> value	<.001	.007	.17
Gender			
ANOVA (<i>F</i>)	0.38	6.43	5.27
<i>P</i> value	.54	.01	.02
Race or ethnicity			
ANOVA (<i>F</i>)	2.36	12.66	3.85
<i>P</i> value	.051	<.001	.004
Political affiliation			
ANOVA (<i>F</i>)	6.23	94.13	49.87
<i>P</i> value	.002	<.001	<.001
COVID-19 vaccination status			
ANOVA (<i>F</i>)	2.7	23.88	69.85
<i>P</i> value	.07	<.001	<.001
Site for COVID-19 news			
ANOVA (<i>F</i>)	2.07	1.64	2.89
<i>P</i> value	.07	.15	.01
Regularity of social media use			
ANOVA (<i>F</i>)	0.53	1.21	1.23
<i>P</i> value	.75	.30	.29

Knowledge of COVID-19 Guidelines

Indicating the level of knowledge related to COVID-19 prevention guidelines, the possible scores participants could receive included 100% (n=14, 1.4%), 75% (n=112, 10.9%), 50% (n=429, 41.7%), 25% (n=368, 35.7%), or 0% (n=107, 10.4%) correct. Model 1 (Table 3) suggests that income, Democratic political affiliation, and use of the social media platform TikTok were associated with COVID-19 prevention guideline knowledge. Specifically, as income ($\beta=.03$, 95% CI

0.005-0.05; $P=.02$) increased, it was found to be associated with a higher level of knowledge of COVID-19 guidelines. Democratic political affiliation ($\beta=-.21$, 95% CI -0.37 to -0.057; $P=.008$) was found to be negatively associated with guideline knowledge. Using TikTok as a source of COVID-19 information ($\beta=-.29$, 95% CI -0.58 to -0.004; $P=.047$) was associated with a lower level of knowledge. This model explained 6% of the variance in knowledge of COVID-19 guidelines.

Table 3. Regression results for knowledge, perceptions, and self-reported adherence.

Independent variables (reference)	Model 1: knowledge ^a		Model 2: perceptions ^a		Model 3: self-reported adherence ^a	
	β (95% CI)	<i>P</i> value	β (95% CI)	<i>P</i> value	β (95% CI)	<i>P</i> value
Age	.002 (-0.003 to 0.007)	.51	-.007 (-0.01 to -0.002)	.007 ^b	.004 (-0.0004 to 0.008)	.07
Gender (men)						
Women	.026 (-0.1 to 0.15)	.69	.16 (0.02 to 0.3)	.02 ^b	.15 (0.04 to 0.26)	.008 ^b
Race or ethnicity (White)						
Black or African American	-.048 (-0.26 to 0.16)	.66	.14 (-0.09 to 0.36)	.24	.21 (0.03 to 0.39)	.02 ^b
Hispanic or Latino	-.079 (-0.27 to 0.11)	.41	.28 (0.08 to 0.49)	.007 ^b	.27 (0.11 to 0.43)	.001 ^b
American Indian or Alaska Native and Native Hawaiian or Pacific Islander	.072 (-0.45 to 0.6)	.79	.92 (0.35 to 1.49)	.002 ^b	.32 (-0.13 to 0.77)	.16
Other	-.09 (-0.35 to 0.17)	.49	.07 (-0.21 to 0.34)	.63	.14 (-0.07 to 0.36)	.19
Education level	.03 (-0.014 to 0.079)	.17	-.015 (-0.065 to 0.036)	.56	.001 (-0.039 to 0.04)	.95
Employment	.029 (-0.008 to 0.066)	.13	-.001 (-0.04 to 0.038)	.95	.005 (-0.027 to 0.036)	.78
Income	.03 (0.005 to 0.05)	.02 ^b	-.03 (-0.053 to -0.005)	.02 ^b	-.015 (-0.03 to 0.004)	.13
Political affiliation (independent)						
Republican	-.12 (-0.28 to 0.04)	.15	-.5 (-0.67 to -0.33)	<.001 ^b	-.23 (-0.37 to -0.09)	.001 ^b
Democrat	-.21 (-0.37 to -0.057)	.008 ^b	.34 (0.17 to 0.5)	<.001 ^b	.17 (-0.04 to 0.31)	.01 ^b
COVID-19 vaccination status (yes, but no booster)						
No	.00 (-0.17 to 0.17)	.99	-.22 (-0.4 to -0.04)	.02 ^b	-.22 (-0.36 to -0.07)	.003 ^b
Yes, including booster	.02 (-0.13 to 0.18)	.78	.31 (0.15 to 0.48)	<.001 ^b	.32 (0.19 to 0.45)	<.001 ^b
Site for COVID-19 news (Reddit)						
Facebook	-.086 (-0.31 to 0.14)	.45	-.23 (-0.47 to 0.009)	.06	-.23 (-0.42 to -0.043)	.02 ^b
Instagram	-.026 (-0.28 to 0.23)	.84	-.40 (-0.68 to -0.12)	.005 ^b	-.22 (-0.44 to 0.0026)	.05
Snapchat	.21 (-0.26 to 0.68)	.38	-.17 (-0.66 to 0.31)	.49	-.33 (-0.71 to 0.057)	.10
TikTok	-.29 (-0.58 to -0.004)	.047 ^b	-.29 (-0.6 to 0.016)	.06	-.25 (-0.5 to -0.009)	.04 ^b
Twitter	.015 (-0.22 to 0.25)	.90	-.33 (-0.58 to -0.08)	.01 ^b	-.08 (-0.28 to 0.12)	.43
Regularity of social media use (less than once per week)						
Several times per day	.27 (-0.18 to 0.71)	.24	-.22 (-0.71 to 0.27)	.37	-.24 (-0.62 to 0.14)	.22
Once per day	.16 (-0.32 to 0.63)	.52	-.12 (-0.63 to 0.4)	.66	-.23 (-0.63 to 0.18)	.27
A few times per week	.2 (-0.29 to 0.69)	.43	-.03 (-0.57 to 0.5)	.91	-.16 (-0.58 to 0.27)	.47

Independent variables (reference)	Model 1: knowledge ^a		Model 2: perceptions ^a		Model 3: self-reported adherence ^a	
	β (95% CI)	<i>P</i> value	β (95% CI)	<i>P</i> value	β (95% CI)	<i>P</i> value
Once per week	-.55 (-1.33 to 0.22)	.16	-.23 (-1.07 to 0.62)	.60	-.03 (-0.7 to 0.63)	.92

^a R^2 values of models 1-3 are 0.06 (knowledge), 0.23 (perceptions), and 0.19 (self-reported adherence) respectively.

^b*P* values indicate statistical significance at the $\alpha=.05$ level.

Perceptions of COVID-19 Guidelines

Model 2 (Table 3) suggests that age, gender, Hispanic or Latino populations, American Indian or Alaska Native populations, income, political affiliation, COVID-19 vaccination status, and the use of the social media sites Instagram and Twitter were associated with perceptions of COVID-19 prevention guidelines. As age ($\beta=-.007$, 95% CI -0.01 to -0.002 ; $P=.007$) increased, it was found to be associated with a perception of the guidelines as strict. Women ($\beta=.16$, 95% CI 0.02 - 0.3 ; $P=.02$) were associated with perceiving the guidelines as relaxed. Hispanic or Latino ($\beta=.28$, 95% CI 0.08 - 0.49 ; $P=.007$) and American Indian or Alaska Native and Native Hawaiian or Pacific Islander ($\beta=.92$, 95% CI 0.35 - 1.49 ; $P=.002$) populations were found to be associated with perceiving the guidelines as relaxed. As income ($\beta=-.03$, 95% CI -0.05 to -0.005 ; $P=.02$) increases, it was found to be associated with stricter perceptions of the guidelines. Republican political affiliation ($\beta=-.5$, 95% CI -0.67 to -0.33 ; $P<.001$) was found to be associated with perceiving the guidelines as strict, while Democratic political affiliation ($\beta=.34$, 95% CI 0.17 - 0.5 ; $P<.001$) was found to be associated with perceiving them as relaxed. Receiving the full vaccination series and booster ($\beta=.31$, 95% CI 0.15 - 0.48 ; $P<.001$) was found to be associated with perceiving the guidelines as relaxed, while receiving no COVID-19 vaccinations ($\beta=-.22$, 95% CI -0.4 to -0.04 ; $P=.02$) was associated with perceiving them as strict. Instagram ($\beta=-.4$, 95% CI -0.68 to -0.12 ; $P=.005$) and Twitter ($\beta=-.33$, 95% CI -0.58 to -0.08 ; $P=.01$) were found to be associated with stricter perceptions of the COVID-19 prevention guidelines. This model explained 23% of the variance in perceptions of COVID-19 guidelines.

Adherence to COVID-19 Guidelines

As related to self-reported COVID-19 guideline adherence, model 3 (Table 3) suggests that women, Black or African American populations, Hispanic or Latino populations, political affiliation, COVID-19 vaccination status, and the use of Facebook and TikTok were associated with adherence to the COVID-19 prevention guidelines. Women ($\beta=.15$, 95% CI 0.04 - 0.26 ; $P=.008$) were found to be positively associated with adherence to the COVID-19 prevention guidelines. Black or African American ($\beta=.21$, 95% CI 0.03 - 0.39 ; $P=.02$) and Hispanic or Latino ($\beta=.27$, 95% CI 0.11 - 0.43 ; $P=.001$) populations were found to be positively associated with adherence to the guidelines. Republican political affiliation ($\beta=-.23$, 95% CI -0.37 to -0.09 ; $P=.001$) was negatively associated with adherence to prevention guidelines, while Democratic political affiliation ($\beta=.17$, 95% CI -0.04 to 0.31 ; $P=.01$) was positively associated with adherence. Receiving the full vaccination series and booster ($\beta=.32$, 95% CI 0.19 - 0.45 ;

$P<.001$) was positively associated with adherence to the COVID-19 prevention guidelines, while receiving no COVID-19 vaccinations ($\beta=-.22$, 95% CI -0.36 to -0.07 ; $P=.003$) was negatively associated with adherence. Facebook ($\beta=-.23$, 95% CI -0.42 to -0.043 ; $P=.02$) and TikTok ($\beta=-.25$, 95% CI -0.5 to -0.009 ; $P=.04$) were found to be negatively associated with self-reported adherence to COVID-19 prevention guidelines. This model explained 19% of the variance in adherence to COVID-19 guidelines.

Discussion

Principal Findings

This study suggests that knowledge, perceptions, and self-reported adherence toward COVID-19 prevention guidelines differ by demographics and social media site use. Notably, marginalized populations (eg, older adults, women, and racial or ethnic minority individuals) were found to perceive the COVID-19 prevention guidelines as relaxed, in addition to their positive association with adherence. Political affiliation and COVID-19 vaccination status mirror assumptions about perceptions and adherence, where those identifying as Republican and reporting no vaccination were associated with perceiving the guidelines as too strict and adhering to a lesser degree, respectively. The popular social media sites TikTok, Instagram, Facebook, and Twitter were found to negatively impact pandemic prevention efforts as they were differentially associated with lower levels of knowledge, perceiving guidelines as strict, and lower self-reported adherence. The findings of this work, while demonstrating complicated interactions between guideline knowledge, perceptions, and adherence, serve to inform tailored public health interventions (ie, on the basis of demographic subgroups and social media site use), platform policies (eg, misinformation prevention), and digital public health policy more broadly.

Demographics and Knowledge, Perceptions, and Adherence Toward Guidelines

When considering the associations between the demographic correlates of income, age, and gender with knowledge, perceptions, and adherence toward prevention guidelines, the findings suggest a complex pandemic landscape. Whereas education and employment were not associated with guideline knowledge, it can be assumed that income reflects a layer of privilege afforded to those of higher income throughout the pandemic. In the case of this study, income may be acting as a proxy for pandemic privilege rather than solely socioeconomic status. Pandemic privilege can be understood here as the role of income in altering the pandemic environment, where those with additional resources are more likely to have access to

prevention methods (eg, working from home, personal protective equipment, vaccination appointment flexibility, transportation, residential privilege, limited disruptions to services and care, and financial buffer for burdens of lost employment and wages) [14,15]. Despite possessing increased knowledge of the guidelines, perceptions of the prevention guidelines as strict reflect privileged protections afforded through increased income. Concordant with the existing literature, among older adults, a higher level of adherence to prevention guidelines, despite perceptions of them as strict, is likely due to the higher risk of severe illness from COVID-19 associated with increased age [16,17]. Gendered differences in perceptions of the guidelines as relaxed with a higher level of adherence reflect disproportionate pandemic burdens experienced by women (eg, occupational exposure, incidence, and post-COVID-19 condition [long COVID]).

The present findings are in accordance with the existing literature that demonstrates the impact of political affiliation on knowledge, perceptions, and adherence toward prevention guidelines. Partisan differences in perceptions of COVID-19 guidelines have been theorized to be explained by differential risk perceptions as influenced by news sources and media consumption [18-21]. Republican political affiliation has been found to be aligned with a preference for reducing the imposition of guidelines, while Democratic political affiliation is aligned with a preference for maintaining guidelines [22]. In accordance with the literature, political affiliation may play a decisive role in impacting knowledge-seeking and comprehension, perceptions, and adherence toward prevention guidelines. Health communication efforts may bolster prevention efforts through the characteristics inherent to partisan politics (eg, collectivism, inequity perceptions, perceived risk, skepticism, and media influence) and their influence on health behaviors [22-24]. The emerging literature attests that although political affiliation may demonstrate explanatory differences in pandemic prevention outcomes, there is a call for public health efforts that extend beyond interventions targeted based on political affiliation, implementing bipartisan efforts that also further consider demographics and individual differences influencing the operationalization of information from news and social media sites in the interest of COVID-19 prevention [18,23].

Social Media Sites and Knowledge, Perceptions, and Adherence Toward Guidelines

The use of the social media sites TikTok, Instagram, Twitter, and Facebook was found to be associated with lower knowledge, stricter perceptions, and lesser adherence toward COVID-19 prevention guidelines. Despite operating under distinct algorithms, all 4 platforms share commonalities in their functions for photo, video, audio, and text sharing, as well as social networking structures. A reliance on user-generated content creates difficulty in regulating the presence and spread of misinformation on social media. All 4 sites implemented, to various degrees, efforts to mitigate misinformation through informational banners on videos discussing the pandemic with off-site links to additional information. Despite these soft moderation efforts to address misinformation by TikTok, Instagram, Twitter, and Facebook, all have been found to contribute to the dissemination of misinformation [25-28].

Therefore, there is a need for improved mechanisms on these social media sites to limit the spread of misinformation due to its negative impacts on COVID-19 prevention guideline knowledge, perceptions, and adherence in the physical environment.

One key consideration of this study is the discrepancy between the demographic profiles of the included social media sites and the study sample. The user base of TikTok (ie, 48% users aged 18-29 years, 22% users aged 30-49 years, 14% users aged 50-64 years, and 4% users aged 65 years and older), Twitter (ie, 42% users aged 18-29 years, 27% users aged 30-49 years, 18% users aged 50-64 years, and 7% users aged 65 years and older), and Instagram (ie, 71% users aged 18-29 years, 48% users aged 30-49 years, 29% users aged 50-64 years, and 13% users aged 65 years and older) tends to be younger than that of Facebook (ie, 70% users aged 18-29 years, 77% users aged 30-49 years, 73% users aged 50-64 years, and 50% users aged 65 years and older) [8]. Although the average age of the study sample is older, it aligns with profiles of users of a similar age range who are active online (ie, 22% users on TikTok, 27% users on Twitter, 48% users on Instagram, and 77% users on Facebook) [8]. Although social media sites have unique demographic user profiles, it is necessary to consider that all individuals are able to access their platforms. Understanding the scope of a platform's typical and atypical users is necessary to systematically address misinformation online, where those who do not align with the average user experience an assumedly differential interaction with the platform and its content.

Public Health Implications

This research is uniquely situated within the COVID-19 pandemic and serves to inform tailored public health interventions, social media platform strategies, and policies. The key implications of this research include addressing knowledge gaps in the literature regarding the impact of social media use and demographic characteristics on COVID-19 prevention guideline knowledge, perceptions, and adherence. Public health interventions should be tailored to relevant platforms to address the impacts of social media sites on prevention guideline knowledge, perceptions, and adherence. Additionally, interventions targeting demographic subgroups may be operationalized on social media platforms with a user base that aligns with the target subgroup (eg, age, income, and political affiliation). In this context, platform functionality should be considered when designing interventions, regulations, and misinformation mitigation policies to alleviate the negative impacts of social media use on COVID-19 prevention efforts. Finally, these findings are necessary to be operationalized within public health interventions to tailor interventions to increase pandemic-related knowledge while enhancing supportive perceptions of the guidelines, aiming to increase and maintain sufficient adherence among subpopulations to mitigate the effects of the pandemic.

Strengths, Limitations, and Future Studies

This study has the strengths of using a country-wide, quota-based sample to investigate emerging trends during the pandemic as related to knowledge of, perceptions of, and adherence to COVID-19 prevention guidelines. Although there

is likely some inherent difference in those who are online and able to participate in the survey as compared with those who are not, this concern may be mitigated in the context of this work, as it centers those active in the online environment. With the goal of identifying the role of social media on the target population, the exclusion of those not online is warranted. The findings should be cautiously interpreted and generalized as selection bias may affect the representativeness of the sample. When interpreting the study's findings, low statistical significance does not imply the absence of a certain phenomenon. One limitation that could persist, as the results are reliant on a self-report measure of prevention guidelines adherence, is participants' ability to approximate habits (eg, wearing a mask and using a social media site). A key limitation of this study is the discrepancy between the demographics of the study sample and the demographic profiles of the users of the various social media sites included. Finally, as a

cross-sectional study, where some potential but key confounders may not have been included, there is the inability to obtain causal inference. Further, work accounting for the interrelations between factors should be conducted to provide a comprehensive assessment of confounders [22]. Future work should consider focusing on the validation of measures to assess knowledge, perceptions, and adherence. Additional research would benefit from an expanded survey considering a variety of potential, influential factors (eg, health literacy and location). Longitudinal explorations of the influence of social media use, knowledge levels, and declining perceptions should be prioritized in efforts to examine their impacts on prevention guideline adherence over time. Future directions for health communication should prioritize implementing programmatic interventions on social media platforms to address misinformation and information oversaturation in a manner that optimizes each platform's social networking functions, algorithms, and user base.

Acknowledgments

The authors would like to acknowledge the efforts of Jacob Long from the College of Information and Communication Sciences at the University of South Carolina for his support in survey development and data analysis. This work was funded by the National Institute of Health grants R01AI174892 and R01AI127203-5S1.

Authors' Contributions

CG developed survey materials, analyzed data, and led manuscript development. SQ collaborated on the development and provided significant contributions in manuscript refinement. XL provided significant contributions to manuscript refinement.

Conflicts of Interest

None declared.

References

1. Liu YC, Kuo RL, Shih SR. COVID-19: the first documented coronavirus pandemic in history. *Biomed J* 2020;43(4):328-333 [FREE Full text] [doi: [10.1016/j.bj.2020.04.007](https://doi.org/10.1016/j.bj.2020.04.007)] [Medline: [32387617](https://pubmed.ncbi.nlm.nih.gov/32387617/)]
2. Coronavirus disease (COVID-19). World Health Organization. URL: <https://www.who.int/health-topics/coronavirus> [accessed 2022-11-17]
3. COVID-19: How to protect yourself and others. Centers for Disease Control and Prevention. 2022. URL: <https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/prevention.html> [accessed 2022-04-13]
4. COVID data tracker. Centers for Disease Control and Prevention. 2022. URL: <https://covid.cdc.gov/covid-data-tracker> [accessed 2022-11-17]
5. McClain C, Vogels EA, Perrin A, Sechopoulos S, Rainie L. The internet and the pandemic. Pew Research Center. 2021. URL: <https://www.pewresearch.org/internet/2021/09/01/the-internet-and-the-pandemic/> [accessed 2022-11-17]
6. Aichner T, Grünfelder M, Maurer O, Jegeni D. Twenty-five years of social media: a review of social media applications and definitions from 1994 to 2019. *Cyberpsychol Behav Soc Netw* 2021;24(4):215-222 [FREE Full text] [doi: [10.1089/cyber.2020.0134](https://doi.org/10.1089/cyber.2020.0134)] [Medline: [33847527](https://pubmed.ncbi.nlm.nih.gov/33847527/)]
7. Mheidly N, Fares J. Leveraging media and health communication strategies to overcome the COVID-19 infodemic. *J Public Health Policy* 2020;41(4):410-420 [FREE Full text] [doi: [10.1057/s41271-020-00247-w](https://doi.org/10.1057/s41271-020-00247-w)] [Medline: [32826935](https://pubmed.ncbi.nlm.nih.gov/32826935/)]
8. Social media fact sheet. Pew Research Center. URL: <https://www.pewresearch.org/internet/fact-sheet/social-media/> [accessed 2022-07-25]
9. Auxier A, Anderson M. Social media use in 2021. Pew Research Center. 2021. URL: <https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/> [accessed 2022-10-24]
10. Joseph AM, Fernandez V, Kritzman S, Eaddy I, Cook OM, Lambros S, et al. COVID-19 misinformation on social media: a scoping review. *Cureus* 2022;14(4):e24601 [FREE Full text] [doi: [10.7759/cureus.24601](https://doi.org/10.7759/cureus.24601)] [Medline: [35664409](https://pubmed.ncbi.nlm.nih.gov/35664409/)]
11. Vaal S, Schofield MB, Baker IS, Roberts BLH. Narcissism, national narcissism, COVID-19 conspiracy belief, and social media use as predictors of compliance with COVID-19 public health guidelines. *Curr Psychol* 2022;1-8 [FREE Full text] [doi: [10.1007/s12144-022-03715-6](https://doi.org/10.1007/s12144-022-03715-6)] [Medline: [36213568](https://pubmed.ncbi.nlm.nih.gov/36213568/)]

12. Volkmer I. Social media and COVID-19: a global study of digital crisis interaction among Gen Z and millennials. University of Melbourne. 2021. URL: <https://minerva-access.unimelb.edu.au/items/db07228d-5b4c-59df-bf3e-e6bb37072aa9> [accessed 2024-01-10]
13. Haktanir A, Can N, Seki T, Kurnaz MF, Dilmaç B. Do we experience pandemic fatigue? current state, predictors, and prevention. *Curr Psychol* 2022;41(10):7314-7325 [FREE Full text] [doi: [10.1007/s12144-021-02397-w](https://doi.org/10.1007/s12144-021-02397-w)] [Medline: [34690475](https://pubmed.ncbi.nlm.nih.gov/34690475/)]
14. Huang X, Lu J, Gao S, Wang S, Liu Z, Wei H. Staying at home is a privilege: evidence from fine-grained mobile phone location data in the United States during the COVID-19 pandemic. *Ann Am Assoc Geogr* 2021;112(1):286-305. [doi: [10.1080/24694452.2021.1904819](https://doi.org/10.1080/24694452.2021.1904819)]
15. Berkman LF, Kawachi I. *Social Epidemiology*. New York: Oxford University Press; 2000.
16. COVID-19 risks and information for older adults. Centers for Disease Control and Prevention. 2022. URL: <https://www.cdc.gov/aging/covid19/index.html#:~:text=Older%20adults%20are%20more%20likely,very%20sick%20from%20COVID%2D19> [accessed 2022-07-23]
17. Coroiu A, Moran C, Campbell T, Geller AC. Barriers and facilitators of adherence to social distancing recommendations during COVID-19 among a large international sample of adults. *PLoS One* 2020;15(10):e0239795 [FREE Full text] [doi: [10.1371/journal.pone.0239795](https://doi.org/10.1371/journal.pone.0239795)] [Medline: [33027281](https://pubmed.ncbi.nlm.nih.gov/33027281/)]
18. Hsiehchen D, Espinoza M, Slovic P. Political partisanship and mobility restriction during the COVID-19 pandemic. *Public Health* 2020;187:111-114 [FREE Full text] [doi: [10.1016/j.puhe.2020.08.009](https://doi.org/10.1016/j.puhe.2020.08.009)] [Medline: [32947252](https://pubmed.ncbi.nlm.nih.gov/32947252/)]
19. Kiviniemi MT, Orom H, Hay JL, Waters EA. Prevention is political: political party affiliation predicts perceived risk and prevention behaviors for COVID-19. *BMC Public Health* 2022;22(1):298 [FREE Full text] [doi: [10.1186/s12889-022-12649-4](https://doi.org/10.1186/s12889-022-12649-4)] [Medline: [35164719](https://pubmed.ncbi.nlm.nih.gov/35164719/)]
20. Limbu YB, Gautam RK, Pham L. The health belief model applied to COVID-19 vaccine hesitancy: a systematic review. *Vaccines (Basel)* 2022;10(6):973 [FREE Full text] [doi: [10.3390/vaccines10060973](https://doi.org/10.3390/vaccines10060973)] [Medline: [35746581](https://pubmed.ncbi.nlm.nih.gov/35746581/)]
21. Zhou Y, Myrick JG, Farrell EL, Cohen O. Perceived risk, emotions, and stress in response to COVID-19: the interplay of media use and partisanship. *Risk Anal* 2023;43(8):1572-1586 [FREE Full text] [doi: [10.1111/risa.14044](https://doi.org/10.1111/risa.14044)] [Medline: [36307383](https://pubmed.ncbi.nlm.nih.gov/36307383/)]
22. Gerace A, Rigney G, Anderson JR. Predicting attitudes towards easing COVID-19 restrictions in the United States of America: the role of health concerns, demographic, political, and individual difference factors. *PLoS One* 2022;17(2):e0263128 [FREE Full text] [doi: [10.1371/journal.pone.0263128](https://doi.org/10.1371/journal.pone.0263128)] [Medline: [35196316](https://pubmed.ncbi.nlm.nih.gov/35196316/)]
23. Franz B, Dhanani LY. Beyond political affiliation: an examination of the relationships between social factors and perceptions of and responses to COVID-19. *J Behav Med* 2021;44(5):641-652 [FREE Full text] [doi: [10.1007/s10865-021-00226-w](https://doi.org/10.1007/s10865-021-00226-w)] [Medline: [33877532](https://pubmed.ncbi.nlm.nih.gov/33877532/)]
24. Gollust SE, Fowler EF, Vogel RI, Rothman AJ, Yzer M, Nagler RH. Americans' perceptions of health disparities over the first year of the COVID-19 pandemic: results from three nationally-representative surveys. *Prev Med* 2022;162:107135 [FREE Full text] [doi: [10.1016/j.ypmed.2022.107135](https://doi.org/10.1016/j.ypmed.2022.107135)] [Medline: [35803354](https://pubmed.ncbi.nlm.nih.gov/35803354/)]
25. Basch CH, Meleo-Erwin Z, Fera J, Jaime C, Basch CE. A global pandemic in the time of viral memes: COVID-19 vaccine misinformation and disinformation on TikTok. *Hum Vaccin Immunother* 2021;17(8):2373-2377 [FREE Full text] [doi: [10.1080/21645515.2021.1894896](https://doi.org/10.1080/21645515.2021.1894896)] [Medline: [33764283](https://pubmed.ncbi.nlm.nih.gov/33764283/)]
26. Shelton CC, Curcio R, Carpenter JP, Schroeder SE. Instagramming for justice: the potentials and pitfalls of culturally relevant professional learning on Instagram. *TechTrends* 2022;66(5):837-854 [FREE Full text] [doi: [10.1007/s11528-022-00758-1](https://doi.org/10.1007/s11528-022-00758-1)] [Medline: [35818416](https://pubmed.ncbi.nlm.nih.gov/35818416/)]
27. Sharevski F, Alsaadi R, Jachim P, Pieroni E. Misinformation warnings: Twitter's soft moderation effects on COVID-19 vaccine belief echoes. *Comput Secur* 2022;114:102577 [FREE Full text] [doi: [10.1016/j.cose.2021.102577](https://doi.org/10.1016/j.cose.2021.102577)] [Medline: [34934255](https://pubmed.ncbi.nlm.nih.gov/34934255/)]
28. Broniatowski DA, Simons JR, Gu J, Jamison AM, Abrams LC. The efficacy of Facebook's vaccine misinformation policies and architecture during the COVID-19 pandemic. *Sci Adv* 2023;9(37):eadh2132 [FREE Full text] [doi: [10.1126/sciadv.adh2132](https://doi.org/10.1126/sciadv.adh2132)] [Medline: [37713497](https://pubmed.ncbi.nlm.nih.gov/37713497/)]

Edited by R Cuomo; submitted 17.11.22; peer-reviewed by A Zheluk, A Rovetta; comments to author 02.05.23; revised version received 26.09.23; accepted 07.01.24; published 16.02.24.

Please cite as:

Garrett C, Qiao S, Li X

The Role of Social Media in Knowledge, Perceptions, and Self-Reported Adherence Toward COVID-19 Prevention Guidelines: Cross-Sectional Study

JMIR Infodemiology 2024;4:e44395

URL: <https://infodemiology.jmir.org/2024/1/e44395>

doi: [10.2196/44395](https://doi.org/10.2196/44395)

PMID: [38194493](https://pubmed.ncbi.nlm.nih.gov/38194493/)

©Camryn Garrett, Shan Qiao, Xiaoming Li. Originally published in JMIR Infodemiology (<https://infodemiology.jmir.org>), 16.02.2024. 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.

Original Paper

Government-Nongovernmental Organization (NGO) Collaboration in Macao's COVID-19 Vaccine Promotion: Social Media Case Study

Xuechang Xian^{1,2*}, MA; Rostam J Neuwirth³, PhD; Angela Chang^{2*}, PhD

¹Department of Publicity, Zhaoqing University, Zhaoqing, China

²Department of Communication, University of Macau, Macao SAR, China

³Department of Global Legal Studies, University of Macau, Macao SAR, China

*these authors contributed equally

Corresponding Author:

Angela Chang, PhD

Department of Communication

University of Macau

Avenida da Universidade

Taipa

Macao SAR, 999078

China

Phone: 86 88228991

Email: wychang@um.edu.mo

Abstract

Background: The COVID-19 pandemic triggered unprecedented global vaccination efforts, with social media being a popular tool for vaccine promotion.

Objective: This study probes into Macao's COVID-19 vaccine communication dynamics, with a focus on the multifaceted impacts of government agendas on social media.

Methods: We scrutinized 22,986 vaccine-related Facebook posts from January 2020 to August 2022 in Macao. Using automated content analysis and advanced statistical methods, we unveiled intricate agenda dynamics between government and nongovernment entities.

Results: "Vaccine importance" and "COVID-19 risk" were the most prominent topics co-occurring in the overall vaccine communication. The government tended to emphasize "COVID-19 risk" and "vaccine effectiveness," while regular users prioritized vaccine safety and distribution, indicating a discrepancy in these agendas. Nonetheless, the government has limited impact on regular users in the aspects of vaccine importance, accessibility, affordability, and trust in experts. The agendas of government and nongovernment users intertwined, illustrating complex interactions.

Conclusions: This study reveals the influence of government agendas on public discourse, impacting environmental awareness, public health education, and the social dynamics of inclusive communication during health crises. Inclusive strategies, accommodating public concerns, and involving diverse stakeholders are paramount for effective social media communication during health crises.

(*JMIR Infodemiology* 2024;4:e51113) doi:[10.2196/51113](https://doi.org/10.2196/51113)

KEYWORDS

COVID-19; government; vaccine; automated content analysis; Granger causality test; network agenda setting; QAP; social media

Introduction

As of December 2022, the global COVID-19 pandemic had resulted in 669 million confirmed cases and 6.8 million deaths [1]. Environmental factors were a key determinant significantly

influencing the pandemic [2], through airborne viral infectivity impacted by air pollution and seasonality effects [3,4].

Vaccination was crucial to contain the spread of virus [5], although complex factors such as the Peltzman effect, emerging viral variants, and socioeconomic conditions also affected

pandemic diffusion [6]. Determining an optimal level of vaccination is complex and multifaceted, requiring a balance to avoid undermining democratic values and triggering larger socioeconomic problems than the pandemic [7,8]. Nonetheless, the willingness to vaccinate hinges on various factors, including safety concerns, sociodemographic characteristics, and individual behaviors and attitudes [9,10]. Other determinants including lack of knowledge, government distrust, skepticism about vaccine development, efficacy concerns, exposure experience, coronaphobia, and workplace mandates also predict vaccine uptake [11-13]. As social media becomes increasingly significant for public communication, social media adaptivity, information availability, and health care infrastructure capabilities are also influential for vaccination decisions [14].

Vaccine communication plays a vital role in addressing public concerns, building trust, and encouraging vaccine uptake. Specifically, effective strategies including trusted sources, health provider guidance, a reasonable quantity of information, cultural tailoring, information contextualization, and cultural sensitivity have the potential to significantly increase vaccination intent [15-17]. Despite the notable antagonism in the discourse surrounding immunization on social media [18], it is worth noting that social media campaigns initiated by health organizations have proven to be effective in increasing public awareness about vaccination [19].

Governance mechanisms are another crucial factor for expediting vaccine distribution and mitigating pandemic-related socioeconomic effects [20]. Evidence has shown that clear, consistent, and transparent communication from governmental bodies engendered higher levels of public compliance and trust [21,22]. Given the major impact of the pandemic on public health and society, involvement of the government in vaccine communication becomes a vital research area.

Governments worldwide have adopted diverse approaches to encourage COVID-19 vaccination. For instance, the New Zealand government promoted vaccination among young people by highlighting community factors such as “protecting others” and “striving for herd immunity” [23]. By promoting the scientific notion that there are more advantages than disadvantages to COVID-19 vaccination, the Chinese government has strengthened risk communication to increase the public's awareness of the benefits of vaccines [24]. Although COVID-19 vaccine communication has received increasing attention, particularly from the research community, scientific evidence focusing specifically on low-risk regions, such as Macao, is scarce. This suggests that the existing literature does not sufficiently reflect the concerns of the Macao population as related to COVID-19 vaccination. As one of the world's most densely populated cities, Macao has maintained a record of relatively low risk of infection and high coverage of COVID-19 vaccines [25]. Throughout the pandemic before June 2022, Macao had only recorded 17 confirmed cases of local infection (with a rate of 2.5 cases per 100,000 population) with no fatalities. By June 19, 2022, the vaccine coverage rate within the entire population in Macao was 85.6% for at least 2 doses and 40.5% for 3 doses [26]. The low prevalence of COVID-19 is believed to be the result of the close connection between Macao and mainland China. Since the outbreak of the pandemic,

Macao has implemented anti-epidemic measures following the “dynamic zero-COVID-19 policy” established by mainland China, with some adaptations based on local socioeconomic circumstances [27]. Given the close link between these entities, it is important to understand how the Macao Government communicated with citizens to drive their demand for vaccinations and the impact of this communication. Researchers have long investigated how governments develop policy agendas and whether a policy agenda is led by the government or the public [28]. However, literature on the role of the government in public health agenda setting, specifically related to vaccine promotion in the COVID-19 context, is limited.

The primary goal of this study was to reveal the patterns of vaccine communication on social media during the COVID-19 pandemic as well as the role of the government in advancing vaccination through a case study of Macao, the special administrative region of China. By conducting this research, we aimed to contribute to the existing knowledge on vaccine communication and provide implications for policymakers to improve health promotion communication strategies for preparedness against future pandemics.

The theory of agenda setting suggests that the media has the ability to influence the public agenda by making a specific issue prevalent and salient [29]. Agenda setting is a competition among issue proponents to gain the attention of media professionals, the public, and policy elites [30]. Recently, research about agenda setting has been extended by incorporating the concept of social networks and the associative network of memory, which has been proposed by Guo [31] as the network agenda setting model (NAS). The NAS underlines the associations between topics or attributes presented in the agenda: The more frequently 2 attributes are correlated in news coverage, the more likely the public will perceive them to be interrelated [32].

The NAS can be used to identify the interconnections between public, media, organizational, and government topics on social media. For instance, a study conducted by Chen et al [33] utilized the NAS to investigate the correlation between individual users and organizational accounts on Weibo in terms of their focus on nationalist concerns. The NAS emphasizes the relationship between topics or attributes in constructed agendas. Hou et al [34] analyzed posts mentioning COVID-19 vaccines on Twitter and found that topics related to COVID-19 vaccines can be divided into the following 9 categories: (1) vaccine importance, (2) vaccine effectiveness, (3) vaccine safety, (4) trust in governments, (5) trust in experts, (6) COVID-19 risk, (7) vaccine accessibility, (8) vaccine distribution, (9) vaccine affordability. Additionally, recent studies examined the concerns of all users, including parents, regarding COVID-19 vaccines (eg, [35]). However, these studies did not distinguish between regular accounts (ie, ordinary individual users), government accounts, organization accounts, and media accounts. This distinction is important to understand the nuances of vaccine promotion engaged by different entities. Governments, for instance, influence public discourse through policymaking [24,28], whereas organizations play a significant role in agenda setting via funding, lobbying, and advertising activities [36]. The public, media, and government may construct different

associations among topics in their respective agendas and impact each other. Our research questions (RQ) thus ask the following:

- RQ1: What are the most prevalent agenda attributes emphasized in the communication of vaccination on Facebook during the COVID-19 outbreak in Macao?
- RQ2: How do the attributes interact in the vaccine agendas of governmental and nongovernmental entities?
- RQ3: What are the associations between the vaccine agenda networks constructed by government and nongovernment users?
- RQ4: How do government and nongovernment users impact each other's vaccine agenda on Facebook?

Methods

Sample and Data

This study retrieved data relevant to COVID-19 vaccines in Macao from January 1, 2020, when the SARS-CoV-2 virus was initially detected in China, to August 31, 2022, when the number of newly reported cases had sharply declined [1]. Facebook was selected as the main source of data to analyze the dynamics of vaccine communication in Macao. Being one of the most widely used social media platforms globally, Facebook accounts for a more dominant market share (65.05%) than other sources (eg, Pinterest: 11.47%; Twitter: 10.54%) in Macao [37,38]. The widespread usage of Facebook suggests that it has a significant impact on the population's perceptions, attitudes, and behaviors, making it an essential platform to study to understand the public agenda. In addition, Facebook's archival nature allows for tracking of the evolution of vaccine-related discussions over time, capturing the core dynamics of vaccine communication online.

A combination of the keywords "COVID-19" and "vaccine" as well as their synonyms (ie, 29 synonyms of COVID-19-related terms and 10 synonyms of vaccine-related terms) in Chinese were used to detect and collect relevant posts (see [Multimedia Appendix 1](#)). Information was also compiled on the various labels given to users on Facebook, such as labels of government, media, and organization accounts. Following the collection of raw data from Facebook, data screening was performed to remove duplicate and irrelevant posts. The preprocessing of data including the removal of stop words (eg, "an," "the," "etc.," punctuation, symbols, and numbers) and word segmentation was implemented using the DivoMiner platform.

Ethics Approval

This research strictly adheres to ethical guidelines by ensuring complete anonymity and de-identification of all data sources. To preserve the confidentiality and privacy of all sources involved, no identifiable information about individual users, their IDs, or direct, non-paraphrased posts are included in the main manuscript or any supplementary materials.

Clarification

All identifiers in the data set (eg, names of the senders) were removed and replaced with a code to mask the information about each sender, ensuring the anonymization of our data. Data were

only collected from publicly available posts that were returned based on the structured keyword search criteria.

Measures of Variables

This study investigated the dynamics of agenda setting between government and nongovernment users on Facebook. To achieve this, we categorized users into the following different categories, drawing from prior research [39,40]: (1) media, (2) civil organizations, (3) regular users, (4) government.

The media functions as information gatekeepers and holds potential influence over people's decision-making [29,32]. To account for significant differences in content, news culture, and viewpoints, the media category in this study was further divided into professional media and alternative media for a thorough investigation [41]. Professional media includes those traditional mass media outlets responsible for information dissemination and public awareness, such as newspapers, radio, and television, while alternative media includes independent and electronic media, which is in contrast to mainstream mass media. By referencing relevant media research [42], this study annotated professional media accounts, alongside alternative media accounts.

Civil organizations, also called civil society organizations, include those organizations or associations that are established by individuals or groups with a common purpose or interest and operate in the community, differing from the government and corporations. Civil organizations work alongside the government and other stakeholders to contribute to public discourse, policy development, and social change [43].

Regular users were defined in this study as individuals who interact with Facebook on a personal basis, without representing any official capacity, media, or organizations. Therefore, regular users can be considered as representatives of the public in this study.

The government in this study was defined as all authorities. We did not categorize the specific levels, instead treating all government authorities as a single entity, to gain a clear understanding of the overall position of the Macao Government in vaccine communication. This was also a practice adopted by previous research (eg, [44]).

The classification of Facebook users into 5 distinct categories was conducted based on the information gathered from users' short biographical profiles and the user identity labels provided by Facebook. We assigned 2 coders to classify the users contributing relevant posts. Any confusion that might have occurred during classification was resolved through discussion. This approach allowed for the categorization of users into specific groups, enabling a systematic analysis of user communication and interactions within the Facebook platform [44].

To investigate the dynamics of vaccine communication, 9 predefined categories that indicate elements influencing vaccine acceptance were established based on a coding framework adapted from prior studies (eg, [34,45,46]). These categories included the following topics: importance of vaccines, effectiveness of vaccines, safety of vaccines, trust in

governments, trust in experts, risk of the COVID-19 pandemic, and vaccine convenience (ie, accessibility, distribution, and affordability). Details of the coding categories are shown in [Multimedia Appendix 2](#).

Data Analysis Procedures

Automated Content Analysis

In this study, an automated content analysis method was used to identify and categorize posts into the predefined categories. Each post could belong to one or more categories or none at all. The effectiveness of automated coding depends on the design of the keywords. To develop accurate keywords, this study followed the approach outlined by Chang et al [37] using the Word2vec word embedding toolkit from the Python 3.7.4 Gensim module [47]. Word2vec, a word embedding technique powered by neural networks, allows the identification of words with similar meanings by analyzing word associations in a large text corpus [48]. Due to the intricacies of the Chinese language, the synonyms suggested by Word2vec were further checked by assessing their relevance to the context. On this basis, the Chinese thesaurus and relevant literature [49] were further consulted for the inclusion of additional synonyms. The list of keywords for machine coding can be found in [Multimedia Appendix 3](#).

DivoMiner, a text mining and automated content analysis platform driven by machine learning algorithms, was used to facilitate the automated content coding task. This platform integrates automated content analysis with traditional content analysis methods and has been widely utilized in health and communication studies [37,50,51]. Following automated coding, manual verification was conducted to ensure the accuracy and reliability of the machine-generated outcomes. To achieve this, 2 coders, both native Cantonese speakers, were recruited and underwent 36 hours of training to independently code 300 messages. Each variable was coded as either present or absent. Discrepancies between the coders were resolved through discussions, with the author intervening only when consensus could not be reached between the coders. The overall intercoder reliability, measured using Krippendorff alpha, demonstrated satisfactory levels across all examined variables, with coefficients ranging from .77 to .82. The consistency between machine coding and manual coding reached an acceptable level, with an average score of 74%. This score aligns with previous studies, in which a threshold value of 70% was considered rational [49-51].

Statistical Analysis

The conventional statistical analysis in this study involved the use of SPSS (version 23; IBM Corp) for analysis. Categorical variables were summarized using counts and percentages. The chi-square test of independence was used, and post hoc comparisons with Bonferroni corrections were further implemented to precisely identify the specific significant differences between user categories and vaccine-related topics and avoid the likelihood of generating false-positive outcomes (type I errors).

Co-Occurrence Network Analysis

Co-occurrence matrices, which represent the strength of ties between 2 topics engaged by different users, were generated as dyadic data sets. Based on the co-occurrence data, this study established undirected and weighted topic co-occurrence networks. Each network represents the co-occurrence relations of the attributes of a certain user category. To clarify, if a particular category of user mentions topic “i” and topic “j,” a band will link “i” and “j.” The width of the band indicates the frequencies of the pair of topics discussed by a user type [52,53]. For example, in the professional media user category’s topic co-occurrence network, if a professional media news report mentions the topics of “vaccine importance” and “vaccine effectiveness” together, the topics will be linked in the network by a band. The more frequently these topics co-occur, the thicker the band becomes. The visualization of topic co-occurrence is presented in a chord diagram by Echarts (The Apache Software Foundation), as indicated by Wang et al [52].

Quadratic Assignment Procedure for Network Analysis

In this study, the quadratic assignment procedure (QAP) method was applied to understand the correlation between the Macao Government’s agenda network and that of other Facebook users, via analysis of the co-occurrence matrices. QAP is a common method in social network or agenda network studies [40,54]. QAP correlation analysis can be used to assess the correlation between 2 matrices with the Pearson correlation coefficient, while QAP regression analysis can determine whether an explanatory variable can predict an outcome variable when the 2 matrices are significantly correlated [55]. In this study, the QAP method used UCINET 6.730 to test whether the Macao Government’s vaccine agenda network has impacted that of nongovernment Facebook users, particularly regular type users, during the COVID-19 pandemic.

Vector Autoregression Modeling

The vector autoregression (VAR) approach was used to examine the dynamic of agenda attributes between government and nongovernment users. This approach evaluates the effect of an observed variable by considering its lagged effect in the earlier period and that of other predictors in previous time points, without presuming the associations between the variables [56]. The VAR modeling technique is widely used in the economic field and, in recent years, has been increasingly applied in research on health science, sociology, neuroimaging, and meteorology (eg, [54,57-59]).

VAR modeling is ideal for measuring the dynamic performance response and interaction between performance and marketing communication variables. A study applied VAR models to construct the dynamic response relationship between news stories and public attention using a combination of survey and news content ranging from 2009 to 2013 [60]. The VAR models captured the dynamic feedback system and gave estimates for the short-term effects of TV news coverage on public perception by demonstrating a unidirectional process wherein changes in news salience led to significant changes in public salience. In addition, VAR models have also been used to investigate the dynamic mapping relationship between the diffusion of political

messages and emotional expression in public messages during the COVID-19 pandemic [61]. The increased diffusion of political messages positively predicted changes in emotional expression among citizens, and the VAR model was able to explain the interdependencies among variables based on the lag values of multiple time series. Overall, the VAR model proves to be an insightful tool for analyzing complex relationships in communication studies, providing insights into the short-term and long-term effects of various factors on outcomes of interest. Hence, using the VAR technique allows the exploration of temporal dynamics and associations between different agenda attributes in this study. For example, the approach enables a better understanding of whether the agenda attributes propagated by the government (AG) at time (t-n) impacts the agenda attributes of nongovernment users (AN) including professional media, alternative media, civil organizations, and regular users. The VAR model was generated as follows:



Within this model, α_i and β_i are the estimated coefficients, ρ represents the optimal number of lags for the model, and ϵ indicates the error term. AG_{t-i} and AN_{t-i} represent the respective variable at the earlier periods. For instance, AG_{t-1} indicates the first lag of AG. The lag length for the VAR model was selected as per the Akaike information criterion. The augmented Dickey-Fuller test was applied to examine the stationarity of the time series. For nonstationary series, differencing at the first or higher level was performed to achieve stationarity [62]. When both time series were stationary at the same level, this study proceeded with the Johansen maximum eigenvalue and trace tests based on the estimation of VAR models to determine whether the time series were cointegrated and suitable for Granger causality tests. Granger causality posits that causes lead to effects and happen before their effects [40]. In this sense, using prior values of a time series can statistically forecast the future status of another time series.

In this study, the Granger causality test was used to provide greater insight into the statistical causal relationship between the government's agenda and the nongovernment users' agenda. To estimate VAR models and enable Granger causality tests, this study transformed the collected data in the form of time series by dividing the data into 32 monthly periods (from January 2020 to August 2022), and each monthly period was treated as an independent unit for analysis. EViews 12 software was used for statistical analysis.

Results

Results of Content Analysis

This research initially collected a sample of 24,089 Facebook posts with relevance to COVID-19 vaccines. Data screening

was further performed on the sample to remove duplicated, irrelevant, and unclear messages, resulting in 23,577 unique and relevant posts. Finally, the results of machine coding presented a total of 22,986 posts that include the examined vaccine topics.

In answering RQ1, we calculated the frequency of the vaccine topics and found that the majority of posts in the sample related to the importance of COVID-19 vaccination (7358/22,986, 32.01%), followed by posts that indicated the high risk of contracting COVID-19 (6877/22,986, 29.92%) and highlighted trust in experts (4320/22,986, 18.79%). In addition, a considerable number of posts mentioned vaccine effectiveness (4163/22,986, 18.11%), safety (3358/22,986, 14.61%), accessibility (2683/22,986, 11.67%), distribution (2492/22,986, 10.84%), and affordability (1685/22,986, 7.33%), while posts related to trust in government were less frequent (1593/22,986, 6.93%). In addition, in the overall vaccine-related discussion, nongovernment users comprised a substantial majority of the posts, at 76.85% (17,665/22,986). When examining the nongovernment user segment at a more granular level, professional media accounted for a significant proportion of the posts, at 33.87% (7555/22,986), followed by alternative media, at 12.24% (2814/22,986); civil organizations, at 3.99% (918/22,986); and regular users, at 27.74% (6377/22,986). The topics associated with vaccine agenda attributes by government and nongovernment users are shown in [Table 1](#).

The chi-square test indicated that the distributions of vaccine-related topics were significantly different across the user categories ($\chi^2_{32}=1579.469, P<.001$). The outcomes of the post hoc comparisons suggested that the government was more concerned with topics of vaccine effectiveness (1003/5322, 18.85%; $P<.001$), COVID-19 risk (1805/5322, 33.92%; $P<.001$), vaccine accessibility (1010/5322, 18.98%; $P<.001$), and vaccine affordability (605/5322, 11.37%; $P<.001$), while discussion of vaccine safety (393/5322, 7.38%; $P<.001$), government trust (133/5322; 2.5%, $P<.001$), expert trust (518/5322, 9.73%; $P<.001$), and vaccine distribution (341/5322, 6.41%; $P<.001$) occurred to a less extent than for other users. In comparison, professional media contributed more to the topics of government trust (752/7555, 9.95%; $P<.001$) and expert trust (1895/7555, 25.08%; $P<.001$). Alternative media, however, were less inclined to discuss vaccine affordability (128/2814, 4.55%; $P<.001$) than other categories of users. Regular users were primarily concerned about vaccine safety (1092/6377, 17.12%; $P<.001$) and vaccine distribution (724/6377, 11.35%; $P<.001$) and were less concerned about vaccine effectiveness (937/6377, 14.69%; $P<.001$), COVID-19 risk (1529/6377, 23.98%; $P<.001$), and vaccine accessibility (416/6377, 6.52%; $P<.001$) than other users. The outcomes of the post hoc tests with details are shown in [Multimedia Appendix 4](#).

Table 1. Overview of the vaccine agenda attributes by government and nongovernment users in Macau from January 1, 2020, to August 31, 2022.

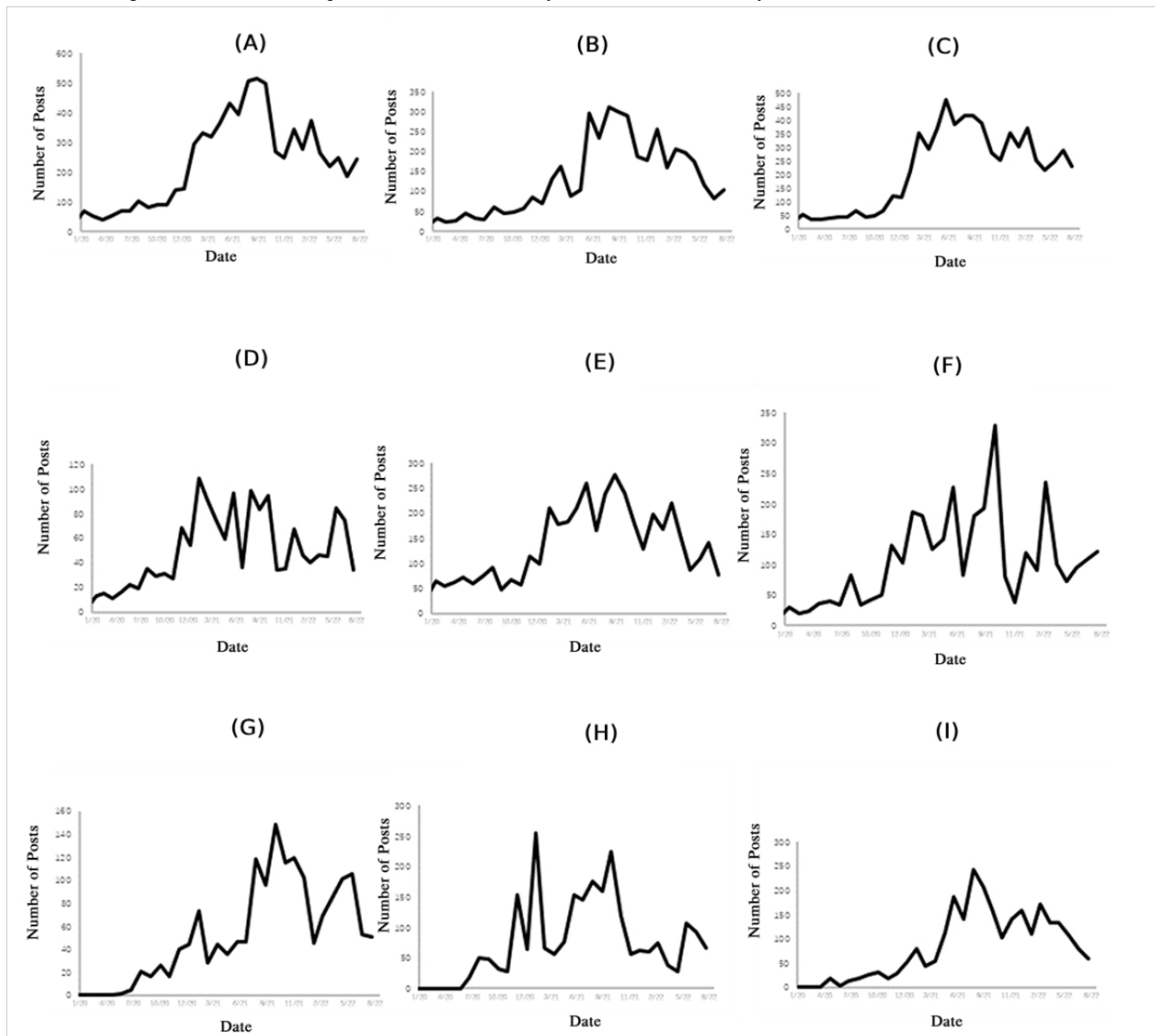
Vaccine topic	Government users, n (%)	Nongovernment users, n (%)				Total, n (%)
		Professional media	Alternative media	Civil organizations	Regular users	
All posts	5322 (23.15)	7555 (32.87)	2814 (12.24)	918 (3.99)	6377 (27.74)	22,986 (100)
Importance	1616 (30.36)	2931 (38.80)	697 (24.77)	298 (32.46)	1816 (28.48)	7358 (32.01)
Effectiveness	1003 (18.85)	1638 (21.68)	404 (14.36)	181 (19.72)	937 (14.69)	4163 (18.11)
Safety	393 (7.38)	1374 (18.19)	359 (12.76)	140 (15.25)	1092 (17.12)	3358 (14.61)
Trust in government	133 (2.5)	752 (9.95)	175 (6.22)	40 (4.36)	493 (7.73)	1593 (6.93)
Trust in experts	518 (9.73)	1895 (25.08)	593 (21.07)	154 (16.78)	1160 (18.19)	4320 (18.79)
COVID-19 risk	1805 (33.92)	2651 (35.09)	681 (24.2)	211 (22.98)	1529 (23.98)	6877 (29.92)
Accessibility	1010 (18.98)	981 (12.98)	196 (6.97)	80 (8.71)	416 (6.52)	2683 (11.67)
Distribution	341 (6.41)	1005 (13.3)	309 (10.98)	113 (12.31)	724 (11.35)	2492 (10.84)
Affordability	605 (11.37)	529 (7)	128 (4.55)	53 (5.77)	370 (5.8)	1685 (7.33)

Trend in Facebook Activities

To reveal the dynamics of different attributes of the vaccine agenda, this study mapped trends of these attributes during the investigated period. All vaccine-relevant content remained at a relatively low volume in 2020 and increased significantly in 2021. The volume of content regarding “vaccine distribution” began to grow at the start of 2021 and showed an observable spike in February of the same year. This was followed by a

sharp acceleration in content regarding the high risk of COVID-19 reaching its peak in June 2021. The highest peak in vaccine-relevant content occurred in September 2021 related to the topic of vaccine importance. Between June 2021 and October 2021, the most debate centered around themes relating to COVID-19 vaccines. Overall, variations in the volume of vaccine communication were observed over time. [Figure 1](#) shows the dynamic of vaccine discussion showing the monthly volume of posts.

Figure 1. Temporal changes in the vaccine agenda attributes (January 2020–August 2022): (A) vaccine importance, (B) vaccine effectiveness, (C) risk of COVID-19, (D) government trust, (E) expert trust, (F) vaccine safety, (G) vaccine affordability, (H) vaccine distribution, (I) vaccine accessibility.

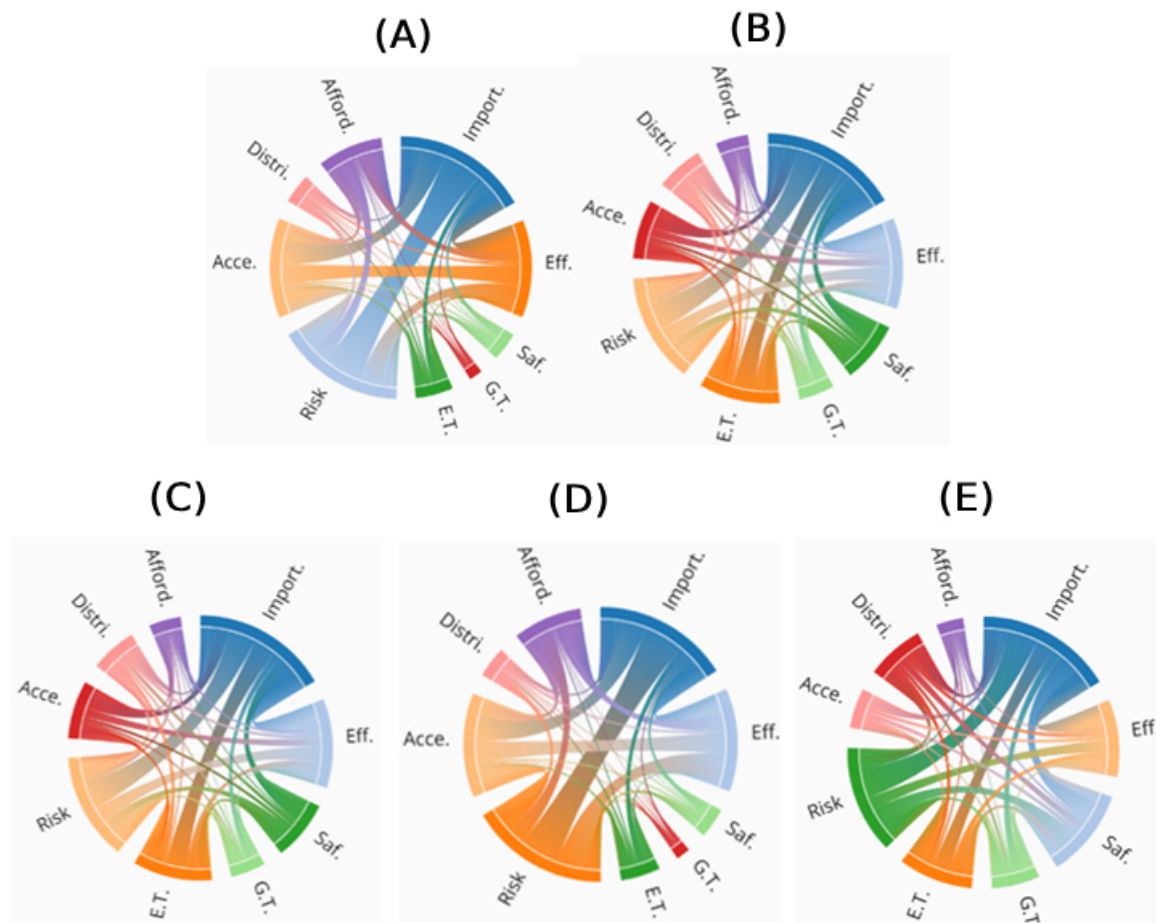


Interactions Between Agenda Attributes in Vaccine Communication

To answer RQ2, this study computed the interrelationships between agenda attributes by the government and nongovernment users by constructing co-occurrence matrices. Results showed that “vaccine importance,” “vaccine effectiveness,” and “COVID-19 risk” were the most prominent attributes interacting with each other in the agendas of government and nongovernment users, except for the regular users’ agenda in which “vaccine safety” (n=2503) rather than “vaccine effectiveness” (n=2161) had more established connections overall with other attributes. Specifically, the government agenda featured strong connections between “vaccine importance” and “COVID-19 risk” (n=1505), followed by “vaccine importance” and “vaccine effectiveness” (n=945), “vaccine importance” and “accessibility” (n=940), and “COVID-19 risk” and “accessibility” (n=816). As for the agenda of professional media, the strongest link was established between “vaccine importance” and “COVID-19 risk” (n=1528), followed

by the link between “vaccine importance” and “vaccine effectiveness” (n=1327) and the link between “vaccine importance” and “trust in experts” (n=1220). In terms of regular users, their agenda highlighted the relationships between “vaccine importance” and “COVID-19 risk” (n=931), “vaccine importance” and “vaccine effectiveness” (n=655), “vaccine importance” and “vaccine safety” (n=644), “vaccine importance” and “trust in experts” (n=536), and “vaccine safety” and “COVID-19 risk” (n=469). Using chord diagrams, this study visualized the interrelationships of agenda attributes by different user categories. The arc in the outer ring represents the attributes of the vaccination agenda and is differentiated by color. The arc length indicates the total number of associations an attribute maintains with other attributes when communicated by users in a specific category. The band within the ring represents the connected relationship between 2 topics, with the thickness of the band indicating the magnitude of the connection. A set of chord diagrams revealing agenda attribute interactions in the agendas with comparison of different users is presented in [Figure 2](#).

Figure 2. Comparison of agenda attribute interactions by different users: (A) government, (B) professional media, (C) alternative media, (D) civil societal organizations, (E) regular users. Acce.: vaccine accessibility, Afford.: vaccine affordability, Distri.: vaccine distribution, Eff.: vaccine effectiveness, E.T.: expert trust, G.T.: government trust, Import.: vaccine importance, Risk: risk of COVID-19, Saf.: vaccine safety.



To assess the evolution of links between attributes over time, this study also divided the co-occurrence dynamics of intragroup agenda attributes into 3 distinct periods: 2020, 2021, and 2022. Our findings revealed that the connections between agenda attributes varied by both the time period and the categories of Facebook users. Notably, in the government agenda, the link between “vaccine effectiveness” and “vaccine affordability” exhibited an increase in strength during 2022 (795/4150, 19.16%), compared with 2020 (18/223, 8.25%) and 2021 (690/4744, 14.54%). Conversely, the connection between “vaccine importance” and “expert trust” within the agenda of regular users demonstrated a decline in frequency over the 3-year span (2020: 119/1165, 10.21%; 2021: 282/3417, 8.25%; 2022: 118/1779, 6.63%). More information about the co-occurrence dynamics of the intragroup agenda attributes over time can be found in [Multimedia Appendix 5](#).

Agenda Network Analysis

In answering RQ3, the results of the QAP tests demonstrated significantly positive and strong correlations between the agenda network of the government and those of professional media ($r=0.745$, $P=.005$) and civil organizations ($r=0.632$, $P=.02$). However, the correlations between the government’s agenda network and the network of alternative media ($r=0.462$, $P=.08$) and regular users ($r=.451$, $P=.07$) were not statistically significant.

The subsequent QAP linear regression analysis tested whether the agenda network of the Macao government can predict that of nongovernment users. For example, by using the government as a predictor and different types of nongovernment users as outcome variables, the results demonstrated that the government has an impact on the agenda network of professional media ($b=0.703$, $P=.006$) and civil organizations ($b=0.051$, $P=.02$). The adjusted R^2 value for professional media indicated that government accounts for around 54% of the variance in the professional media’s agenda network, while government only accounts for 38% of the variance in the agenda network of civil organizations. The results of the QAP linear regression analysis with the government as a predictor are shown in [Table 2](#).

In the QAP linear regression model predicting the agenda of regular users, the results revealed significant impacts of alternative media ($b=2.46$, $P=.001$), professional media ($b=0.52$, $P=.001$), and civil organizations ($b=6.16$, $P=.001$) on the agenda of regular users. The adjusted R^2 value for professional media, civil organizations, and alternative media ranged from 0.81 to 0.86, suggesting that all 3 categories of users can explain 81%-86% of the variance in the regular users’ agenda network. The results of the QAP linear regression analysis with regular users as the outcome variable are shown in [Table 3](#).

Table 2. Quadratic assignment procedure regression analysis with government as the predictor.

User category	Unstandardized coefficient	<i>P</i> value ^a	<i>R</i> ² value	Adjusted <i>R</i> ²
Civil organizations	0.051	.02	0.399	0.382
Professional media	0.703	.006	0.556	0.543
Alternative media	0.095	.10	0.214	0.191
Regular users	0.246	.12	0.204	0.180

^aOutcomes were considered statistically significant at $P < .05$.

Table 3. Quadratic assignment procedure regression analysis with regular users as the outcome variable.

User category	Unstandardized coefficient	<i>P</i> value ^a	<i>R</i> ² value	Adjusted <i>R</i> ²
Government	0.246	.12	0.204	0.180
Alternative media	2.462	.001	0.868	0.864
Professional media	0.521	.001	0.811	0.805
Civil organizations	6.164	.001	0.832	0.827

^aOutcomes were considered statistically significant at $P < .05$.

Impacts of Government and Nongovernment Users on Each Other's Vaccine Agenda

To answer RQ4, the Granger causality test was further performed to examine whether the 9 attributes in the government's agenda statistically predicted the future intensity of topics discussed by different categories of users and vice versa. Specifically, the results showed that attributes such as "vaccine safety" ($F_{3,13}=3.817$; $P=.04$) and "trust in experts" ($F_{3,13}=3.916$; $P=.03$) in the government's agenda significantly affected such attributes in the agenda of nongovernment users, while the attributes associated with "trust in government" ($F_{3,13}=4.590$; $P=.02$) and "vaccine affordability" ($F_{3,13}=3.851$; $P=.04$) in the agenda of nongovernment users affected these attributes in the agenda of the government at the significance level of $P < .05$.

By classifying nongovernment users into different user categories, the results suggested a unidirectional trend in the attribute of "vaccine safety" flowing from the government's agenda to that of professional media ($F_{5,15}=3.247$; $P=.03$), while professional media affected the agenda of the government unilaterally through the attributes of "vaccine importance" ($F_{5,12}=7.192$; $P=.003$), "vaccine effectiveness" ($F_{3,13}=4.391$; $P=.02$), "COVID-19 risk" ($F_{5,15}=5.173$; $P=.006$), and "vaccine

affordability" ($F_{3,13}=4.754$; $P=.02$). Additionally, alternative media affected the government by setting the agenda with attributes such as "COVID-19 risk" ($F_{5,15}=8.769$; $P < .001$) and "vaccine accessibility" ($F_{5,15}=2.963$; $P=.047$), while there was no temporal causation from the government to alternative media for the attributes identified.

Regarding civil organizations, the government predicted the agenda of civil organizations through the attributes of "vaccine importance" ($F_{5,15}=4.111$; $P=.01$), "vaccine effectiveness" ($F_{3,13}=6.264$; $P=.007$), and "trust in experts" ($F_{3,9}=15.877$; $P=.001$), while the causation from civil organizations to the government was absent for all attributes except "vaccine safety" ($F_{3,12}=4.405$; $P=.03$).

Most notably, the Granger causality analysis revealed that the government had a significant impact on the agenda of regular users through the attributes of "vaccine importance" ($F_{5,15}=3.809$; $P=.02$), "trust in experts" ($F_{5,15}=16.639$; $P < .001$), "vaccine accessibility" ($F_{5,15}=3.343$; $P=.03$), and "vaccine affordability" ($F_{3,13}=6.012$; $P=.008$). Despite the absence of Granger causality from regular users to the government for most attributes, there was a reciprocal relationship between the government and regular users in the attribute of "vaccine affordability." The results of the Granger causality tests between the government and other types of users are shown in [Table 4](#).

Table 4. Granger causality tests between government users and other types of users for each vaccine attribute.

Vaccine attribute	Nongovernment users		Professional media		Alternative media		Civil societal organizations		Regular users	
	Outcome variable	Antecedent variable	Outcome variable	Antecedent variable	Outcome variable	Antecedent variable	Outcome variable	Antecedent variable	Outcome variable	Antecedent variable
Importance										
<i>F</i> value (df)	1.410 (5,20)	1.209 (2,20)	1.413 (5,20)	7.192 (5,12)	2.412 (5,15)	2.407 (5,15)	4.111 (5,15)	1.801 (5,15)	3.809 (5,15)	2.259 (5,15)
<i>P</i> value	.26	.32	.26	.003	.09	.09	.01	.17	.02	.10
Effectiveness										
<i>F</i> value (df)	0.449 (2,30)	3.029 (3,13)	0.133 (2,30)	4.391 (3,13)	0.293 (3,13)	1.319 (5,9)	6.264 (3,13)	0.567 (2,10)	0.968 (2,30)	0.858 (2,30)
<i>P</i> value	.64	.07	.88	.02	.83	.34	.007	.58	.39	.44
Safety										
<i>F</i> value (df)	3.817 (3,13)	3.222 (3,13)	3.247 (5,15)	2.565 (5,15)	0.706 (1,15)	2.419 (5,15)	2.923 (3,13)	4.405 (3,12)	2.004 (5,15)	2.912 (3,22)
<i>P</i> value	.04	.057	.03	.07	.41	.08	.07	.03	.14	.057
Trust in government										
<i>F</i> value (df)	2.017 (3,15)	4.590 (3,13)	3.270 (3,13)	3.924 (3,13)	1.228 (2,20)	2.296 (3,13)	2.705 (3,9)	3.585 (2,10)	0.304 (2,10)	3.373 (3,13)
<i>P</i> value	.15	.02	.055	.03	.31	.12	.11	.07	.74	.051
Trust in experts										
<i>F</i> value (df)	3.916 (3,13)	0.402 (2,20)	3.753 (2,10)	1.437 (5,30)	0.401 (2,20)	1.146 (2,20)	15.877 (3,9)	1.058 (1,22)	16.639 (5,15)	4.189 (2,9)
<i>P</i> value	.03	.67	.06	.24	.67	.34	.001	.31	<.001	.051
COVID-19 risk										
<i>F</i> value (df)	0.255 (2,9)	1.124 (2,30)	1.890 (2,30)	5.173 (5,15)	0.665 (3,3)	8.769 (5,15)	2.442 (3,20)	2.275 (3,20)	0.655 (2,15)	0.235 (2,15)
<i>P</i> value	.78	.34	.16	.006	.63	<.001	.09	.11	.53	.79
Accessibility										
<i>F</i> value (df)	0.248 (2,15)	2.781 (3,13)	0.045 (2,10)	1.362 (5,20)	1.461 (5,15)	2.963 (5,15)	1.546 (5,15)	2.763 (5,10)	3.343 (5,15)	2.376 (5,15)
<i>P</i> value	.78	.08	.96	.28	.26	.047	.23	.08	.03	.09
Distribution										
<i>F</i> value (df)	0.756 (2,20)	0.104 (2,20)	0.596 (2,20)	0.283 (2,20)	0.147 (1,25)	0.005 (1,15)	1.264 (1,20)	0.382 (1,20)	4.175 (1,25)	0.458 (1,25)
<i>P</i> value	.48	.90	.56	.76	.70	.94	.27	.54	.051	.50
Affordability										
<i>F</i> value (df)	2.500 (3,13)	3.851 (3,13)	0.745 (2,20)	4.754 (3,13)	0.479 (2,20)	0.688 (2,20)	0.525 (2,20)	0.495 (2,20)	6.012 (3,13)	5.067 (2,20)
<i>P</i> value	.10	.04	.49	.02	.63	.51	.60	.62	.008	.02

Discussion

Principal Findings

This study examined the dynamics and patterns of vaccine communication on Facebook in Macao during the COVID-19 pandemic. The principal findings demonstrated that “vaccine importance” was the most prevalent attribute in the vaccination agenda on Facebook, followed by the attributes of “COVID-19

risk” and “trust in experts.” The overall vaccination agenda revealed the highest co-occurrences were between “vaccine importance” and “COVID-19 risk.” Differences existed in agenda priorities between the government and regular users. The government primarily focused on the risks of COVID-19 and the effectiveness of vaccines, whereas regular users were more concerned with the safety and distribution of vaccines. The Macao government played a role in shaping the agenda for

regular users by highlighting vaccine importance (Granger causality result: $F_{5,15}=3.809$; $P=.02$), trust in experts (Granger causality result: $F_{5,15}=16.639$; $P<.001$), and vaccine accessibility (Granger causality result: $F_{5,15}=3.343$; $P=.03$) and affordability (Granger causality result: $F_{3,13}=6.012$; $P=.008$), while its impact on the agenda network of regular users remained insignificant (QAP result: $b=0.246$; $P=.12$). Both government and nongovernment users (eg, professional media, alternative media, civil organizations, and regular users) had intertwined agendas with mutual influence.

Unlike previous studies that predominantly focused on single aspects of vaccine communication (eg, [17,34]), this study used a more holistic approach to reveal the role of various actors including the government, professional media, alternative media, civil organizations, and regular users in promoting vaccination agendas and the interplay of diverse actors in the vaccine agenda setting process. The results of this study suggest that professional media acts as more than simple information providers to the government but rather effectively pushed agenda setting as a supplementary process to vaccine promotion by raising salient topics that the government fails to identify due to lack of information and experience. The government, however, is more likely to respond to professional media to receive timely feedback on vaccination issues for the purpose of learning and improvement. This can be observed from the impact that professional media has on the government in the agenda setting process through topics of “vaccine importance” (Granger causality results: $F_{5,12}=7.192$; $P=.003$), “vaccine effectiveness” (Granger causality results: $F_{3,13}=4.391$; $P=.02$), “trust in government” (Granger causality results: $F_{3,13}=3.924$; $P=.03$), “COVID-19 risk” (Granger causality results: $F_{5,15}=5.173$; $P=.006$), and vaccine affordability (Granger causality results: $F_{3,13}=4.754$; $P=.02$).

Who Leads the Vaccine Agenda of Whom?

Despite a significant correlation between the government agenda network and the agenda network of nongovernment users, the government had a limited impact on the agenda attributes of different Facebook user categories and vice versa. As Facebook is an open platform where information from a wide variety of sources freely circulates and interacts, it is difficult to determine the driving force behind the vaccine promotion agenda on the platform [55]. In other words, nongovernment users' vaccine promotion agendas may have been impacted by other sources, such as the World Health Organization or other health professionals, which indicates a multidirectional effect.

As such, it appears that the government did not unilaterally set the agenda of nongovernment users. Instead, there is a “2-way” interaction between government and nongovernment user agendas. Due to their mutual effect, neither the government nor nongovernment users lead the agenda on social media. It is likely that the government and different types of nongovernment users pay attention to the agendas of one another and interact with one another to build the overall vaccine agenda network on Facebook. This corresponds with the argument by Finset et al [63] that, amid the near-chaotic flow of information, every individual, in different roles and with varied responsibilities,

can contribute to the development of the information flow and agenda on COVID-19. A plausible explanation for this outcome could be the unprecedented nature of the health crisis. The lack of up-to-date crisis communication planning and experience with coping with a novel crisis may challenge the government's agenda-setting process, particularly in terms of vaccine promotion.

Comparison With Prior Work

Previous agenda setting research found that changes in the government agenda led to changes in the public agenda [64]. However, during the COVID-19 pandemic, the public was no longer passive consumers of social media. Our results indicating the different concerns of vaccination between the government and regular users corroborate previous findings by Zhou and Zheng [44] who found that, during the COVID-19 pandemic, the government's Weibo account exhibited a more propaganda-oriented approach, whereas public accounts were more attentive to issues that directly pertained to self-interest, such as protective measures against the virus and minimizing financial losses. Unlike other political issues, the government may have less impact on shaping public agenda due to the more collected information possessed by the public. This is partly consistent with some recent research indicating that shaping public opinion in a fragmented digital environment such as social media is challenging [54,65]. Additionally, the case of Macao also indicates selective public responsiveness on topics that are clear and straightforward, which partially verifies the observation by Kim [66] that individuals are more receptive to topics that are unambiguous and do not demand extensive background knowledge as they may not have enough background information with which to fully process any new information on complex topics.

Practical Implications

Our study provides several implications to inform the management of future pandemics. First, given the disparity between the government and public agenda networks, it is crucial to bridge the gap to enable effective vaccine communication. Policymakers should strive for alignment between government messaging and public concerns, addressing issues that are prominent within the public discourse. Social media listening activities are invaluable tools for understanding public health concerns. By monitoring public conversation through social media listening, policymakers can develop targeted messaging and communication strategies that effectively address public concerns and provide accurate information to dispel misconceptions.

Second, the low responsiveness of the public agenda to the government agenda indicates the need to enhance the government impact on the public agenda. Governments can streamline their messaging by using plain language, which helps individuals with different levels of knowledge understand information easily. Clear and concise presentation avoids unnecessary complexity. Visual aids and interactive media can also be used to improve public involvement and responsiveness, overcoming barriers caused by limited background information.

Third, policymakers' efforts to convince the public to receive vaccines in response to potential health risks have been shown in our study to lead to a spillover of media attention that significantly drives the vaccination agenda among the public. Collaboration with influential media, including professional and alternative media, thus offers a powerful means to facilitate vaccination policy and improve public health. Governments can utilize the extensive reach and persuasive power of media outlets to actively involve and inform the public about specific issues that should receive priority attention, thereby advancing the government's crisis management initiatives.

Fourth, civil organizations' ability to shape public attention toward vaccination issues by influencing the public agenda network suggests that their impact on shaping the vaccination agenda may be underestimated or overlooked. Driven by social responsibility, civil organizations often dedicate their efforts to promoting public health by increasing awareness and advocating for public health policies [43]. The close ties to communities enable them to be trusted sources of information for the public. Therefore, through partnerships with civil organizations, governments can leverage their networks, expertise, and community trust to effectively promote vaccination initiatives.

Limitations

Several limitations warrant consideration. First, broadening the scope beyond vaccines to encompass diverse policies could offer a more comprehensive understanding of public attention allocation mechanisms. Researchers are encouraged to explore various policies to enhance generalizability. Second, although Facebook data provided valuable insights, the findings are

platform-specific and may not apply universally. Future studies should incorporate a diverse set of social media platforms and combine quantitative data with surveys and interviews for a more nuanced perspective. Third, although this study explored temporal agenda dynamics, it did not delve into the determinants driving public attention intensity, such as government transparency and issue salience. Investigating these factors could provide valuable insights into the agenda setting process at the government level.

Conclusions

This study investigated the communication dynamics of COVID-19 vaccines in Macao, with a specific focus on how government agendas impact other entities on Facebook. Our results reveal that the Macao Government's efforts to set the vaccination agenda on Facebook have shown limited effectiveness in shaping the public's discourse and priorities regarding vaccines. Such findings have profound implications for shaping government responses to future pandemics. Authorities, in their endeavor to legitimize policies, must recognize the intricate interplay between their agendas and public reception. Although agenda setting serves as a strategic tool to promote vaccination, it also exhibits limitations. This requires a shift toward more nuanced, strategy-focused research. This study offers indispensable insights in the area of crisis communication, underscoring the urgent necessity of bridging the gap between government and public agendas. Furthermore, it illuminates the potential of collaborations with influential media outlets and civil organizations as formidable channels to augment the reach and influence of vaccination agendas set by the government.

Authors' Contributions

XX conceptualized the study, curated the data, wrote the original manuscript draft, and created the visualizations. XX and AC performed the formal analysis. XX and RJN validated the data. AC and RJN reviewed and edited the manuscript. AC supervised the study and served as project administrator. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Keywords for vaccine-related data acquisition.

[DOCX File, 16 KB - [infodemiology_v4i1e51113_app1.docx](#)]

Multimedia Appendix 2

Coding framework for COVID-19 vaccine posts on Facebook.

[DOCX File, 21 KB - [infodemiology_v4i1e51113_app2.docx](#)]

Multimedia Appendix 3

Keywords for machine coding of vaccine-related topics.

[DOCX File, 19 KB - [infodemiology_v4i1e51113_app3.docx](#)]

Multimedia Appendix 4

Outcomes of post hoc tests on the significant difference between user categories and vaccine-related topics.

[DOCX File, 21 KB - [infodemiology_v4i1e51113_app4.docx](#)]

Multimedia Appendix 5

Intra-group co-occurrences dynamics of agenda attributes for the years of 2020, 2021, and 2022.

[[DOCX File, 2025 KB - infodemiology_v4i1e51113_app5.docx](#)]

References

1. COVID-19 Coronavirus Pandemic Statistics. Worldometer. URL: <https://www.worldometers.info/coronavirus/> [accessed 2024-02-10]
2. Coccia M. Sources, diffusion and prediction in COVID-19 pandemic: lessons learned to face next health emergency. *AIMS Public Health* 2023;10(1):145-168 [FREE Full text] [doi: [10.3934/publichealth.2023012](https://doi.org/10.3934/publichealth.2023012)] [Medline: [37063362](https://pubmed.ncbi.nlm.nih.gov/37063362/)]
3. Coccia M. Factors determining the diffusion of COVID-19 and suggested strategy to prevent future accelerated viral infectivity similar to COVID. *Sci Total Environ* 2020 Aug 10;729:138474 [FREE Full text] [doi: [10.1016/j.scitotenv.2020.138474](https://doi.org/10.1016/j.scitotenv.2020.138474)] [Medline: [32498152](https://pubmed.ncbi.nlm.nih.gov/32498152/)]
4. Coccia M. COVID-19 pandemic over 2020 (withlockdowns) and 2021 (with vaccinations): similar effects for seasonality and environmental factors. *Environ Res* 2022 May 15;208:112711 [FREE Full text] [doi: [10.1016/j.envres.2022.112711](https://doi.org/10.1016/j.envres.2022.112711)] [Medline: [35033552](https://pubmed.ncbi.nlm.nih.gov/35033552/)]
5. Magazzino C, Mele M, Coccia M. A machine learning algorithm to analyse the effects of vaccination on COVID-19 mortality. *Epidemiol Infect* 2022 Sep 12;150:e168 [FREE Full text] [doi: [10.1017/S0950268822001418](https://doi.org/10.1017/S0950268822001418)] [Medline: [36093862](https://pubmed.ncbi.nlm.nih.gov/36093862/)]
6. Coccia M. COVID-19 vaccination is not a sufficient public policy to face crisis management of next pandemic threats. *Public Organiz Rev* 2022 Oct 17;23(4):1353-1367 [FREE Full text] [doi: [10.1007/s11115-022-00661-6](https://doi.org/10.1007/s11115-022-00661-6)]
7. Coccia M. Optimal levels of vaccination to reduce COVID-19 infected individuals and deaths: A global analysis. *Environ Res* 2022 Mar;204(Pt C):112314 [FREE Full text] [doi: [10.1016/j.envres.2021.112314](https://doi.org/10.1016/j.envres.2021.112314)] [Medline: [34736923](https://pubmed.ncbi.nlm.nih.gov/34736923/)]
8. Coccia M. Improving preparedness for next pandemics: Max level of COVID-19 vaccinations without social impositions to design effective health policy and avoid flawed democracies. *Environ Res* 2022 Oct;213:113566 [FREE Full text] [doi: [10.1016/j.envres.2022.113566](https://doi.org/10.1016/j.envres.2022.113566)] [Medline: [35660409](https://pubmed.ncbi.nlm.nih.gov/35660409/)]
9. Wagner AL, Zhang F, Kerekes S, Shih S, Zhao L. COVID-19 vaccination preferences during a pause in Johnson and Johnson vaccine administration. *Vaccine X* 2023 Dec;15:100373. [doi: [10.1016/j.jvacx.2023.100373](https://doi.org/10.1016/j.jvacx.2023.100373)] [Medline: [37674932](https://pubmed.ncbi.nlm.nih.gov/37674932/)]
10. Reichmuth ML, Heron L, Riou J, Moser A, Hauser A, Low N, et al. Socio-demographic characteristics associated with COVID-19 vaccination uptake in Switzerland: longitudinal analysis of the CoMix study. *BMC Public Health* 2023 Aug 10;23(1):1523 [FREE Full text] [doi: [10.1186/s12889-023-16405-0](https://doi.org/10.1186/s12889-023-16405-0)] [Medline: [37563550](https://pubmed.ncbi.nlm.nih.gov/37563550/)]
11. Ali F, Kaura A, Russell C, Bonn M, Bruneau J, Dasgupta N, et al. Identifying barriers and facilitators to COVID-19 vaccination uptake among people who use drugs in Canada: a national qualitative study. *Harm Reduct J* 2023 Jul 29;20(1):99 [FREE Full text] [doi: [10.1186/s12954-023-00826-6](https://doi.org/10.1186/s12954-023-00826-6)] [Medline: [37516836](https://pubmed.ncbi.nlm.nih.gov/37516836/)]
12. Blahut R, Flint A, Orlando E, DesChatelets J, Khowaja A. A scoping review on the decision-making dynamics for accepting or refusing the COVID-19 vaccination among adolescent and youth populations. *BMC Public Health* 2023 Apr 28;23(1):784 [FREE Full text] [doi: [10.1186/s12889-023-15717-5](https://doi.org/10.1186/s12889-023-15717-5)] [Medline: [37118794](https://pubmed.ncbi.nlm.nih.gov/37118794/)]
13. Soleimanpour H, Sarbazi E, Esmaeili ED, Mehri A, Fam SG, Nikbakht H, et al. Predictors of receiving COVID-19 vaccine among adult population in Iran: an observational study. *BMC Public Health* 2023 Mar 14;23(1):490 [FREE Full text] [doi: [10.1186/s12889-023-15409-0](https://doi.org/10.1186/s12889-023-15409-0)] [Medline: [36918858](https://pubmed.ncbi.nlm.nih.gov/36918858/)]
14. Yadav H, Sagar M. Exploring COVID-19 vaccine hesitancy and behavioral themes using social media big-data: a text mining approach. *K* 2023 Jun 09;52(7):2616-2648. [doi: [10.1108/k-06-2022-0810](https://doi.org/10.1108/k-06-2022-0810)]
15. Dickson K, Aboltins C, Pelly J, Jessup RL. Effective communication of COVID-19 vaccine information to recently-arrived culturally and linguistically diverse communities from the perspective of community engagement and partnership organisations: a qualitative study. *BMC Health Serv Res* 2023 Aug 21;23(1):877 [FREE Full text] [doi: [10.1186/s12913-023-09836-3](https://doi.org/10.1186/s12913-023-09836-3)] [Medline: [37605184](https://pubmed.ncbi.nlm.nih.gov/37605184/)]
16. Park M, Park J, Kim H, Lee JH, Park H. Measuring the impacts of quantity and trustworthiness of information on COVID - 19 vaccination intent. *Asso for Info Science & Tech* 2023 Apr 17;74(7):846-865 [FREE Full text] [doi: [10.1002/asi.24760](https://doi.org/10.1002/asi.24760)]
17. Pierz AJ, Rauh L, Masoud D, Cruz AK, Palmedo PC, Ratzan SC, et al. Supporting US healthcare providers for successful vaccine communication. *BMC Health Serv Res* 2023 May 02;23(1):423 [FREE Full text] [doi: [10.1186/s12913-023-09348-0](https://doi.org/10.1186/s12913-023-09348-0)] [Medline: [37131261](https://pubmed.ncbi.nlm.nih.gov/37131261/)]
18. Cmeciu C. (De)legitimation of COVID-19 vaccination narratives on Facebook comments in Romania: Beyond the co-occurrence patterns of discursive strategies. *Discourse Soc* 2023 May 26;4(5-29):09579265231174793 [FREE Full text] [doi: [10.1177/09579265231174793](https://doi.org/10.1177/09579265231174793)] [Medline: [37831753](https://pubmed.ncbi.nlm.nih.gov/37831753/)]
19. Athey S, Grabarz K, Luca M, Wernerfelt N. Digital public health interventions at scale: The impact of social media advertising on beliefs and outcomes related to COVID vaccines. *Proc Natl Acad Sci U S A* 2023 Jan 31;120(5):e2208110120 [FREE Full text] [doi: [10.1073/pnas.2208110120](https://doi.org/10.1073/pnas.2208110120)] [Medline: [36701366](https://pubmed.ncbi.nlm.nih.gov/36701366/)]
20. Benati I, Coccia M. Global analysis of timely COVID-19 vaccinations: improving governance to reinforce response policies for pandemic crises. *IJHG* 2022 May 31;27(3):240-253 [FREE Full text] [doi: [10.1108/ijhg-07-2021-0072](https://doi.org/10.1108/ijhg-07-2021-0072)]
21. Han Q, Zheng B, Cristea M, Agostini M, Bélanger JJ, Gützkow B, PsyCorona Collaboration, et al. Trust in government regarding COVID-19 and its associations with preventive health behaviour and prosocial behaviour during the pandemic:

- a cross-sectional and longitudinal study. *Psychol Med* 2023 Jan;53(1):149-159 [FREE Full text] [doi: [10.1017/S0033291721001306](https://doi.org/10.1017/S0033291721001306)] [Medline: [33769242](https://pubmed.ncbi.nlm.nih.gov/33769242/)]
22. Liu J, Shahab Y, Hoque H. Government response measures and public trust during the COVID - 19 pandemic: evidence from around the world. *British J of Management* 2021 Dec 23;33(2):571-602 [FREE Full text] [doi: [10.1111/1467-8551.12577](https://doi.org/10.1111/1467-8551.12577)]
 23. Asia Pacific Immunization Coalition, L.E.K. Consulting. Best Practices for Vaccine Communications: Lessons Learned on Preparedness and Policy Implications From India, Indonesia and the Philippines. National University of Singapore. 2022 Jan. URL: https://medicine.nus.edu.sg/apic/wp-content/uploads/sites/28/2022/01/VF_Vaccine-Communication-Preparedness-Special-report.pdf [accessed 2024-02-10]
 24. Zheng H, Jiang S, Rosenthal S. Linking online vaccine information seeking to vaccination intention in the context of the COVID-19 pandemic. *Science Communication* 2022 Jun 06;44(3):320-346 [FREE Full text] [doi: [10.1177/10755470221101067](https://doi.org/10.1177/10755470221101067)]
 25. Ung COL, Hu Y, Hu H, Bian Y. Investigating the intention to receive the COVID-19 vaccination in Macao: implications for vaccination strategies. *BMC Infect Dis* 2022 Mar 04;22(1):218 [FREE Full text] [doi: [10.1186/s12879-022-07191-y](https://doi.org/10.1186/s12879-022-07191-y)] [Medline: [35246072](https://pubmed.ncbi.nlm.nih.gov/35246072/)]
 26. Xiong W, Peng L, Tsang TK, Cowling BJ. Epidemiology of SARS-CoV-2 Omicron BA.5 infections, Macau, June-July 2022. *Emerg Infect Dis* 2023 Feb;29(2):453-456 [FREE Full text] [doi: [10.3201/eid2902.221243](https://doi.org/10.3201/eid2902.221243)] [Medline: [36648122](https://pubmed.ncbi.nlm.nih.gov/36648122/)]
 27. Zhao Y, Xu J, Bai W, Sun H, Shui B, Yang Z, et al. COVID-19 prevention and control strategies: learning from the Macau model. *Int J Biol Sci* 2022;18(14):5317-5328 [FREE Full text] [doi: [10.7150/ijbs.70177](https://doi.org/10.7150/ijbs.70177)] [Medline: [36147478](https://pubmed.ncbi.nlm.nih.gov/36147478/)]
 28. Dai Y, Li Y, Cheng C, Zhao H, Meng T. Government-led or public-led? Chinese policy agenda setting during the COVID-19 pandemic. *Journal of Comparative Policy Analysis: Research and Practice* 2021 Apr 21;23(2):157-175 [FREE Full text] [doi: [10.1080/13876988.2021.1878887](https://doi.org/10.1080/13876988.2021.1878887)]
 29. McCombs ME, Shaw DL. The agenda-setting function of mass media. *Public Opinion Quarterly* 1972;36(2):176-187 [FREE Full text] [doi: [10.1086/267990](https://doi.org/10.1086/267990)]
 30. Rogers EM, Dearing JW, Bregman D. The anatomy of agenda-setting research. *Journal of Communication* 1993;43(2):68-84. [doi: [10.1111/j.1460-2466.1993.tb01263.x](https://doi.org/10.1111/j.1460-2466.1993.tb01263.x)]
 31. Guo L. The application of social network analysis in agenda setting research: a methodological exploration. *Journal of Broadcasting & Electronic Media* 2012 Oct;56(4):616-631 [FREE Full text] [doi: [10.1080/08838151.2012.732148](https://doi.org/10.1080/08838151.2012.732148)]
 32. McCombs ME, Guo L. Agenda-Setting Influence of the Media in the Public Sphere. In: Forner RS, Fackler PM, editors. *The Handbook of Media and Mass Communication Theory*. Hoboken, NJ: John Wiley & Sons, Inc; 2014.
 33. Chen Z, Su CC, Chen A. Top-down or bottom-up? A network agenda-setting study of Chinese nationalism on social media. *Journal of Broadcasting & Electronic Media* 2019 Sep 20;63(3):512-533 [FREE Full text] [doi: [10.1080/08838151.2019.1653104](https://doi.org/10.1080/08838151.2019.1653104)]
 34. Hou Z, Tong Y, Du F, Lu L, Zhao S, Yu K, et al. Assessing COVID-19 vaccine hesitancy, confidence, and public engagement: a global social listening study. *J Med Internet Res* 2021 Jun 11;23(6):e27632 [FREE Full text] [doi: [10.2196/27632](https://doi.org/10.2196/27632)] [Medline: [34061757](https://pubmed.ncbi.nlm.nih.gov/34061757/)]
 35. Buller D, Walkosz B, Henry K, Woodall WG, Pagoto S, Berteletti J, et al. Promoting social distancing and COVID-19 vaccine intentions to mothers: randomized comparison of information sources in social media messages. *JMIR Infodemiology* 2022 Aug 23;2(2):e36210 [FREE Full text] [doi: [10.2196/36210](https://doi.org/10.2196/36210)] [Medline: [36039372](https://pubmed.ncbi.nlm.nih.gov/36039372/)]
 36. Albalawi Y, Sixsmith J. Agenda setting for health promotion: exploring an adapted model for the social media era. *JMIR Public Health Surveill* 2015;1(2):e21 [FREE Full text] [doi: [10.2196/publichealth.5014](https://doi.org/10.2196/publichealth.5014)] [Medline: [27227139](https://pubmed.ncbi.nlm.nih.gov/27227139/)]
 37. Chang A, Xian X, Liu MT, Zhao X. Health communication through positive and solidarity messages amid the COVID-19 pandemic: automated content analysis of Facebook uses. *Int J Environ Res Public Health* 2022 May 19;19(10):6159 [FREE Full text] [doi: [10.3390/ijerph19106159](https://doi.org/10.3390/ijerph19106159)] [Medline: [35627696](https://pubmed.ncbi.nlm.nih.gov/35627696/)]
 38. Social Media Stats Macaots. Statcounter. URL: <https://gs.statcounter.com/social-media-stats/all/macao> [accessed 2024-02-10]
 39. Della Rosa S, Sen F. Health topics on Facebook groups: content analysis of posts in multiple sclerosis communities. *Interact J Med Res* 2019 Feb 11;8(1):e10146 [FREE Full text] [doi: [10.2196/10146](https://doi.org/10.2196/10146)] [Medline: [30741640](https://pubmed.ncbi.nlm.nih.gov/30741640/)]
 40. Tahamtan I, Potnis D, Mohammadi E, Singh V, Miller LE. The mutual influence of the World Health Organization (WHO) and Twitter users during COVID-19: network agenda-setting analysis. *J Med Internet Res* 2022 Apr 26;24(4):e34321 [FREE Full text] [doi: [10.2196/34321](https://doi.org/10.2196/34321)] [Medline: [35275836](https://pubmed.ncbi.nlm.nih.gov/35275836/)]
 41. Holt K, Ustad Figenschou T, Frischlich L. Key dimensions of alternative news media. *Digital Journalism* 2019 Jun 13;7(7):860-869 [FREE Full text] [doi: [10.1080/21670811.2019.1625715](https://doi.org/10.1080/21670811.2019.1625715)]
 42. Lin ZX, Liu S. [The change of media ecology and online news practice in Macao]. *Ershiyi Shi Ji* 2019;176:95-112 [FREE Full text]
 43. Knabe A, McCarthy M. Civil society organisations and public health research--evidence from eight European union new member states. *Cent Eur J Public Health* 2012 Dec;20(4):287-293 [FREE Full text] [doi: [10.21101/cejph.a3764](https://doi.org/10.21101/cejph.a3764)] [Medline: [23441396](https://pubmed.ncbi.nlm.nih.gov/23441396/)]

44. Zhou S, Zheng X. Agenda dynamics on social media during COVID-19 pandemic: interactions between public, media, and government agendas. *Communication Studies* 2022 Jun 10;73(3):211-228 [[FREE Full text](#)] [doi: [10.1080/10510974.2022.2082504](https://doi.org/10.1080/10510974.2022.2082504)]
45. Chen L, Ling Q, Cao T, Han K. Mislabeled, fragmented, and conspiracy-driven: a content analysis of the social media discourse about the HPV vaccine in China. *Asian Journal of Communication* 2020 Sep 08;30(6):450-469 [[FREE Full text](#)] [doi: [10.1080/01292986.2020.1817113](https://doi.org/10.1080/01292986.2020.1817113)]
46. Li J, Zheng H. Coverage of HPV-related information on Chinese social media: a content analysis of articles in Zhihu. *Hum Vaccin Immunother* 2020 Oct 02;16(10):2548-2554 [[FREE Full text](#)] [doi: [10.1080/21645515.2020.1729028](https://doi.org/10.1080/21645515.2020.1729028)] [Medline: [32159420](https://pubmed.ncbi.nlm.nih.gov/32159420/)]
47. Thavareesan S, Mahesan S. Sentiment Lexicon Expansion using Word2vec and fastText for Sentiment Prediction in Tamil texts. 2020 Presented at: Moratuwa Engineering Research Conference (MERCCon); July 28-30, 2020; Moratuwa, Sri Lanka URL: <https://ieeexplore.ieee.org/document/9185369> [doi: [10.1109/MERCCon50084.2020.9185369](https://doi.org/10.1109/MERCCon50084.2020.9185369)]
48. Orkphol K, Yang W. Word sense disambiguation using cosine similarity collaborates with Word2vec and WordNet. *Future Internet* 2019 May 12;11(5):114 [[FREE Full text](#)] [doi: [10.3390/fi11050114](https://doi.org/10.3390/fi11050114)]
49. Zhou F, Zhang W, Cai H, Cao Y. Portrayals of 2v, 4v and 9vHPV vaccines on Chinese social media: a content analysis of hot posts on Sina Weibo. *Hum Vaccin Immunother* 2021 Nov 02;17(11):4433-4441 [[FREE Full text](#)] [doi: [10.1080/21645515.2021.1971016](https://doi.org/10.1080/21645515.2021.1971016)] [Medline: [34543155](https://pubmed.ncbi.nlm.nih.gov/34543155/)]
50. Chang A, Schulz PJ, Jiao W, Liu MT. Obesity-related communication in digital Chinese news from mainland China, Hong Kong, and Taiwan: automated content analysis. *JMIR Public Health Surveill* 2021 Nov 23;7(11):e26660 [[FREE Full text](#)] [doi: [10.2196/26660](https://doi.org/10.2196/26660)] [Medline: [34817383](https://pubmed.ncbi.nlm.nih.gov/34817383/)]
51. Chang A, Schulz PJ, Tu S, Liu MT. Communicative blame in online communication of the COVID-19 pandemic: computational approach of stigmatizing cues and negative sentiment gauged with automated analytic techniques. *J Med Internet Res* 2020 Nov 25;22(11):e21504 [[FREE Full text](#)] [doi: [10.2196/21504](https://doi.org/10.2196/21504)] [Medline: [33108306](https://pubmed.ncbi.nlm.nih.gov/33108306/)]
52. Wang D, Zhou Y, Ma F. Opinion leaders and structural hole spanners influencing echo chambers in discussions about COVID-19 vaccines on social media in China: network analysis. *J Med Internet Res* 2022 Nov 18;24(11):e40701 [[FREE Full text](#)] [doi: [10.2196/40701](https://doi.org/10.2196/40701)] [Medline: [36367965](https://pubmed.ncbi.nlm.nih.gov/36367965/)]
53. Feicheng M, Yating L. Utilising social network analysis to study the characteristics and functions of the co-occurrence network of online tags. *Online Information Review* 2014;38(2):232-247. [doi: [10.1108/OIR-11-2012-0124](https://doi.org/10.1108/OIR-11-2012-0124)]
54. Wang X, Chen L, Shi J, Tang H. Who sets the agenda? the dynamic agenda setting of the wildlife issue on social media. *Environmental Communication* 2021 Apr 12;17(3):245-262 [[FREE Full text](#)] [doi: [10.1080/17524032.2021.1901760](https://doi.org/10.1080/17524032.2021.1901760)]
55. Vu H, Guo L, McCombs ME. Exploring “the World Outside and the Pictures in Our Heads”. *Journalism & Mass Communication Quarterly* 2014 Sep 09;91(4):669-686 [[FREE Full text](#)] [doi: [10.1177/1077699014550090](https://doi.org/10.1177/1077699014550090)]
56. Feng N, Feng H, Li D, Li M. Online media coverage, consumer engagement and movie sales: A PVAR approach. *Decision Support Systems* 2020 Apr;131:113267 [[FREE Full text](#)] [doi: [10.1016/j.dss.2020.113267](https://doi.org/10.1016/j.dss.2020.113267)]
57. Lorenz N, Sander C, Ivanova G, Hegerl U. Temporal associations of daily changes in sleep and depression core symptoms in patients suffering from major depressive disorder: idiographic time-series analysis. *JMIR Ment Health* 2020 Apr 23;7(4):e17071 [[FREE Full text](#)] [doi: [10.2196/17071](https://doi.org/10.2196/17071)] [Medline: [32324147](https://pubmed.ncbi.nlm.nih.gov/32324147/)]
58. Yang H, Wang S, Ren Z, Liu H, Tong Y, Wang N. Life expectancy, air pollution, and socioeconomic factors: a multivariate time-series analysis of Beijing City, China. *Soc Indic Res* 2022 Jan 03;162(3):979-994 [[FREE Full text](#)] [doi: [10.1007/s11205-021-02872-8](https://doi.org/10.1007/s11205-021-02872-8)]
59. Golovko D, Schumann JH. Influence of company Facebook activities on recruitment success. *Journal of Business Research* 2019 Nov;104:161-169 [[FREE Full text](#)] [doi: [10.1016/j.jbusres.2019.06.029](https://doi.org/10.1016/j.jbusres.2019.06.029)]
60. Geiß S. The media’s conditional agenda-setting power: how baselines and spikes of issue salience affect likelihood and strength of agenda-setting. *Communication Research* 2019 Sep 19;49(2):296-323 [[FREE Full text](#)] [doi: [10.1177/0093650219874968](https://doi.org/10.1177/0093650219874968)]
61. Widmann T. Fear, hope, and COVID-19: emotional elite rhetoric and its impact on the public during the first wave of the COVID-19 pandemic. *Polit Psychol* 2022 May 09;43(5):827-850 [[FREE Full text](#)] [doi: [10.1111/pops.12831](https://doi.org/10.1111/pops.12831)] [Medline: [35941918](https://pubmed.ncbi.nlm.nih.gov/35941918/)]
62. Barberá P, Casas A, Nagler J, Egan PJ, Bonneau R, Jost JT, et al. Who leads? Who follows? Measuring issue attention and agenda setting by legislators and the mass public using social media data. *Am Polit Sci Rev* 2019 Jul 12;113(4):883-901 [[FREE Full text](#)] [doi: [10.1017/S0003055419000352](https://doi.org/10.1017/S0003055419000352)] [Medline: [33303996](https://pubmed.ncbi.nlm.nih.gov/33303996/)]
63. Finset A, Bosworth H, Butow P, Gulbrandsen P, Hulsman RL, Pieterse AH, et al. Effective health communication - a key factor in fighting the COVID-19 pandemic. *Patient Educ Couns* 2020 May;103(5):873-876 [[FREE Full text](#)] [doi: [10.1016/j.pec.2020.03.027](https://doi.org/10.1016/j.pec.2020.03.027)] [Medline: [32336348](https://pubmed.ncbi.nlm.nih.gov/32336348/)]
64. Luo Y. The Internet and agenda setting in China: The influence of online public opinion on media coverage and government policy. *International Journal of Communication* 2014;8:1289-1312 [[FREE Full text](#)]
65. Gruszczynski M, Wagner MW. Information flow in the 21st century: the dynamics of agenda-uptake. *Mass Communication and Society* 2016 Nov 18;20(3):378-402 [[FREE Full text](#)] [doi: [10.1080/15205436.2016.1255757](https://doi.org/10.1080/15205436.2016.1255757)]

66. Kim Y. Contingent factors of agenda-setting effects: How need for orientation, issue obtrusiveness, message tone influence issue salience attitude strength. In: Johnson TJ, editor. Agenda setting in a 2.0 world. New York, NY: Routledge; 2013:65-81.

Abbreviations

NAS: network agenda setting model
QAP: quadratic assignment procedure
RQ: research question
VAR: vector autoregression

Edited by W Ahmed; submitted 21.07.23; peer-reviewed by A Staffini, M Coccia; comments to author 29.09.23; revised version received 19.11.23; accepted 29.11.23; published 19.03.24.

Please cite as:

Xian X, Neuwirth RJ, Chang A

Government-Nongovernmental Organization (NGO) Collaboration in Macao's COVID-19 Vaccine Promotion: Social Media Case Study

JMIR Infodemiology 2024;4:e51113

URL: <https://infodemiology.jmir.org/2024/1/e51113>

doi: [10.2196/51113](https://doi.org/10.2196/51113)

PMID: [38502184](https://pubmed.ncbi.nlm.nih.gov/38502184/)

©Xuechang Xian, Rostam J Neuwirth, Angela Chang. Originally published in JMIR Infodemiology (<https://infodemiology.jmir.org>), 19.03.2024. 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.

Original Paper

Development of a Medical Social Media Ethics Scale and Assessment of #IRad, #CardioTwitter, and #MedTwitter Posts: Mixed Methods Study

Vongai Christine Mlambo¹, BSc; Eric Keller², MA, MD; Caroline Mussatto³, BSc, MD; Gloria Hwang², MD

¹School of Medicine, Stanford University, Palo Alto, CA, United States

²Division of Interventional Radiology, Stanford Health Care, Stanford, CA, United States

³School of Medicine, University of Kansas Medical Center, Kansas City, KS, United States

Corresponding Author:

Vongai Christine Mlambo, BSc

School of Medicine

Stanford University

751 Layne Court

Apartment 25

Palo Alto, CA, 94306

United States

Phone: 1 6506605737

Email: vmlambo@stanford.edu

Abstract

Background: Social media posts by clinicians are not bound by the same rules as peer-reviewed publications, raising ethical concerns that have not been extensively characterized or quantified.

Objective: We aim to develop a scale to assess ethical issues on medical social media (SoMe) and use it to determine the prevalence of these issues among posts with 3 different hashtags: #MedTwitter, #IRad, and #CardioTwitter.

Methods: A scale was developed based on previous descriptions of professionalism and validated via semistructured cognitive interviewing with a sample of 11 clinicians and trainees, interrater agreement, and correlation of 100 posts. The final scale assessed social media posts in 6 domains. This was used to analyze 1500 Twitter posts, 500 each from the 3 hashtags. Analysis of posts was limited to original Twitter posts in English made by health care professionals in North America. The prevalence of potential issues was determined using descriptive statistics and compared across hashtags using the Fisher exact and χ^2 tests with Yates correction.

Results: The final scale was considered reflective of potential ethical issues of SoMe by participants. There was good interrater agreement (Cohen $\kappa=0.620$, $P<.01$) and moderate to strong positive interrater correlation ($=0.602$, $P<.001$). The 6 scale domains showed minimal to no interrelation (Cronbach $\alpha=0.206$). Ethical concerns across all hashtags had a prevalence of 1.5% or less except the conflict of interest concerns on #IRad, which had a prevalence of 3.6% ($n=18$). Compared to #MedTwitter, posts with specialty-specific hashtags had more patient privacy and conflict of interest concerns.

Conclusions: The SoMe professionalism scale we developed reliably reflects potential ethical issues. Ethical issues on SoMe are rare but important and vary in prevalence across medical communities.

(*JMIR Infodemiology* 2024;4:e47770) doi:[10.2196/47770](https://doi.org/10.2196/47770)

KEYWORDS

ethics; social media; conflict of interest; interventional radiology; X; Twitter; cardiology; privacy; ethical issues; medical social media; prevalence; professional; professionalism

Introduction

The digital footprint of clinicians on social media has increased over the past 10 years with an estimated 90% and 65% of

clinicians using social media for personal and professional purposes, respectively [1]. Medical social media (SoMe) has blossomed, offering clinicians opportunities to collaborate across distances, debate treatment approaches for challenging cases,

and engage in public health advocacy [2-4]. However, this rapid integration of social media in health care has outpaced guidance that counsels on how to avoid ethical concerns that can occur with SoMe [2].

The risks of SoMe have not gone unnoticed. Several professional organizations have released statements outlining guiding principles for online clinician behavior, including the American Medical Association and the Federation of State Medical Boards [5,6]. There have also been opinion pieces and recommendations published within various specialties such as neurology, dermatology, and vascular surgery [7-9]. Guidelines and opinion pieces are helpful starting points but may not address subtle but important breaches in professionalism [10] and may fail to resonate with the majority of users' experiences and values [2].

A few studies have assessed the prevalence of issues such as violations of the HIPAA (Health Insurance Portability and Accountability Act) [10]. However, the potential issues are much broader than explicit patient privacy violations [10,11]. This study sought to develop a more complete scale of ethical issues related to medical SoMe to provide empirical data on these issues. The authors hypothesized that a scale could be developed that captures the most salient ethical issues with good interrater agreement and correlation. The authors also hypothesized that applying such a scale would find that the prevalence of issues was small and varied across different professional groups.

Methods

Scale Development

This study was approved by the Stanford University Institutional Review Board (eProtocol 60351). An initial draft of the scale

was developed based on medical professionalism in the new millennium: a physician charter created by the American Board of Internal Medicine Foundation, American College of Physicians Foundation, and the European Federation of Internal Medicine as well as a study by Chandratilake et al [12] assessing definitions of medical professionalism across cultures [13]. These sources were selected to attempt to define medical SoMe ethics that would be reflective of common definitions of medical professionalism. The initial draft consisted of 5 criteria rated on a 3-point scale: no ethical concern (0), potential ethical concern (1), and clear ethical concern (2). The 3-point scale was selected to reflect a concept raised by both initial sources that ethical issues occur on a continuum, allowing the scale to also capture less overt violations of professionalism.

The initial scale was then vetted for validity via semistructured cognitive interviewing with a group of clinicians and trainees [14]. Interviewees were recruited via email and were primarily a convenience sample at the authors' institutions. They were invited to provide feedback on a draft of the scale, which included fabricated posts and example scoring for demonstration. Purposeful recruiting was used to ensure that interviewees were diverse in terms of specialty, training level, and gender identity. Iterative adjustments were made to the initial scale based on interviewee feedback until additional interviews continued suggesting that the scale was reflective of interviewee perceptions of potential ethical issues related to medical SoMe. This occurred after 11 interviews with interviewees from 6 different specialties whose demographics are shown in [Table 1](#).

Table 1. Demographic characteristics of interviewees (N=11).

Characteristic	Interviewees, n (%)
Training level	
1st-year MD ^a	2 (18)
2nd-year MD	0 (0)
3rd-year MD	2 (18)
4th+ year MD	1 (9)
1st-year resident	0 (0)
2nd-year resident	1 (9)
3rd+ year resident	1 (9)
Attending	4 (36)
Institution	
Stanford University School of Medicine	8 (73)
University of California San Diego	1 (9)
University of Kansas Medical Center	2 (18)
Specialty	
Anesthesiology	2 (18)
DR ^b /IR ^c	1 (9)
Emergency medicine	1 (9)
Primary care	1 (9)
Psychiatry	1 (9)
Otolaryngology	1 (9)
Undeclared	4 (36)
Sex	
Female	7 (64)
Male	4 (36)

^aMD: Doctor of Medicine.

^bDR: Diagnostic Radiology.

^cIR: Interventional Radiology.

The vetted scale scored posts on 6 domains, using the same 3-point scale (Table 2). Scale item interrelation as well as scale interrater agreement and correlation were assessed by having 2 researchers use the scale to independently rate 50 random posts each from #MedTwitter between June 15, 2021, and August 15, 2021, with an overlap of 10 tweets. Posts were identified

using the Healthcare Hashtag Project (Symplur, LLC). The interrelation of scale items was assessed via Cronbach α . Interrater agreement was assessed via Cohen κ and interrater correlation was assessed via Spearman correlation coefficient, assuming a nonlinear relationship. An α of $<.05$ was predefined as statistical significance.

Table 2. Medical social media professionalism scale.

Principle	Score		
	0=no concern	1=minor concern	2=major concern
Patient privacy			
Does the post maintain patient privacy by applying appropriate safeguards for patient information and removing patient identifiers?	Post omits HIPAA ^a identifiers and any other details that in combination would enable patient identification.	Post omits HIPAA identifiers but uses information that could potentially allow for patient identification, particularly when combined with the author's known practice location, medical specialty, or rarity of medical condition.	Post uses one or more HIPAA identifiers that allows for easy identification.
Patient dignity			
Does the post treat patients with respect and avoid the use of degrading language or images?	Post treats patients as individuals worthy of respect and does not demean the patient in any way.	Post contains references, images, or language that could be negatively construed such that some may take offense.	Post is objectifying or dehumanizing, treating patients as being of lesser intelligence or caliber.
Information accuracy			
Is the information medically accurate with no counterfactual, exaggerated, or otherwise misleading content?	Information in the post is reasonably supported by current evidence and does not make superlative claims.	Information in the post is ambiguous or exaggerated in a manner that could lead to misinterpretation.	Information in the post is overtly sensational and makes baseless claims.
Conflict of interest			
Is the post unduly influenced by ulterior motives for private gain without proper acknowledgment or disclosure in a way that could affect information accuracy?	The post does not promote or endorse products or services without an appropriate declaration of any associated financial ties.	The post promotes or endorses products or services without a declaration of conflicts, however, it does not make authoritative claims about these products.	The post promotes or endorses products or services without a proper declaration of conflicts and also makes authoritative claims about these products.
Justice and equity			
Is the text or images in the post discriminatory based on race, gender, socioeconomic status, ethnicity, religion, sexual orientation, or any other social category and does the post promote further inequities in health care?	The post does not express or imply any discriminatory sentiments or propagate a stance that either sustains or widens inequities in health care.	The post contains ideas associated with stereotypes or broad generalizations <i>without</i> suggesting the differential treatment of individuals based on these stereotypes.	The post explicitly expresses sentiments that are discriminatory and is a proponent for the differential treatment of individuals based on these prejudiced notions.
Interprofessional respect			
Does the post treat colleagues and other health care professionals with respect and avoid the use of stereotypes, mockery, and incivility?	Post treats colleagues and other health care professionals with esteem and does not demean them in any way.	Post contains references, images, or language that could be negatively construed by other colleagues as offensive.	Post clearly mocks or disrespects colleagues, portraying them as inferior or of lesser intelligence or caliber.

^aHIPAA: Health Insurance Portability and Accountability Act.

Evaluation of Posts

The validated scale was then used to assess the prevalence of ethical issues among posts using 3 distinct hashtags: #MedTwitter, #IRad, and #CardioTwitter. These were selected as they are the most frequently used hashtags among the general medical community, interventional radiologists, and cardiologists, respectively, as indicated by the number of posts per day for each hashtag on the Symplur software. Interventional Radiology (IR) and cardiology were selected to provide examples of more specialty-specific posts to contrast with #MedTwitter as they are primarily used by physician specialists in those fields to discuss more expert medical content compared to #MedTwitter. Posts were limited to those in English posted by individuals (rather than societies or bots) who are clinicians or health care trainees in North America between December 10, 2021, and January 10, 2022. Retweets were also excluded.

A total of 1500 posts were analyzed, 500 from each hashtag. Data were analyzed using descriptive statistics as well as Fisher exact tests and χ^2 tests with Yates correction to compare the prevalence of ethical issues across hashtags. These statistical tests were selected to adjust for the low rates of ethical issues. All statistical analyses were performed using SPSS software (IBM, Inc).

Ethical Considerations

All procedures were approved by the Stanford University Institutional Review Board (IRB#: 60351) and were per the legal and ethical standards of the responsible committee on human experimentation institutionally. Additionally, we adhered to local, national, regional, and international laws and regulations regarding the protection of personal information, privacy, and human rights.

Results

Scale Development

Cognitive interviewing supported the validity of the initial 5 domains. However, the initial interviewees felt the initial scale did not address interspecialty and inter-health care professional cyberbullying, leading to the addition of interprofessional respect as a 6th domain. Interviewees also suggested the addition of language to better delineate a minor concern (1) rating from a major concern (2) rating. Subsequent interviews confirmed that the 6-domain scale, each rated from 0 to 2, was reflective of their perceptions of SoMe ethics.

The scale demonstrated good interrater agreement (Cohen $\kappa=0.620$, $P<.01$) and moderate to strong positive correlation between the scores given by the independent raters (Spearman correlation coefficient=0.602, 95% CI 0.515-0.677; $P<.001$). The scale domains showed minimal to no interrelation (Cronbach $\alpha=0.206$).

Evaluation of Posts

Application of the scale to 1500 Twitter posts showed that ethical concerns across all 6 domains were infrequent with the majority in the range of 0.2% ($n=1$) to 1.2% ($n=6$). Further, 1 exception was a minor conflict of interest concern among posts using #IRad, which demonstrated a prevalence of 3.6% ($n=18$). Relative to posts using #MedTwitter, posts using #IRad or #CardioTwitter were more likely to have patient privacy concerns ($n=7$, 1.4% vs 0%, $P=.02$; $n=6$, 1.2% vs 0%, $P=.04$; respectively). Posts using #IRad were also more likely to have conflicts of interest concerns relative to #MedTwitter and #CardioTwitter ($n=18$, 3.6% vs $n=3$, 0.6%, $P<.001$; $n=18$, 3.6% vs $n=4$, 0.8%, $P=.005$; respectively). Issues related to interprofessional respect were also more prevalent in #IRad posts than #CardioTwitter ($n=8$, 1.6% vs $n=1$, 0.2%, $P=.04$) but similar to #MedTwitter ($n=8$, 1.6% vs $n=6$, 1.2%, $P=.79$). As a result, across all domains, #IRad posts had the greatest overall prevalence of ethical concerns. Table 3 summarizes the prevalence of ethical concerns by hashtag and domain and Tables 4-6 summarize comparisons between hashtags.

Table 3. Prevalence of ethical concerns on medical social media by hashtag (N=500).

	No issue (0), n (%)	Minor concern (1), n (%)	Major concern (2), n (%)
MedTwitter prevalence			
Patient privacy	500 (100)	0 (0)	0 (0)
Patient dignity	495 (99)	3 (0.6)	2 (0.4)
Information accuracy	497 (99.4)	2 (0.4)	1 (0.2)
Conflict of interest	500 (100)	0 (0)	0 (0)
Justice and equity	499 (99.8)	1 (0.2)	0 (0)
Interprofessional respect	494 (98.8)	4 (0.8)	2 (0.4)
IR^a prevalence			
Patient privacy	493 (98.6)	6 (1.2)	1 (0.2)
Patient dignity	497 (99.4)	1 (0.2)	2 (0.4)
Information accuracy	497 (99.4)	2 (0.4)	1 (0.2)
Conflict of interest	482 (96.4)	18 (3.6)	0 (0)
Justice and equity	500 (100)	0 (0)	0 (0)
Interprofessional respect	492 (98.4)	7 (1.4)	1 (0.2)
Cardiology prevalence			
Patient privacy	494 (98.8)	6 (1.2)	0 (0)
Patient dignity	499 (99.8)	1 (0.2)	0 (0)
Information accuracy	500 (100)	0 (0)	0 (0)
Conflict of interest	496 (99.2)	2 (0.4)	2 (0.4)
Justice and equity	500 (100)	0 (0)	0 (0)
Interprofessional respect	499 (99.8)	1 (0.2)	0 (0)

^aIR: Interventional Radiology.

Table 4. Comparison of ethical concerns on medical social media by hashtag^a: #IRad vs #MedTwitter.”

	#IRad, n (%)	#MedTwitter, n (%)	Fisher exact <i>P</i> value	Chi-squared with Yates correction <i>P</i> value
#IRad vs #MedTwitter				
Patient privacy	7 (1.4) ^b	0 (0) ^b	.02 ^b	.02 ^b
Patient dignity	3 (0.6)	5 (1)	.73	.72
Information accuracy	3 (0.6)	3 (0.6)	≥.99	≥.99
Conflict of interest	18 (3.6) ^b	0 (0) ^b	<.001 ^b	<.001 ^b
Justice and equity	0 (0)	1 (0.2)	≥.99	.32
Interprofessional respect	8 (1.6)	6 (1.2)	.79	.79

^aComparisons reflect the composite of major and minor concerns for each scale criterion. $P < .05$ on a 2-tailed analysis was considered significant.

^bComparisons that are significant.

Table 5. Comparison of ethical concerns on medical social media by hashtag^a: #CardioTwitter vs #MedTwitter.”

	#CardioTwitter, n (%)	#MedTwitter, n (%)	Fisher exact <i>P</i> value	Chi-squared with Yates correction <i>P</i> value
#CardioTwitter vs MedTwitter				
Patient privacy	6 (1.2) ^b	0 (0) ^b	.03 ^b	.04 ^b
Patient dignity	1 (0.2)	5 (1)	.22	.22
Information accuracy	0 (0)	3 (0.6)	.37	.62
Conflict of interest	4 (0.8)	0 (0)	.22	.37
Justice and equity	0 (0)	1 (0.2)	≥.99	.32
Interprofessional respect	1 (0.2)	6 (1.2)	.12	.13

^aComparisons reflect the composite of major and minor concerns for each scale criterion. $P < .05$ on a 2-tailed analysis was considered significant.

^bComparisons that are significant.

Table 6. Comparison of ethical concerns on medical social media by hashtag^a: #IRad vs #CardioTwitter.”

	#IRad, n (%)	#CardioTwitter, n (%)	Fisher exact <i>P</i> value	Chi-squared with Yates correction <i>P</i> value
#IRad vs #CardioTwitter				
Patient privacy	7 (1.4)	6 (1.2)	≥.99	.78
Patient dignity	3 (0.6)	1 (0.2)	.62	.62
Information accuracy	3 (0.6)	0 (0)	.37	.62
Conflict of interest	18 (3.6) ^b	4 (0.8) ^b	.004 ^b	.005 ^b
Justice and equity	0 (0)	0 (0)	≥.99	≥.99
Interprofessional respect	8 (1.6) ^b	1 (0.2) ^b	.04 ^b	.04 ^b

^aComparisons reflect the composite of major and minor concerns for each scale criterion. $P < .05$ on a 2-tailed analysis was considered significant.

^bComparisons that are significant.

Discussion

Principal Results

This study sought to develop a scale to characterize and quantitate ethical issues on SoMe and then apply the scale to 3 different SoMe communities based on Twitter hashtags. Although some guidelines and opinion pieces exist describing potential ethical issues on SoMe, to the best of the authors'

knowledge, no scales had been created, making it difficult to assess the prevalence of ethical issues and guide efforts to mitigate potential harm [10]. This is important not only because of legal implications, but this behavior can exacerbate existing hierarchies and damage mutual trust.

The scale proposed in this study was developed via a structured deductive and inductive approach. Key domains were identified based on literature review as well as qualitative interviews,

consistent with best practices in scale development [15,16]. This helped ensure that the scale was comprehensive and perceived as valid. Interrater agreement and correlation were good but likely limited by the qualitative nature of these assessments. The lack of interrelation between domains is not unexpected. A post with a patient privacy concern would not necessarily be more likely to have a conflict of interest as well.

Application of the scale to Twitter posts with #MedTwitter, #CardioTwitter, and #IRad yielded a couple of important observations. First, the prevalence of ethical concerns is low, often around 1% (n=5) across domains. However, such a number is not insignificant. According to Symplur software, there are approximately 5000 to 8000 posts per day made using #MedTwitter, equating to approximately 50-80 ethically concerning posts per day. These findings are similar to a 2011 study of over 5000 general tweets from health care providers, which found 3% of tweets were unprofessional and 0.7% were concerning for breaches in patient privacy [17].

A second interesting observation was how the prevalence of ethical concerns varied across the 3 groups of posts analyzed. For example, posts with the specialty-specific hashtags #CardioTwitter and #IRad had more patient privacy and conflict of interest concerns than general #MedTwitter posts. This may be due to a higher likelihood of posting specific patient cases in specialty-specific communities to illustrate an approach or solicit recommendations compared to the general #MedTwitter community. Posts with conflict of interest were also most prevalent in #IRad posts, which may be due to IR being a more procedural specialty than cardiology in general, and a specialty whose professional identity is closely tied to specific procedures and devices rather than patient populations [18]. Previous authors have observed similar variations in posts across specialties. The dominating content among IR posts tends to be images of an intervention performed on a patient to share new techniques or gather recommendations for superior approaches [19]. In contrast, cardiology posts are dominated by short synopses of trending research papers with reactive commentary [20]. However, interventional cardiology posts can share similar traits to IR [20,21], likely accounting for some of the overlap in the ethical issues among these posts.

Practical Implications

The persistence of posts with ethical issues among medical professionals and trainees invites evaluation of current social media training programs. The domains in the scale offer a useful framework with validated language and examples to offer caution against ethical concerns that go beyond HIPAA violations. The framework can also foster a mental model to assist in evaluating personal tweets before publishing a post. This is important as once a post is made; it is difficult to retract it completely before it is shared or copied by other users.

The results from this study also provide a foundation for evidence-based social media guidelines by professional bodies and specialty-specific societies. As demonstrated by differences in the prevalence of ethical concerns between #CardioTwitter and #IRad, not all ethical issues are equally problematic, and with this data, guidelines can be tailored to the target group. This scale can be applied to hashtags used by other specialists

to uncover trends in ethical issues and address those weak points more specifically. For example, social media statements for interventional radiologists may include more specific and detailed guidance on avoiding conflict of interest concerns.

From an academic perspective, the scale and methodology described in this study offer a way to assess the efficacy of interventions aimed at reducing the frequency of ethical issues on SoMe. Previously, there were limited ways to quantify and characterize the landscape of SoMe professionalism. However, now it is possible to perform pre- and poststudies with a specific intervention of interest.

Although this study focused on the application of the professionalism scale to Twitter posts as a proof of concept, the principles could be translated to other platforms as they do not include any evaluation metric that is inherent to Twitter, since the development of the scale was independent of any specific platform. From a validation perspective, this translation would be easiest for platforms that mimic Twitter by using a combination of texts and images, such as Facebook and Instagram posts. Importantly, videos were not assessed in this study, which would be of interest in analyzing Reels, TikTok, and YouTube videos. However, the methodology of this study can be applied to these different social media contexts to assess the generalizability of the scale.

Limitations and Future Directions

This study had important limitations. The scale provides a good estimate of the prevalence of ethical issues, but it is not a thorough investigation of whether a given issue definitively exists especially for domains like conflict of interest that are challenging to verify without collateral information. Although the scale development incorporated input from a diverse group of clinicians and trainees in terms of training level, specialty, and gender identity, the sample was a small convenience sample from academic settings that could have missed important input from other clinicians in different contexts, for example, private practice. The sample was limited to posts in English from North America due to language restrictions and greater cultural familiarity. However, this may limit the external validity of the scale and results in other cultures. The authors relied on self-described Twitter biographies to limit posts to health care professionals, which could have been inaccurate.

To address some of these limitations, future steps to continue improving the scale would include expanding the sample to include more physicians and trainees from private practice, community hospitals, and primary care so that these additional perspectives can further refine the scale. Additionally, although the Cohen κ for interrater reliability already suggests good agreement, there may be domains with greater discrepancies than others. The language of these domains can be made more precise or explicit based on a bigger sample feedback to potentially improve consistency. Lastly, a comparison among different platforms would help directly assess if scale validity transcends social media contexts.

Conclusions

The developed SoMe ethics scale is reliable, relevant, and concisely captures the myriad ethical tensions that can arise on

these platforms. Ethical issues are present in a small but meaningful percentage of posts among health care professionals, which vary in important ways across different specialties and professional groups. The authors hope this scale will allow

researchers to better characterize and assess the prevalence of ethical issues on SoMe while guiding more targeted interventions to mitigate these issues.

Conflicts of Interest

None declared.

References

1. Ventola CL. Social media and health care professionals: benefits, risks, and best practices. *P T* 2014;39(7):491-520 [FREE Full text] [Medline: 25083128]
2. Panahi S, Watson J, Partridge H. Social media and physicians: exploring the benefits and challenges. *Health Informatics J* 2016;22(2):99-112 [FREE Full text] [doi: 10.1177/1460458214540907] [Medline: 25038200]
3. McGowan BS, Wasko M, Vartabedian BS, Miller RS, Freiherr DD, Abdolrasulnia M. Understanding the factors that influence the adoption and meaningful use of social media by physicians to share medical information. *J Med Internet Res* 2012;14(5):e117 [FREE Full text] [doi: 10.2196/jmir.2138] [Medline: 23006336]
4. O'Glasser AY, Jaffe RC, Brooks M. To tweet or not to tweet, that is the question. *Semin Nephrol* 2020;40(3):249-263. [doi: 10.1016/j.semnephrol.2020.04.003] [Medline: 32560773]
5. Farnan JM, Sulmasy LS, Worster BK, Chaudhry HJ, Rhyne JA, Arora VM. Online medical professionalism: patient and public relationships: policy statement from the American College of Physicians and the Federation of State Medical Boards. *Ann Intern Med* 2013;158(8):620-627 [FREE Full text] [doi: 10.7326/0003-4819-158-8-201304160-00100] [Medline: 23579867]
6. Association AM. The AMA code of medical ethics' opinions on observing professional boundaries and meeting professional responsibilities. *AMA J Ethics* 2015;17:432-434 [FREE Full text] [doi: 10.1001/journalofethics.2015.17.5.coet1-1505]
7. Busl KM, Rubin MA, Tolchin BD, Larriviere D, Epstein L, Kirschen M, et al. Use of social media in health care-opportunities, challenges, and ethical considerations: a position statement of the American Academy of Neurology. *Neurology* 2021;97(12):585-594 [FREE Full text] [doi: 10.1212/WNL.0000000000012557] [Medline: 34864637]
8. Militello M, Yang RA, Anderson JB, Szeto MD, Presley CL, Laughter MR. Social media and ethical challenges for the dermatologist. *Curr Dermatol Rep* 2021;10(4):120-127 [FREE Full text] [doi: 10.1007/s13671-021-00340-7] [Medline: 34540357]
9. Gifford ED, Mouawad NJ, Bowser KE, Bush RL, Chandra V, Coleman DM, et al. Society for vascular surgery best practice recommendations for use of social media. *J Vasc Surg* 2021;74(6):1783.e1-1791.e1 [FREE Full text] [doi: 10.1016/j.jvs.2021.08.073] [Medline: 34673169]
10. Ahmed W, Jagsi R, Gutheil TG, Katz MS. Public disclosure on social media of identifiable patient information by health professionals: content analysis of Twitter data. *J Med Internet Res* 2020;22(9):e19746 [FREE Full text] [doi: 10.2196/19746] [Medline: 32870160]
11. Attai DJ, Anderson PF, Fisch MJ, Graham DL, Katz MS, Kesselheim J, et al. Risks and benefits of Twitter use by hematologists/oncologists in the era of digital medicine. *Semin Hematol* 2017;54(4):198-204 [FREE Full text] [doi: 10.1053/j.seminhematol.2017.08.001] [Medline: 29153081]
12. Chandratilake M, McAleer S, Gibson J. Cultural similarities and differences in medical professionalism: a multi-region study. *Med Educ* 2012;46(3):257-266. [doi: 10.1111/j.1365-2923.2011.04153.x] [Medline: 22324525]
13. ABIM Foundation, ACP-ASIM Foundation, European Federation of Internal Medicine. Medical professionalism in the new millennium: a physician charter. *Ann Intern Med* 2002;136(3):243-246 [FREE Full text] [doi: 10.7326/0003-4819-136-3-200202050-00012] [Medline: 11827500]
14. Schwarz NE, Sudman SE. *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research*. Francisco, CA: Jossey-Bass/Wiley; 1996.
15. Morgado FFR, Meireles JFF, Neves CM, Amaral ACS, Ferreira MEC. Scale development: ten main limitations and recommendations to improve future research practices. *Psicol Reflex Crit* 2018;30(1):3 [FREE Full text] [doi: 10.1186/s41155-016-0057-1] [Medline: 32025957]
16. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Front Public Health* 2018;6:149 [FREE Full text] [doi: 10.3389/fpubh.2018.00149] [Medline: 29942800]
17. Ahmed O, Jilani S, Ginsburg M, Hadied O, Tasse J, Loanzon R, et al. You are what you tweet: navigating legal issues in social media for interventional radiologists. *J Vasc Interv Radiol* 2018;29(6):816-818. [doi: 10.1016/j.jvir.2018.02.012] [Medline: 29798756]
18. Keller EJ, Vogelzang RL. Who we are and what we can become: the anthropology of IR and challenges of forming a new specialty. *J Vasc Interv Radiol* 2018;29(12):1703.e2-1704.e2. [doi: 10.1016/j.jvir.2018.07.021] [Medline: 30502878]

19. Rostampour S, Hamady MS, Alsafi A. To tweet or not to tweet? A look at radiology societies' use of Twitter. *Cardiovasc Intervent Radiol* 2020;43(7):1070-1074. [doi: [10.1007/s00270-020-02437-1](https://doi.org/10.1007/s00270-020-02437-1)] [Medline: [32239244](https://pubmed.ncbi.nlm.nih.gov/32239244/)]
20. Goldsweig AM, Galper BZ, Alraies C, Arnold SV, Daniels M, Capodanno D, et al. #SoMe for #IC: optimal use of social media in interventional cardiology. *Catheter Cardiovasc Interv* 2021;98(1):97-106 [FREE Full text] [doi: [10.1002/ccd.29643](https://doi.org/10.1002/ccd.29643)] [Medline: [33686726](https://pubmed.ncbi.nlm.nih.gov/33686726/)]
21. Fischman DL, Savage MP. Cardiotwitter: new virtual tools to advance skillsets in interventional cardiology. *Curr Cardiol Rev* 2021;17(2):157-160 [FREE Full text] [doi: [10.2174/1573403X15666191126104402](https://doi.org/10.2174/1573403X15666191126104402)] [Medline: [31769364](https://pubmed.ncbi.nlm.nih.gov/31769364/)]

Abbreviations

HIPAA: Health Insurance Portability and Accountability Act

IR: Interventional Radiology

SoMe: medical social media

Edited by E Lee; submitted 31.03.23; peer-reviewed by B Chakalov; comments to author 18.09.23; revised version received 25.09.23; accepted 30.09.23; published 27.03.24.

Please cite as:

Mlambo VC, Keller E, Mussatto C, Hwang G

Development of a Medical Social Media Ethics Scale and Assessment of #IRad, #CardioTwitter, and #MedTwitter Posts: Mixed Methods Study

JMIR Infodemiology 2024;4:e47770

URL: <https://infodemiology.jmir.org/2024/1/e47770>

doi: [10.2196/47770](https://doi.org/10.2196/47770)

PMID: [38536206](https://pubmed.ncbi.nlm.nih.gov/38536206/)

©Vongai Christine Mlambo, Eric Keller, Caroline Mussatto, Gloria Hwang. Originally published in JMIR Infodemiology (<https://infodemiology.jmir.org>), 27.03.2024. 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.

Corrigenda and Addenda

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

Ryuichiro Ueda¹, MHA; Feng Han², MHA; Hongjian Zhang¹, PhD; Tomohiro Aoki¹, MHA; Katsuhiko Ogasawara¹, MBA, PhD

¹Faculty of Health Sciences, Hokkaido University, Sapporo, Japan

²Graduate School of Medicine, Hokkaido University, Sapporo, Japan

Corresponding Author:

Katsuhiko Ogasawara, MBA, PhD

Faculty of Health Sciences

Hokkaido University

N12-W5, Kita-ku

Sapporo, 060-0812

Japan

Phone: 81 11 706 3409

Email: oga@hs.hokudai.ac.jp

Related Article:

Correction of: <https://infodemiology.jmir.org/2024/1/e37881>

(*JMIR Infodemiology* 2024;4:e57880) doi:[10.2196/57880](https://doi.org/10.2196/57880)

In “Verification in the Early Stages of the COVID-19 Pandemic: Sentiment Analysis of Japanese Twitter Users” (*JMIR Infod* 2024;3(1):e37881) the authors made 3 corrections.

1. The authorship list was previously listed as:

Ryuichiro Ueda, MA; Feng Han, MA; Hongjian Zhang, MD; Tomohiro Aoki, MA; Katsuhiko Ogasawara, Prof Dr

And has now been changed to:

Ryuichiro Ueda, MHA; Feng Han, MHA; Hongjian Zhang, PhD; Tomohiro Aoki, MHA; Katsuhiko Ogasawara, MBA, PhD

2. Author Feng Han’s affiliation was originally:

Faculty of Health Sciences, Hokkaido University, Sapporo, Japan

And was changed to:

Graduate School of Medicine, Hokkaido University, Sapporo, Japan

3. The phone number listed for the corresponding author was originally:

81 011 716 2111

And was changed to:

81 11 706 3409

The correction will appear in the online version of the paper on the JMIR Publications website on March 14, 2024 together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 28.02.24; this is a non-peer-reviewed article; accepted 28.02.24; published 14.03.24.

Please cite as:

Ueda R, Han F, Zhang H, Aoki T, Ogasawara K

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

JMIR Infodemiology 2024;4:e57880

URL: <https://infodemiology.jmir.org/2024/1/e57880>

doi:[10.2196/57880](https://doi.org/10.2196/57880)

PMID:[38484301](https://pubmed.ncbi.nlm.nih.gov/38484301/)

©Ryuichiro Ueda, Feng Han, Hongjian Zhang, Tomohiro Aoki, Katsuhiko Ogasawara. Originally published in JMIR Infodemiology (<https://infodemiology.jmir.org>), 14.03.2024. 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.

Research Letter

Using Social Listening for Digital Public Health Surveillance of Human Papillomavirus Vaccine Misinformation Online: Exploratory Study

Dannell Boatman^{1,2}, MS, EdD; Abby Starkey^{1,2}, MS; Lori Acciavatti^{1,2}, MS; Zachary Jarrett^{1,2}, MA; Amy Allen^{1,2}, MA, MS; Stephenie Kennedy-Rea^{1,2}, MA, EdD

¹Department of Cancer Prevention & Control, School of Medicine, West Virginia University, Morgantown, WV, United States

²West Virginia University Cancer Institute, Morgantown, WV, United States

Corresponding Author:

Dannell Boatman, MS, EdD

Department of Cancer Prevention & Control

School of Medicine

West Virginia University

PO Box 9350

Morgantown, WV, 26506

United States

Phone: 1 304 293 7883

Email: dboatman@hsc.wvu.edu

Abstract

Despite challenges related to the data quality, representativeness, and accuracy of artificial intelligence–driven tools, commercially available social listening platforms have many of the attributes needed to be used for digital public health surveillance of human papillomavirus vaccination misinformation in the online ecosystem.

(*JMIR Infodemiology* 2024;4:e54000) doi:[10.2196/54000](https://doi.org/10.2196/54000)

KEYWORDS

human papillomavirus; HPV; vaccine; vaccines; vaccination; vaccinations; sexually transmitted infection; STI; sexually transmitted disease; STD; sexual transmission; sexually transmitted; social media; social listening; cancer; surveillance; health communication; misinformation; artificial intelligence; AI; infodemiology; infoveillance; oncology

Introduction

The COVID-19 pandemic accelerated the spread of misinformation online, creating an “infodemic” that had profound effects on health behavior [1]. The breadth and depth of COVID-19 misinformation expanded to include all vaccinations, such as human papillomavirus (HPV) vaccination, depressing already suboptimal vaccination uptake in the United States [1,2]. As HPV vaccination is critical to the prevention of various cancers, this could pose significant cancer control challenges in the future [2]. There is an urgent need to address HPV vaccination misinformation to increase HPV vaccination uptake [2]. Behavioral interventions can counter misinformation online, but they are typically limited to a single social media platform without geographic specificity [3].

Public health surveillance (PHS) is defined by the Centers for Disease Control and Prevention (CDC) as “the ongoing, systematic collection, analysis, interpretation, and dissemination

of data regarding a health-related event for use in public health action to reduce morbidity and mortality and to improve health” [4]. Digital PHS (DPHS) uses data from online sources, often collected outside of traditional PHS, for similar purposes [5]. There has been debate as to the ethics of using publicly available online data for DPHS [5]. However, the pandemic illustrated the need for user-friendly, timely, interactive digital tools to drive health-related intervention [6].

Social listening (SL) is the process of aggregating data from across online channels to collect real-time measures of emotions, opinions, and themes, typically through platform algorithms that rely on machine learning and artificial intelligence (AI) [7]. While SL platforms’ AI-driven tools for emotion and sentiment detection can be unreliable, machine learning provides an opportunity to “train” SL platforms for greater accuracy over time in the automated recognition of emotions and sentiments [8]. The World Health Organization Early AI-Supported Response With Social Listening Platform (WHO EARS) uses

an SL dashboard to provide health professionals access to information from across the internet to assist in the development of timely responses to COVID-19 narratives that occur online at the global and country levels, highlighting the growing acceptance of such tools in public health [7].

The purpose of this exploratory study was to assess the feasibility of using a commercially available SL platform to monitor HPV vaccination misinformation online at the national (ie, within the United States overall) and state (ie, within Mississippi and Rhode Island) levels.

Methods

Ethical Considerations

This study received institutional review board exemption from West Virginia University (protocol #00152755).

Study Design

Brandwatch was the commercially available SL platform selected for this exploratory study. It was selected after reviewing functionalities of leading SL platforms and having conversations about capabilities with representatives from Agorapulse, Brandwatch, Hootsuite, and Sprout Social. While most platforms had similar functionalities and data access, Brandwatch was selected based on opportunities to build queries with greater geographic specificity. While there is limited research on SL platform functionality within public health, Brandwatch was previously studied for the accuracy of AI-driven analyses [8]. The previously cited limitations of Brandwatch AI-driven tools informed the study team's systematic, routine approach to training.

The research team received onboarding from Brandwatch through 5 structured, live training sessions. Two research team members completed a self-paced online training certificate. After onboarding was complete, the research team's SL lead analyst (AS) built an HPV vaccination query within Brandwatch, using keywords and phrases identified through previous research and with research team consensus [9]. From this query, AS, with support from Brandwatch developers, created a dashboard to monitor online conversations within the United States overall and in 2 states—Mississippi, the US state with the lowest HPV vaccination rate, and Rhode Island, the US state with the highest vaccination rate. The research team regularly reviewed the query keywords and updated them as needed for increased relevancy and accuracy.

Brandwatch AI-driven tools were trained to recognize sentiments and emotions related to HPV vaccination. Sentiment categories for this study were different from the ones provided automatically by Brandwatch within the platform and were determined by the research team based on previous research [9]. Sentiment categories included “fact-based information,” “pro-vaccine opinions,” “misinformation,” “anti-vaccine opinions,” and “neutral comments.” These sentiment categories were built into the dashboard by a Brandwatch developer in conjunction with AS. The initial AI-driven recognition of these content categories was inaccurate. For example, all content that mentioned “cancer” was automatically considered negative by the SL platform AI. AS trained the AI-driven sentiment tool to

recognize the intended content by reviewing aggregated social media comments, as well as other online articles and posts within Brandwatch, and adding them to the appropriate categories to spur AI recognition. During this AI training process, another sentiment category—“irrelevant”—was added, as content that used similar language but was not directly related to HPV was identified. The Brandwatch AI-driven sentiment tool was trained by AS routinely over a 6-month period to enhance the recognition of categories. This routine training significantly improved category recognition within the SL platform but was not completely accurate upon periodic spot reviews by the research team. The AI-driven tool for recognizing emotions automatically included categories such as “anger,” “disgust,” “fear,” “joy,” “sadness,” and “surprise.” Like the AI-driven sentiment tool, the identification of correct emotion categories was initially incorrect and required routine training by AS to improve accuracy.

Once the SL platform was built, the research team evaluated the dashboard, query, and implementation process notes to assess the feasibility of using a commercially available SL platform for HPV vaccination misinformation DPHS. This assessment was completed by using an adaptation of the CDC's attributes for an effective PHS system [4]. The attributes adapted in this study were identified from CDC iterations published since 1988 [10]. The adaption of attributes involved the inclusion of consistent elements and associated definitions from across these CDC iterations; the addition of “cost” as a potential challenge to scaling; and the removal of “predictive value positive,” as the proposed DPHS approach would assess online narratives as opposed to a specific health condition. Consensus on each attribute was reached among the research team members.

Results

Table 1 details each adapted PHS system attribute and the opportunities and limitations with regard to using a commercially available SL platform for HPV vaccination misinformation DPHS. Opportunities include user-friendly dashboards with real-time data monitoring and platform adaptability. For example, from June 21 to 24, 2023, the research team was able follow the spread of misinformation through social media posts related to a lawsuit filed by the Children's Health Defense Fund, an organization led by prominent antivaccine activist Robert Kennedy Jr. However, while the SL platform dashboards are user-friendly, it took significant staff time, expertise, and routine maintenance to keep them relevant and as accurate as possible. Brandwatch was also found to be adaptable to the ever-changing online information ecosystem; however, the quality of this information was dependent on data access agreements with individual social media companies, which could change at any time. Additional challenges to using an SL platform for DPHS include concerns with data quality, representativeness, and the accuracy of AI-driven tools. There are limited ways to validate data within the SL platform itself. Data may be downloaded from Brandwatch and externally analyzed for sentiments and emotions, but this process would remove the AI-driven, automated nature of the SL platform and reduce the effectiveness of real-time monitoring in DPHS.

Table 1. Feasibility of using a commercial social listening platform for human papillomavirus vaccination misinformation digital public health surveillance. This was assessed based on attributes of public health surveillance systems adapted from the Centers of Disease Control and Prevention [4].

Attribute	Attribute description	Social listening opportunities	Social listening limitations
Usefulness	Contribution to prevention and control of misinformation	Events that may trigger misinformation spread can be identified in real time, providing an opportunity to target intervention	Unclear if targeted interventions can effectively shift online narratives
Simplicity	Simplicity of structure and ease of use	Dashboards can automate monitoring and provide easy-to-use tools to dig deeper into observable trends	Building effective queries requires a specialized skill set, including content area knowledge and experience with social media and online ecosystems
Flexibility	Adaptable to changing information and conditions	Queries can be adapted to new information and trends by changing keywords and phrases	Requires consistent monitoring by skilled personnel to ensure queries are reflective of current conditions
Data quality	Validity and completeness of data	Queries can include data beyond social media, providing a window into narratives in online public spaces	Data are limited by access provided by specific social media companies and the effectiveness of the query, along with a current lack of external data validation
Representativeness	Accurately describes flow of health information over time and distribution by place and person	Queries can monitor conversation trends over time, such as trends among audience panels and in various locations, which provide insights into demographics and geographic boundaries	Demographic and geographic information is imprecise and is limited based on availability
Timeliness	Lapse of time between misinformation and intervention	Conversations can be monitored in real time, providing opportunities for quick responses to misinformation	Lack of evidence-based responses to counter misinformation spread
Sensitivity	Ability to identify true cases and detect misinformation	Dashboard algorithms can be trained to detect changes in sentiments and emotions, providing an opportunity to respond to trends	Effectively training algorithms to detect sentiments and emotions is time-consuming and requires a specialized skill set
Stability	System is resilient to change	Can collect new sources of online data as they emerge to remain relevant in the shifting social media and online ecosystem	Changes to social media company policies can affect access to data sources
Acceptability	Willingness of persons and organizations to participate	Data collection is passive and does not burden participants with active data requests	Ethical concerns with online public data collection
Portability	Duplication of system in another setting	Social listening platforms can be purchased and adapted to different settings and health conditions, with no specialized hardware required for operation	Effectiveness of the queries may be limited by the personnel developing them and the sophistication of the selected social listening platform
Costs	Cost-effectiveness of the system	Online services can vary in price (\geq US \$2500 annually) based on the services needed for social listening	Sophisticated social listening platforms are more costly, although they provide greater access to data and tools

While Brandwatch was selected due to opportunities for greater geographic specificity, this functionality was limited in scope to only certain social media platforms, such as X (formerly Twitter). Furthermore, geographic specificity was limited based on whether social media users used geolocation functionalities and whether locations were mentioned in profiles or posts. Despite this, the research team identified and monitored different narratives in misinformation within the two states included in this exploratory study—Rhode Island and Mississippi—suggesting the potential importance of assessing online misinformation narratives based on geographic location. For example, on the same day in January 2024, the top trending story for Rhode Island focused on the Children’s Health Defense Fund lawsuit, while in Mississippi, the top story focused on childhood injury due to vaccination.

Discussion

Our findings suggest that there are opportunities and challenges associated with using commercially available SL platforms to monitor HPV vaccination misinformation online at the national and state levels. While there were strengths across all PHS system attributes, there were also significant weaknesses. These weaknesses, particularly those related to data quality, representativeness, and the accuracy of AI-driven tools, reflect limitations to using current SL platforms for DPHS. If these challenges are addressed over time however, this level of DPHS could provide the foundation for different intervention opportunities, such as using skilled infodemiologists to counter online misinformation [11]. While the research team identified challenges with the accuracy of Brandwatch AI-driven tools, which matched previously published research [8], building DPHS capabilities now could provide critical infrastructure if

and when such tools improve over time. If found to be effective in monitoring HPV vaccine misinformation, commercially available SL platforms may be adapted to other fields and health conditions. Findings may differ based on the SL platform used and vendor access agreements with social media companies.

Future research should focus on increasing the specificity of geographic location, studying strategies to increase the accuracy of SL platform AI-driven tools, and testing targeted interventions using SL platforms.

Acknowledgments

This work was supported by a grant from the Merck Investigator Studies Program, Merck and Company, Inc.

Authors' Contributions

DB, AA, and SKR contributed to the conception of this work. DB designed this work. AS acquired the data. DB and AS contributed to the analysis of data. DB, AS, LA, and ZJ contributed to the interpretation of data. All authors contributed to drafting the manuscript, and DB approved the final version for publication.

Conflicts of Interest

None declared.

References

1. Verma G, Bhardwaj A, Aledavood T, De Choudhury M, Kumar S. Examining the impact of sharing COVID-19 misinformation online on mental health. *Sci Rep* 2022 May 16;12(1):8045 [FREE Full text] [doi: [10.1038/s41598-022-11488-y](https://doi.org/10.1038/s41598-022-11488-y)] [Medline: [35577820](https://pubmed.ncbi.nlm.nih.gov/35577820/)]
2. Troiano G, Nardi A. Vaccine hesitancy in the era of COVID-19. *Public Health* 2021 May;194:245-251 [FREE Full text] [doi: [10.1016/j.puhe.2021.02.025](https://doi.org/10.1016/j.puhe.2021.02.025)] [Medline: [33965796](https://pubmed.ncbi.nlm.nih.gov/33965796/)]
3. Gwiaździński P, Gundersen AB, Piksa M, Krysińska I, Kunst JR, Noworyta K, et al. Psychological interventions countering misinformation in social media: a scoping review. *Front Psychiatry* 2022 Jan 5;13:974782 [FREE Full text] [doi: [10.3389/fpsy.2022.974782](https://doi.org/10.3389/fpsy.2022.974782)] [Medline: [36684016](https://pubmed.ncbi.nlm.nih.gov/36684016/)]
4. Buehler JW, Sosin DM, Platt R. Evaluation of surveillance systems for early epidemic detection. In: *Infectious Disease Surveillance*. Hoboken, New Jersey: Blackwell Publishing; 2008:432-442.
5. Aiello AE, Renson A, Zivich PN. Social media- and internet-based disease surveillance for public health. *Annu Rev Public Health* 2020 Apr 2;41:101-118 [FREE Full text] [doi: [10.1146/annurev-publhealth-040119-094402](https://doi.org/10.1146/annurev-publhealth-040119-094402)] [Medline: [31905322](https://pubmed.ncbi.nlm.nih.gov/31905322/)]
6. Cocoros NM, Kirby C, Zambarano B, Ochoa A, Eberhardt K, Sb CR, et al. RiskScape: a data visualization and aggregation platform for public health surveillance using routine electronic health record data. *Am J Public Health* 2021 Feb;111(2):269-276. [doi: [10.2105/AJPH.2020.305963](https://doi.org/10.2105/AJPH.2020.305963)] [Medline: [33351660](https://pubmed.ncbi.nlm.nih.gov/33351660/)]
7. McGowan BS. World Health Organization's early AI-supported response with social listening platform. *J Med Libr Assoc* 2022 Apr 1;110(2):273-275 [FREE Full text] [doi: [10.5195/jmla.2022.1398](https://doi.org/10.5195/jmla.2022.1398)] [Medline: [35440908](https://pubmed.ncbi.nlm.nih.gov/35440908/)]
8. Hayes JL, Britt BC, Evans W, Rush SW, Towery NA, Adamson AC. Can social media listening platforms' artificial intelligence be trusted? examining the accuracy of Crimson Hexagon's (now Brandwatch Consumer Research's) AI-driven analyses. *J Advert* 2020 Sep 17;50(1):81-91. [doi: [10.1080/00913367.2020.1809576](https://doi.org/10.1080/00913367.2020.1809576)]
9. Boatman DD, Eason S, Conn ME, Kennedy-Rea SK. Human papillomavirus vaccine messaging on TikTok: social media content analysis. *Health Promot Pract* 2022 May;23(3):382-387 [FREE Full text] [doi: [10.1177/15248399211013002](https://doi.org/10.1177/15248399211013002)] [Medline: [33969725](https://pubmed.ncbi.nlm.nih.gov/33969725/)]
10. Centers for Disease Control (CDC). Guidelines for evaluating surveillance systems. *MMWR Suppl* 1988 May 6;37(5):1-18. [Medline: [3131659](https://pubmed.ncbi.nlm.nih.gov/3131659/)]
11. Gorman JM, Scales DA. Leveraging infodemiologists to counteract online misinformation: experience with COVID-19 vaccines. *Harv Kennedy Sch Misinformation Rev* 2022 Feb 16;3(1):1-7 [FREE Full text] [doi: [10.37016/mr-2020-92](https://doi.org/10.37016/mr-2020-92)]

Abbreviations

AI: artificial intelligence

CDC: Centers for Disease Control and Prevention

DPHS: digital public health surveillance

HPV: human papillomavirus

PHS: public health surveillance

SL: social listening

WHO EARS: World Health Organization Early AI-Supported Response With Social Listening Platform

Edited by T Mackey; submitted 26.10.23; peer-reviewed by R Hart-Malloy, E Miyagi; comments to author 17.01.24; accepted 07.02.24; published 08.03.24.

Please cite as:

Boatman D, Starkey A, Acciavatti L, Jarrett Z, Allen A, Kennedy-Rea S

Using Social Listening for Digital Public Health Surveillance of Human Papillomavirus Vaccine Misinformation Online: Exploratory Study

JMIR Infodemiology 2024;4:e54000

URL: <https://infodemiology.jmir.org/2024/1/e54000>

doi: [10.2196/54000](https://doi.org/10.2196/54000)

PMID: [38457224](https://pubmed.ncbi.nlm.nih.gov/38457224/)

©Dannell Boatman, Abby Starkey, Lori Acciavatti, Zachary Jarrett, Amy Allen, Stephenie Kennedy-Rea. Originally published in JMIR Infodemiology (<https://infodemiology.jmir.org>), 08.03.2024. 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.

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

<https://www.jmirpublications.com/>