

Original Paper

Public Discourse Against Masks in the COVID-19 Era: Infodemiology Study of Twitter Data

Mohammad Al-Ramahi¹, DSc; Ahmed Elnoshokaty², DSc; Omar El-Gayar³, PhD; Tareq Nasrallah⁴, PhD; Abdullah Wahbeh⁵, DSc

¹Texas A&M University-San Antonio, San Antonio, TX, United States

²Northern Michigan University, Marquette, MI, United States

³Dakota State University, Madison, SD, United States

⁴Supply Chain and Information Management Group, D'Amore-McKim School of Business, Northeastern University, Boston, MA, United States

⁵Slippery Rock University of Pennsylvania, Slippery Rock, PA, United States

Corresponding Author:

Tareq Nasrallah, PhD

Supply Chain and Information Management Group

D'Amore-McKim School of Business

Northeastern University

360 Huntington Ave

Boston, MA, 02115

United States

Phone: 1 617 373 4460

Email: t.nasrallah@northeastern.edu

Abstract

Background: Despite scientific evidence supporting the importance of wearing masks to curtail the spread of COVID-19, wearing masks has stirred up a significant debate particularly on social media.

Objective: This study aimed to investigate the topics associated with the public discourse against wearing masks in the United States. We also studied the relationship between the anti-mask discourse on social media and the number of new COVID-19 cases.

Methods: We collected a total of 51,170 English tweets between January 1, 2020, and October 27, 2020, by searching for hashtags against wearing masks. We used machine learning techniques to analyze the data collected. We investigated the relationship between the volume of tweets against mask-wearing and the daily volume of new COVID-19 cases using a Pearson correlation analysis between the two-time series.

Results: The results and analysis showed that social media could help identify important insights related to wearing masks. The results of topic mining identified 10 categories or themes of user concerns dominated by (1) constitutional rights and freedom of choice; (2) conspiracy theory, population control, and big pharma; and (3) fake news, fake numbers, and fake pandemic. Altogether, these three categories represent almost 65% of the volume of tweets against wearing masks. The relationship between the volume of tweets against wearing masks and newly reported COVID-19 cases depicted a strong correlation wherein the rise in the volume of negative tweets led the rise in the number of new cases by 9 days.

Conclusions: These findings demonstrated the potential of mining social media for understanding the public discourse about public health issues such as wearing masks during the COVID-19 pandemic. The results emphasized the relationship between the discourse on social media and the potential impact on real events such as changing the course of the pandemic. Policy makers are advised to proactively address public perception and work on shaping this perception through raising awareness, debunking negative sentiments, and prioritizing early policy intervention toward the most prevalent topics.

(*JMIR Public Health Surveill* 2021;7(4):e26780) doi: [10.2196/26780](https://doi.org/10.2196/26780)

KEYWORDS

pandemic; coronavirus; masks; social medial, opinion analysis; COVID-19

Introduction

COVID-19 is an infection caused by the novel coronavirus SARS-CoV-2 that is known to cause severe acute respiratory syndrome [1]. As of December 26, 2020, COVID-19 had affected 192 countries around the world, with a total of 80,416,535 reported cases and 1,757,888 resulting deaths [2]. The World Health Organization, the Center for Disease Control and Prevention, and other leading public health organizations have outlined several guidelines to mitigate the COVID-19 pandemic. These guidelines have also been reported in recent scientific studies regarding the spread of COVID-19. The success of initiatives aimed at reopening the national and regional (state) economies ultimately relies on public awareness and acceptance of these guidelines for limiting the transmission of COVID-19. Among these guidelines is the importance of wearing masks.

Existing studies have shown that masks could have a substantial impact on virus transmission and wearing masks might significantly decrease the number of new COVID-19 cases [3,4]. Wearing a mask was found to be more effective than just handwashing [5]. Studies have also shown that mask-wearing diminishes disease spread by reducing the transmission probability per contact. Wearing masks in the public is most effective in stopping the spread of the virus when compliance is high [6] and presents a rational way to implement as a nonpharmaceutical intervention to fight COVID-19 [7]. Wearing a face mask can be effectively combined with social distancing to flatten the epidemic curve [7]; it is also an effective method of adequate isolation for individuals [8]. Ma et al [9] found that N95 masks, medical masks, and even homemade masks could block at least 90% of the virus in aerosols. Wang et al [10] found that the necessity of wearing masks during the COVID-19 pandemic has been underemphasized by the public. Despite its importance, as supported by scientific evidence, wearing masks has stirred up a significant debate, particularly in the United States.

With millions of people forced out of public spaces, many conversations about wearing masks take place on social media [11]. Popular social media platforms, including Twitter, have enabled new channels for users to share information and their experiences [12]. These platforms provide efficient methods of information access for health surveillance and social intelligence [13-15], and they have a growing popularity for sharing and debating scientific information [16-18]. Several studies have used Twitter as a data source to demonstrate the potential to identify the public's reactions to a variety of public health

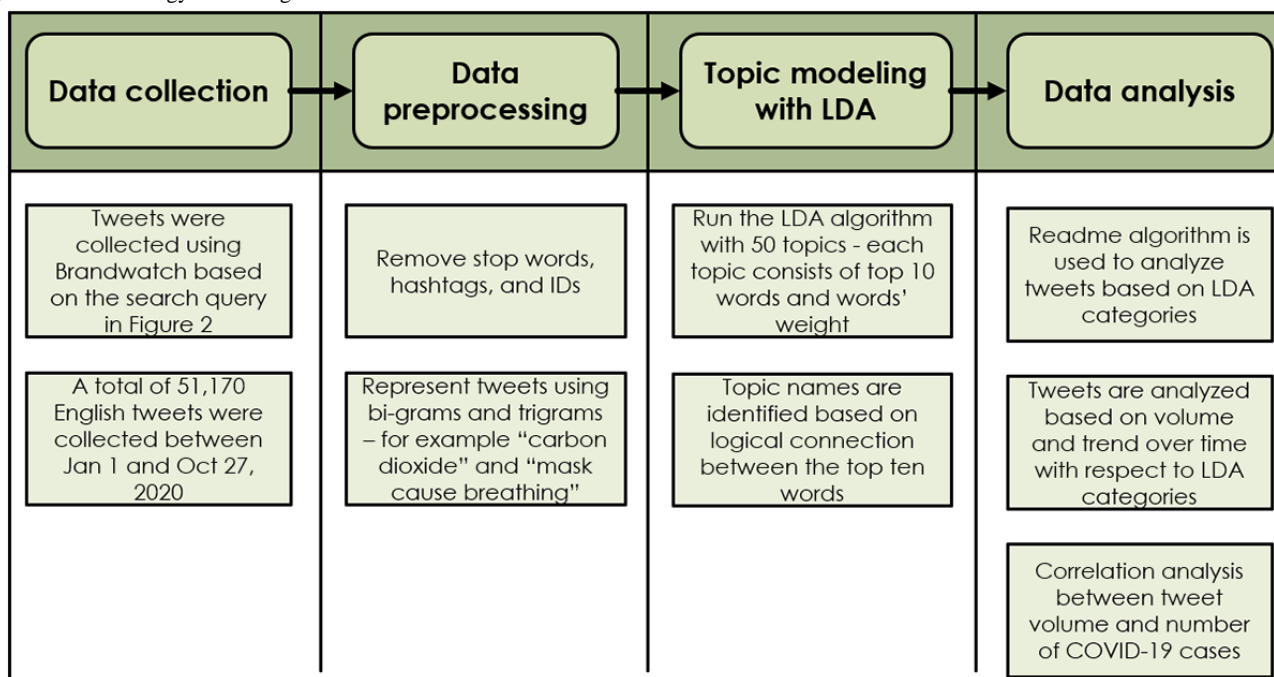
concerns, including the opioid crisis [19], marijuana [20-22], and vaping [23]. However, there are limited studies that have examined the public discourse against wearing masks on social media and its potential relation to the rise of COVID-19 cases.

With plenty of evidence supporting the effectiveness of masks in mitigating the spread of COVID-19, the vigorous public debate about masks is still ongoing [24]. Accordingly, in this study, we aim to provide insights into factors and topics encompassing the ongoing (and sometimes contentious) debate surrounding mask-wearing. Specifically, our research objective is to investigate the topics associated with the public discourse against wearing masks. The study also analyzed trends over time for each topic, with a particular emphasis on the relative volume for each topic, and the spikes in volume. Further, we studied the relationship between the anti-mask discourse on social media and the number of new COVID-19 cases. The time-lagged cross-correlation (TLCC) is used to identify directionality between two signals—volume of tweets and COVID-19 cases—to determine which signal occurs first by analyzing cross-correlations, wherein a peak correlation may have a different offset if one signal leads another. The analysis provided insights into the potential relationship between the cyber world represented by activities on social media and the physical world represented by individuals' actions and possibly reflected in increased infection rates. Such an understanding is needed as governments and public health officials grapple with reopening the economies, and keeping them open, in a manner that does not aggravate the COVID-19 pandemic as a public health crisis of epic proportions.

Methods

Figure 1 shows the methodology adopted in this study for mining social media. The first stage involved data collection. The researchers agreed on a time period of interest to collect data and keywords (ie, hashtags) to search for relevant tweets. Second, the tweets collected were preprocessed by removing stop words, keywords with IDs, and hashtags; these were then represented using bi- and trigrams. Third, a topic modeling technique, the latent Dirichlet allocation (LDA) algorithm [25], was used to analyze the preprocessed tweets to identify the prominent topics or categories in the posts. Finally, a social media analytics tool by Brandwatch was used to analyze the frequency and track the volume of the predefined categories over time. Brandwatch employs unsupervised and supervised machine learning techniques and a text analysis model developed by Hopkins and King [26].

Figure 1. Methodology for mining social media. LDA: latent Dirichlet allocation.



Data Collection

Our target social media platform for data collection was the microblogging platform Twitter. Initially, we identified all hashtags against wearing masks that were being actively used on Twitter. Next, using Brandwatch with the search query shown in Figure 2, we extracted all tweets for the identified hashtags between January 1, 2020, and October 27, 2020. A total of

51,170 English tweets were collected. The hashtags were identified by reviewing the literature [27] as well as by exploring similar trending hashtags used against wearing masks on websites such as hashtags.org [28] and hashtagify.me [29]. A key advantage of using a social media analytics platform such as Brandwatch is that it provides access to the “Twitter firehose” (ie, every public tweet ever posted on Twitter in any language and from any geographic location that meets the search criteria).

Figure 2. Hashtags and search query used for data collection.

```
(#maskoffamerica OR #NoMask* OR #takeoffthemask OR #maskburning OR #BurnYourMask OR #burnyourmaskchallenge OR #masksdontwork OR #nomaskonme OR #NoMasksOnMe OR #nomaskselfie OR #maskhoax OR #maskshoax OR #NoMasksEVER OR #NoMoreMasks OR #IWillNotWearAMask OR #MaskOff* OR #MasksOff* OR #facefreedom OR #masksmakemesweaty OR #MasksAreDangerous OR #TakeTheMaskOff OR #Stopforcingmasksonme OR #takeoffyourmask* OR #refusemask* OR #NeverMasker OR #StopWearingMask* OR #StopWearingTheDamnMasks OR #MasksdontMatter OR #stopmasking OR #stopthestupidmask OR #maskingchildrenschildabuse OR #MomsAgainstMasks OR #MasksRUNhealthy OR #SheepWearMasks OR #MasksAreForSheep OR #RefuseToWearMasks OR #MasksAreMurderingMe) AND - (RT OR http OR https)
```

For comparing the volume of tweets against wearing masks and the number of COVID-19 cases, we collected a time series of the daily number of newly reported COVID-19 cases in the United States from January to October 2020 by using data from John Hopkins University [30]. We also collected data on new COVID-19 cases reported daily in the USA from January 22 to October 27, 2020.

In acquiring data from Twitter, we considered all the common regulatory concerns that arise with social media research. Specifically, the study conforms with federal regulations on research about human subjects by using only public information that requires no interaction with the poster [31]. Moreover, the use of Brandwatch ensured that the study conformed with all the common ethical questions raised when performing web mining [32].

Data Preprocessing

We excluded retweets and addresses to focus solely on personal opinions or statements. First, the collected tweets were preprocessed by removing stop words as well as keywords with IDs and hashtags. Second, tweets were represented using unigrams, bigrams, and trigrams, such as “results,” “lab results,” and “check test results.” Word-level n-grams features were selected to represent tweets instead of the bag-of-words (ie, single words) feature because the latter has two major drawbacks: (1) they lose the ordering of the words and (2) they ignore semantics of the words [33,34].

Data Analysis Using the LDA Algorithm (Unsupervised Learning)

To discover the abstract “topics” that occur in the collected posts, we ran a topic mining model, specifically the LDA

algorithm, with 50 topics. Given a set of documents, $D = \{d_1, d_2, \dots, d_n\}$; a number of topics, $T = \{t_1, t_2, \dots, t_m\}$; and a number of words in each topic, $W = \{w_1, w_2, \dots, w_k\}$; the LDA algorithm generates the following:

- A $D \times T$ matrix with $n \times m$ size, where the weight w_{ij} is the association between a document d_i and a topic t_j [35].
- A $T \times W$ matrix with $m \times k$ size, where the weight w_{ij} is the association between the topic t_i and a word w_j [35].

The corresponding reproductive process is shown below [35,36]:

1. For each topic $t \in \{1, \dots, m\}$,
 - i. generate a probability distribution over words
 $\beta_t \sim \text{Dirichlet}(\eta)$
2. For each document d ,
 - i. generate a vector of the topic probability distribution
 $\theta_d \sim \text{Dirichlet}(\alpha)$
 - ii. For each word w_i in document d ,
 - a. generate a topic assignment
 $z_i \sim \text{Multinomial}(\theta_d)$;
 - b. generate a word
 $w_i \sim \text{Multinomial}(\beta_{z_i})$

β_t is the word distribution for topic t , and θ_d is the topic distribution for document d . The notations η and α are model parameters.

Topic models are statistical based models for uncovering the main themes (ie, set of topics) that depict a large and unstructured collection of documents. Topic models make it possible to summarize textual data at a scale that cannot possibly be tackled by human annotation. In this study, we chose the LDA algorithm [25] owing to its conceptual advantage over other latent topic models [35-38].

The 50 topics from the LDA were labeled by first author and validated by second author. The identified topics were further analyzed and grouped into 10 representative categories. The grouping was done based on semantic similarities between the topics identified. For example, the topics “build herd immunity,” “herd immunity,” and “build immune system” could be grouped into in one main topic, namely, “herd immunity and dependency on the immune system.” Overall, we discovered and collected 10 different categories.

Analysis of Tweets Using Categories Obtained (Supervised Learning)

Brandwatch employs ReadMe, a supervised algorithm developed by Hopkins and King [26]. The algorithm is particularly suited when the objective is to know the proportion of all posts that fit in specific categories. Rather than calculating these proportions based on the categorization of individual posts, ReadMe gives approximately unbiased estimates of category proportions even when the optimal classifier performs poorly [26].

The ReadMe algorithm requires the researcher to hand-code a “training set” of documents into a set of predefined categories.

In this study, the tweets represent the set of documents and the predefined categories are obtained using the LDA algorithms. The authors hand-coded 20 tweets into each predefined category obtained from the LDA and then ran the ReadMe algorithm iteratively on the remaining posts, ensuring that the examples clearly outline each category. Then, based on the training phase, the algorithm builds a model that can automatically assign the remaining tweets into categories and obtain the total number of tweets in each category. Brandwatch automatically generates the trends of tweet volumes over time.

Analyzing the Relationship Between the Tweet Volume and the Number of COVID-19 Cases

To analyze the relationship between the volume of tweets against mask-wearing and the daily volume of new COVID-19 cases, we plotted two time-series over the time span from January to October 2020 and calculated the Pearson correlation coefficient, which measures how two continuous waves co-vary over time and indicate the linear relationship as a number ranging from -1 (negatively correlated) to 0 (not correlated) to 1 (perfectly correlated) [39]. The correlation is a snapshot measure of global synchrony. Although the Pearson correlation coefficient provides a very simple way to compute both global and local synchrony, it does not provide insights into signal dynamics such as which signal occurs first or which can be measured via cross-correlations. A TLCC can identify directionality between two signals such as a leader-follower relationship. We can get a sense of which signal occurs first by looking at cross-correlations. A TLCC is measured by incrementally shifting one time-series vector and repeatedly calculating the correlation between two signals. If the peak correlation is at the center (offset=0), this indicates that the two time-series are perfectly synchronized at that time. However, the peak correlation may have a different offset if one signal leads another [40]. To analyze the relationship between the two time-series, the volume of tweets against mask-wearing, and the daily volume of new COVID-19 cases, we calculated the Pearson correlation coefficient and TLCC in Python using the SciPy package.

Results

Tweet Distribution and Categories

Overview

A total of 51,170 tweets were analyzed with respect to categories identified from the LDA model. These categories were mainly related to (ordered per their frequency in posts) (1) constitutional rights and freedom of choice; (2) conspiracy theory, population control, and big pharma; (3) fake news, fake numbers, fake pandemic, and lies; (4) unhealthy, low oxygen, carbon dioxide, lung infections, and weakened immune system; (5) political, fear, and control people; (6) masks ineffective and cannot block tiny particles; (7) mental health and suicide; (8) herd immunity and dependency on the immune system; (9) child abuse and dehumanization; and (10) virus-related statistics (high recovery rates and low mortality rates). Figure 3 shows the word clouds for the first three categories. The distribution of the tweets over the categories identified is shown in Figure 4.

Figure 3. Word cloud for the most common categories identified: (1) constitutional rights and freedom of choice; (2) conspiracy theory, population control and big pharma; AND (3) fake news, fake numbers, fake pandemic, and lies.



Figure 4. Distribution of 51,170 tweets across the top 10 categories obtained using the latent Dirichlet allocation model.

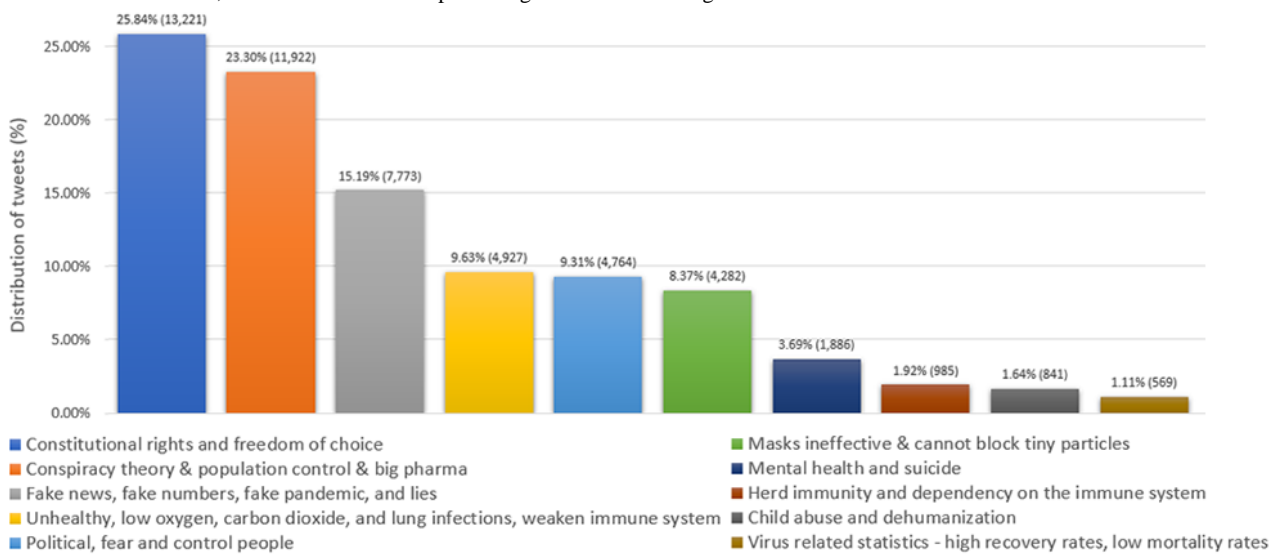
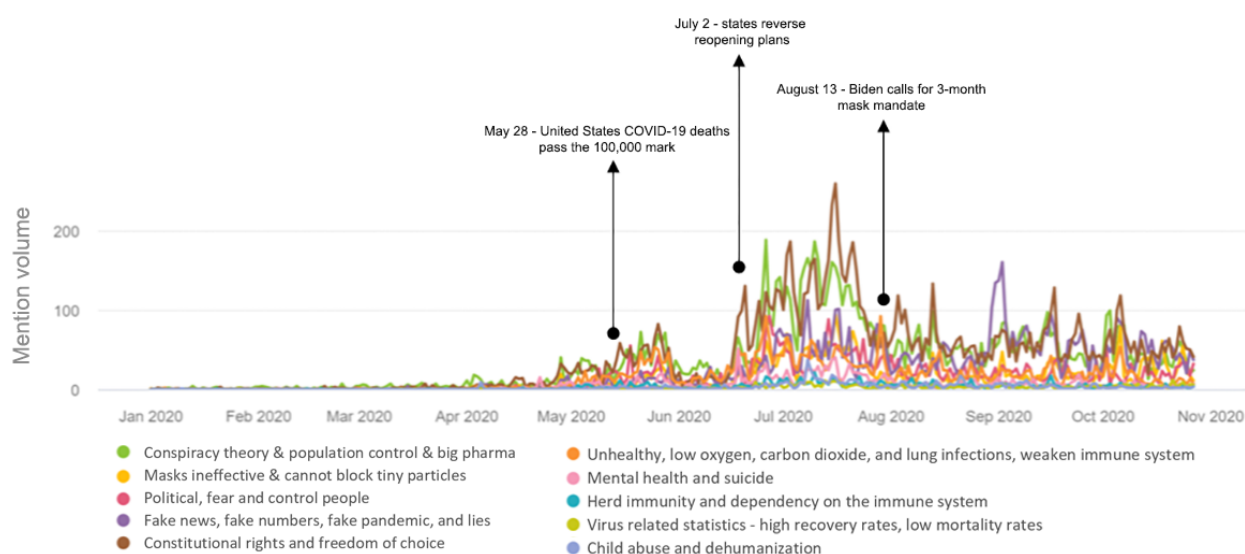


Figure 5 shows the volume of tweets over time by category. Overall, the number of tweets posted increased with time, with the highest volume of tweets recorded in July 2020. Between April 8, 2020, and May 29, 2020, a total of 15 states issued a mask mandate, which could be related to the spike in tweets posted on masks between April and the beginning of July 2020. Furthermore, between June 18 and August 11, 2020, another 20 states issued mask mandates [41]; this could explain the increase in tweets posted about masks between late-June and mid-August. Figure 5 also shows three relevant milestones between May and August 2020 [42]. These three milestones are related to the number of deaths reported in late-May, states reversing reopening plans, and the call for 3-month mask

mandates. These milestones could also relate to the increasing number of Twitter posts on masks. Furthermore, after August 13, 2020, we noted consistent debates on masks across all post categories.

Figure 5 also shows that more tweets were posted as governments and public health officials relaxed the lockdown restrictions but requested people to continue wearing masks. The number of tweets posted about constitutional rights and freedom of choice increased noticeably, followed by tweets about conspiracy theory, population control, and big pharma. The following paragraphs provide a synopsis of each of the categories of tweets posted.

Figure 5. Volume of tweets and trend analysis over 10 categories based on the latent Dirichlet allocation model as well as three significant milestones during the pandemic between January 1, 2020 and October 27, 2020.



Constitutional Rights and Freedom of Choice

Our results revealed several reasons why some Americans refuse to wear face masks despite the overwhelming evidence that wearing masks saves lives. One important reason discussed during the study period was constitutional rights and freedom of choice. Many say mandatory masks violate their constitutional right and freedom of choice. An example tweet is shown below:

*Dear #****, I am an American citizen with constitutional rights. I have the right & freedom to choose #NoMask. If u try to enforce this ridiculous order, I will sue your ass 2 hell & back. Kentucky is a #redstate & you don't belong. GTFO. Signed a pissed of Kentucky girl*

Conspiracy Theory, Population Control, and Big Pharma

Americans also discussed concerns related to conspiracy theory, population control, and big pharma. They believed that COVID-19 was human-engineered. Example tweets are shown below:

Won't have to listen to people blabbering on about their latest favourite conspiracy theory

*You can have a ridiculous opinion. Democrats follow blindly, I do not. **** IS Big Pharma. Masks = Control = Submission that will lead to mandatory inoculation of a genetically modifying vaccine. If dems win, we all lose. #MasksOffAmerica*

Fake News, Fake Numbers, Fake Pandemic, and Lies

Many also believed the pandemic is fake and there was fake news, misinformation, and lies spread about COVID-19. Example tweets are shown below:

*@**** Seasonal flu kills more people EVERY year. You and the fake news media are losing credibility FAST. #nomasks #nonewnormal*

*@**** So how many other false positives are out there...this makes the numbers even more questionable*

Unhealthy, Low Oxygen, Carbon Dioxide, Lung Infections, and Weakened Immune System

Tweets posted also discussed the health impact of wearing masks. Many believed masks limit oxygen intake and cause rebreathing of carbon dioxide, which can lead to lung disease and weaken the immune system. Example tweets are shown below:

Wearing it blocks oxygen and recycles carbon dioxide and carries the bacteria to your respiratory system. #nomasks

Masks weaken the immune system. Masks allow oral bacteria to affect gums, throat & lungs. Masks limit oxygen intake. Masks cause rebreathing of carbon dioxide

#COVID-19 #NoMasks Hypercapnia is generally caused by hypoventilation, lung disease, or diminished consciousness

Political, Fear, and Control People

Another topic discussed by Americans on Twitter was fearmongering. Many users believed that politicians and media have only focused on the numbers that present a negative picture of the COVID-19 pandemic rather than a more balanced and honest overview of the case numbers. Example tweets are shown below:

*@**** Nor do they speak about the low death rate. They want us living in fear. Fear controls the masses! #SheepNoMore #MaskOff*

FEAR MONGERING!!! THIS IS WHAT IT LEADS TO! ENOUGH! NO MORE MASKS!!

Masks Ineffective and Cannot Block Tiny Particles

Many users also had an opinion that masks are ineffective and cannot block tiny particles. Example tweets are shown below:

People wearing #masks and shaming others for NOT wearing them though all #science deems them almost

totally ineffective in protecting against the nano particles of the coronavirus. #DumbPandemicDecisions #Masks4All #MasksOff

*@**** says masks are ineffective to stop the virus. Why is there a state execution/executive order now to mandate masks? #NoMasks #ControlRemedy*

Mental Health and Suicide

Many users thought that wearing mask could also have impact on the mental health of people and could lead to suicidal thoughts. Example tweets are shown below:

they are causing a severe mental health issue. #NoMasks #MasksOff

Masks are causing horrible harm with the mental health of children. Stop wearing them before these damages are irreversible! #NoMasks #MasksOffArizona

Masks are causing serious mental health issues in children. Stop with the masks before it's too late! #MasksOff

*Where is the **** physician saying that this lockdown needs to end b/c suicide is up? Mental health has been ignored completely*

Herd Immunity and Dependency on the Immune System

People should not be forced to wear masks in order to build herd immunity and maintain a healthy and strong immune system. Example tweets are shown below:

*It's time we focus on REAL solutions like herd immunity. #NOMASK for me. @*****

You need INTERACTION with people and #NoMasks to maintain a healthy immune system #OpenAmericaNOW #OPenHawaiiNow

I will NOT wear a damn mask!! It is my right to come in to contact with germs that strengthen my immune system!

Child Abuse and Dehumanization

Asking children to wear masks was considered child abuse according to many Twitter users in USA. Example tweets are shown below:

Masking children is child abuse! Kids are not at risk and not carriers of the virus! Kids need to see and communicate clearly. They need to see facial expressions. A mask desensitizes kids! #maskingchildrenschildabuse

Mandating our young children to wear a mask for 7hrs per day while attending school is tantamount to child abuse. #OpenTheSchools #NoMasks

Masks in this case are a tool for soft torture and dehumanization #NoMasks

Virus-Related Statistics (High Recovery Rates and Low Mortality Rates)

Twitter users also discussed that the high recovery and low mortality rate of the virus that make wearing mask not necessary. Example tweets are shown below:

I will not comply and wear a useless mask that has potential health risks to me for a virus that has a 98% recovery rate. #NoMask

COVID-19 Mortality Rate in CA is .00006925% that means 99.999932% are forced 2 destroy R lives 4the weakest virus on the planet! Stop Quarantining the Healthy, Open up Businesses & only Quarantine the Sick! #UnMaskAmerica

Tweets Versus New COVID-19 Cases

Figure 6 depicts the volume of tweets against wearing masks and the number of newly reported COVID-19 cases over the study period. The two time-series exhibit a high positive Pearson correlation ($r=0.77$). Since information about directionality between the two waves—leading and following—cannot be interpreted solely from this data, we further studied the relationship between both waves (Figure 7). Overall, the results show a 9-day lead for tweet volume over the number of new COVID-19 cases. This 9-day lag is considered comparable to the number of days after which people can develop COVID-19 symptoms. According to a previous study, approximately, 97% of people infected with COVID-19 developed symptoms within 12 days after exposure [43].

Figure 6. Pearson correlation of tweets against wearing masks and newly confirmed COVID-19 cases over time (days) between January 2020 and October 2020.

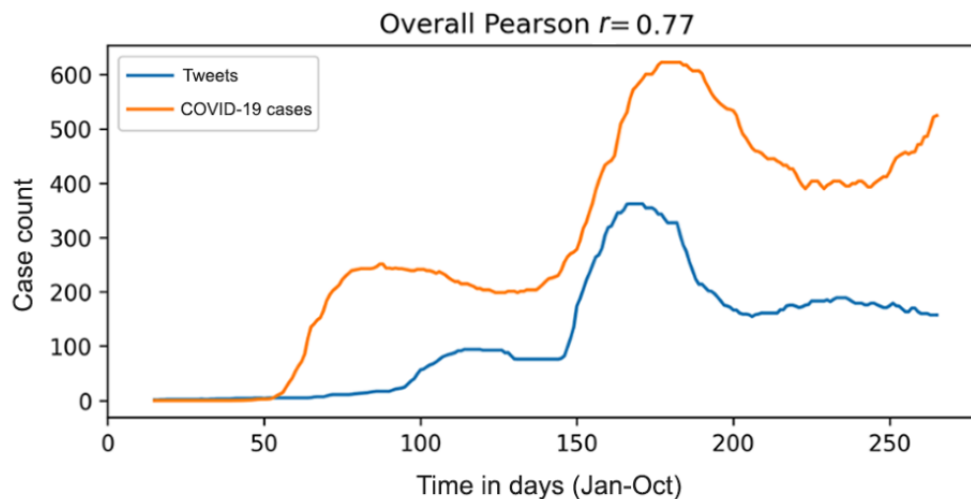
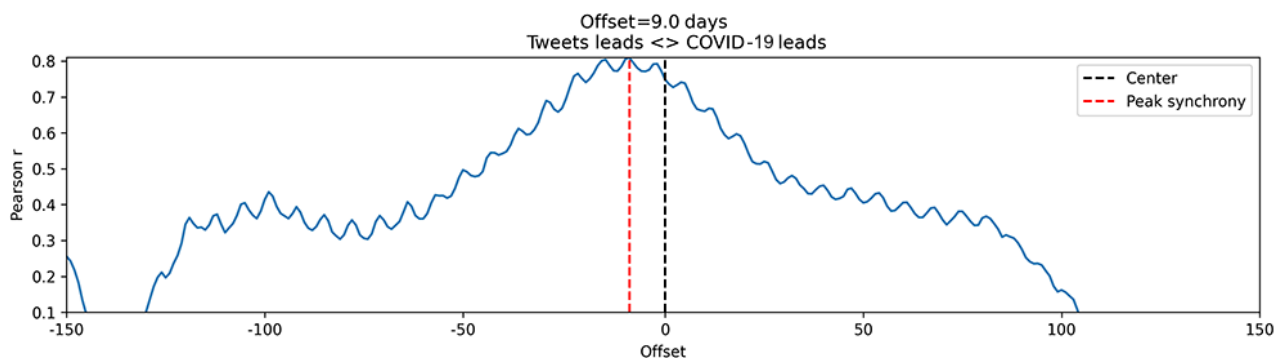


Figure 7. Graph illustrating the 9-day lead in the volume of tweets against mask-wearing compared with new COVID-19 cases by 9 days (study period: January to October 2020).



Discussion

Principal Findings

This study analyzed the negative stance regarding masks on social media, the specific themes within this discourse, and how this discourse could be associated with the prevalence of new COVID-19 cases. The study reported Twitter users' concerns related to constitutional rights and freedom of choice, conspiracy theory, misinformation, health issues, fearmongering, and other concerns related to the use of face masks during the COVID-19 pandemic. Furthermore, the time-series analysis demonstrated a strong correlation between the number of tweets posted against mask-wearing and the actual number of COVID-19 new cases, with the volume of negative tweets leading the number of newly reported COVID-19 cases by 9 days.

The study findings emphasize the potential relationship between social media behavior and its manifestation in the physical world. Such findings highlight the importance of listening to social media and proactively reacting to public perception in fighting COVID-19. Lyu and Wehby [44] showed that mask mandates in a number of states were associated with lowering infection rates by 0.9%-2% after wearing masks for 1-21 days. However, when the government mandates mask-wearing in public, many people feel their constitutional rights and freedom of choice are being violated [45]. As a result, there is a need to

increase awareness about the fact that wearing masks can protect others from contracting COVID-19 even though they do not fully protect the person wearing the mask from the infection [46]. The government should also address the challenges faced by implementing a balanced mask-wearing mandate that considers protecting people's lives while also protecting their freedom of choice [47].

Social media platforms have been used to spread fake news, lies, and conspiracy theories, all of which have a strong impact on people and society [48]. As a result of such an impact, the public is less likely view actions like wearing masks as a necessity to mitigate the spread of the virus during a pandemic [49]. Therefore, it is crucial that, as we seek to control the spread of COVID-19 and future viruses, we develop policies to fight against misleading and damaging conspiracy rhetoric. Similarly, there should be policies in place to combat fake news, lies, and misinformation, especially on social media, which could negatively affect the public's trust in science [49].

Health care professionals should actively engage in the conversation with the public in order to discuss scientific evidence supporting the importance of wearing a mask and debunk rumors on social media that promotes discussions related to masks causing low oxygen levels or lung infections. They should also discuss evidence and guidelines such as "wearing a mask does not raise the carbon dioxide (CO₂) level in the air

you breathe” [50] and “people aged 2 and older should wear masks in public settings and when around people who don’t live in their household” [50] to increase awareness regarding the effectiveness of masks in protecting the wearer from inhaling and spreading airborne particles.

Children of specific age groups should be encouraged to wear masks to protect them from COVID-19. However, protecting these age groups only by using a mask could prove very difficult [51]. To overcome these challenges, there is a need to advocate for parental involvement and support for the initiatives aimed at increasing mask-wearing among children [51]. Children should be encouraged to “take off their masks to breathe in fresh air after wearing masks for a certain amount of time,” and they should not wear masks in certain cases, such as while exercising [51]. In the case of noncompliance, it would be a better option for children to not wear masks and follow other measures to reduce infection risk and remain at home [51].

Following an empathetic approach to motivate people to wear masks and adhere to physical distancing could be an effective alternative [52] for fearmongering that focuses only on presenting a negative picture of the COVID-19 pandemic [53]. In addition, policy makers could use positive messaging to curb the spread of fear while still maintaining a transparent and accurate depiction of the situation [53].

With physical, mental, social, and economic burdens imposed by the pandemic, many populations may experience increased suicide risk [54]. Furthermore, the prevalence of anxiety, depression, posttraumatic stress disorder, and stress was reported to have increased in a number of countries during the COVID-19 pandemic [55]. Data analysis and event surveillance conducted during the first 6 months of the pandemic have shown impacts on suicide risk [54]. As a result, knowing the facts about masks and containing the spread of rumors can reduce stress and the adverse impacts on people’s mental health [56]. Finally, since many people believe that herd immunity is the best solution to this public health crisis and to strengthen their immune systems, a scientific and fact-driven view should be shared with the public explaining why herd immunity is not an ideal solution as has been reported by many researchers [57].

By carefully analyzing social media posts, policy and decision makers are in a better position to tailor public health awareness campaigns to respond to specific themes and thereby improve their effectiveness in a crisis situation such as the COVID-19 pandemic. Thus, exploring the categories of tweets surrounding the topic of mask-wearing during the COVID-19 pandemic may help reveal a number of insights that could help better design and implement awareness campaigns.

Limitations and Future Work

This study has some limitations that could be addressed in future research. First, although we identified a very strong correlation

between the increase in the volume of tweets against wearing masks and the rise in the number of COVID-19 cases, we cannot claim causality, as the rise in COVID-19 cases could be attributed to population density, government-enforced lockdown restrictions, and other factors that are beyond the scope of this study. Second, the study focused on analyzing English tweets in the United States. Future studies need to address and compare the public discourse on masks across different social media platforms and in different countries. Third, given the number of tweets collected and the focus on Twitter as a data source, the public discourse might not reflect the actual public opinion against masks. According to Wojcik and Hughes [58], Twitter has been found to have much younger audiences, with the most prolific 10% of users creating 80% of all tweets published. Finally, we did not separately analyze the opinions of Twitter users against masks in the early and later stages of the pandemic. Such analysis could unmask other important trends that are not discussed in this paper.

Conclusions

In this study, we analyzed tweets against wearing masks on social media to understand topics, insights, and information about user-reported issues. We used data analytics to identify trending themes and topics of concern by the public about wearing face masks. The most discussed issues were related to the constitutional rights and the freedom of choice, conspiracy theory, misinformation, health issues, fearmongering, and the ineffectiveness of masks, followed by issues related to mental health, herd immunity, child abuse, and virus-related statistics. Another key finding of this study is that it highlights the strong correlation between the increase in the volume of tweets against wearing masks and new COVID-19 cases and the lead of negative tweets published in comparison with the rise in new COVID-19 cases in the time-series analysis. In effect, these findings demonstrated the impact of social media not only on people’s opinion or perceptions about public topics but also the potential impact on real events such as changing the course of the pandemic. The significance and implication of this research transcends the COVID-19 pandemic, as it demonstrates the importance of social media mining and its potential to support public health-related policies and decisions. Government officials and decision makers could tailor and fine-tune public awareness campaigns and prioritize policy interventions toward the most discussed topics. In case of a future massive-scale health crisis such as the COVID-19 pandemic, government officials and policy makers could leverage social media analytics and surveillance as important tools in proactively responding to the impending crisis. Policy makers need to proactively address public perception and work on shaping this perception through raising awareness, debunking negative sentiments, and adopting early policy intervention to steer the wheel towards public acceptance of more precautionary measures and thereby containing the situation.

Conflicts of Interest

None declared.

References

1. Feng S, Shen C, Xia N, Song W, Fan M, Cowling BJ. Rational use of face masks in the COVID-19 pandemic. *The Lancet Respiratory Medicine* 2020 May;8(5):434-436. [doi: [10.1016/s2213-2600\(20\)30134-x](https://doi.org/10.1016/s2213-2600(20)30134-x)]
2. Hannah RE, Diana B, Edouard M, Joe H, Bobbie M, Charlie G, et al. Mortality Risk of COVID-19. *Our World in Data*. URL: <https://ourworldindata.org/mortality-risk-covid> [accessed 2021-03-22]
3. Fisman DN, Greer AL, Tuite AR. Bidirectional impact of imperfect mask use on reproduction number of COVID-19: a next generation matrix approach. *Infect Dis Model* 2020;5:405-408 [FREE Full text] [doi: [10.1016/j.idm.2020.06.004](https://doi.org/10.1016/j.idm.2020.06.004)] [Medline: [32691014](https://pubmed.ncbi.nlm.nih.gov/32691014/)]
4. Greenhalgh T, Schmid MB, Czypionka T, Bassler D, Gruer L. Face masks for the public during the covid-19 crisis. *BMJ* 2020 Apr 09;369:m1435. [doi: [10.1136/bmj.m1435](https://doi.org/10.1136/bmj.m1435)] [Medline: [32273267](https://pubmed.ncbi.nlm.nih.gov/32273267/)]
5. MacIntyre CR, Chughtai AA. A rapid systematic review of the efficacy of face masks and respirators against coronaviruses and other respiratory transmissible viruses for the community, healthcare workers and sick patients. *Int J Nurs Stud* 2020 Aug;108:103629 [FREE Full text] [doi: [10.1016/j.ijnurstu.2020.103629](https://doi.org/10.1016/j.ijnurstu.2020.103629)] [Medline: [32512240](https://pubmed.ncbi.nlm.nih.gov/32512240/)]
6. Howard J, Huang A, Li Z, Tufekci Z, Zdimal V, van der Westhuizen HM, et al. An evidence review of face masks against COVID-19. *Proc Natl Acad Sci U S A* 2021 Jan 26;118(4):e2014564118. [doi: [10.1073/pnas.2014564118](https://doi.org/10.1073/pnas.2014564118)] [Medline: [33431650](https://pubmed.ncbi.nlm.nih.gov/33431650/)]
7. Li T, Liu Y, Li M, Qian X, Dai SY. Mask or no mask for COVID-19: a public health and market study. *PLoS One* 2020;15(8):e0237691 [FREE Full text] [doi: [10.1371/journal.pone.0237691](https://doi.org/10.1371/journal.pone.0237691)] [Medline: [32797067](https://pubmed.ncbi.nlm.nih.gov/32797067/)]
8. Zhou Z, Yue D, Mu C, Zhang L. Mask is the possible key for self-isolation in COVID-19 pandemic. *J Med Virol* 2020 Oct;92(10):1745-1746 [FREE Full text] [doi: [10.1002/jmv.25846](https://doi.org/10.1002/jmv.25846)] [Medline: [32267002](https://pubmed.ncbi.nlm.nih.gov/32267002/)]
9. Ma Q, Shan H, Zhang H, Li G, Yang R, Chen J. Potential utilities of mask-wearing and instant hand hygiene for fighting SARS-CoV-2. *J Med Virol* 2020 Sep;92(9):1567-1571 [FREE Full text] [doi: [10.1002/jmv.25805](https://doi.org/10.1002/jmv.25805)] [Medline: [32232986](https://pubmed.ncbi.nlm.nih.gov/32232986/)]
10. Wang J, Pan L, Tang S, Ji JS, Shi X. Mask use during COVID-19: a risk adjusted strategy. *Environ Pollut* 2020 Nov;266(Pt 1):115099 [FREE Full text] [doi: [10.1016/j.envpol.2020.115099](https://doi.org/10.1016/j.envpol.2020.115099)] [Medline: [32623270](https://pubmed.ncbi.nlm.nih.gov/32623270/)]
11. Chen E, Lerman K, Ferrara E. Tracking social media discourse about the COVID-19 pandemic: development of a public coronavirus Twitter data set. *JMIR Public Health Surveill* 2020 May 29;6(2):e19273 [FREE Full text] [doi: [10.2196/19273](https://doi.org/10.2196/19273)] [Medline: [32427106](https://pubmed.ncbi.nlm.nih.gov/32427106/)]
12. Zhang Y, Fan Y, Ye Y, Li X, Zheng W. Detecting opioid users from twitter and understanding their perceptions toward mat. In: *IEEE Xplore*.: IEEE; 2017 Presented at: 2017 IEEE International Conference on Data Mining Workshops (ICDMW); November 18-21, 2017; New Orleans, LA, USA p. 502-509. [doi: [10.1109/icdmw.2017.71](https://doi.org/10.1109/icdmw.2017.71)]
13. Wang F, Carley KM, Zeng D, Mao W. Social computing: from social informatics to social intelligence. *IEEE Intell Syst* 2007 Mar;22(2):79-83. [doi: [10.1109/mis.2007.41](https://doi.org/10.1109/mis.2007.41)]
14. Wahbeh A, Nasralah T, El-Gayar O, Al-Ramahi M, Elnoshokaty A. Adverse Health Effects of Kratom: An Analysis of Social Media Data. 2021 Jan 05 Presented at: Proceedings of the 54th Hawaii International Conference on System Sciences; January 5-8, 2021; Grand Wailea, Maui, Hawaii (Online) p. 3934-3943. [doi: [10.24251/HICSS.2021.477](https://doi.org/10.24251/HICSS.2021.477)]
15. Wahbeh A, Nasralah T, Al-Ramahi M, El-Gayar O. Mining physicians' opinions on social media to obtain insights into COVID-19: mixed methods analysis. *JMIR Public Health Surveill* 2020 Jun 18;6(2):e19276 [FREE Full text] [doi: [10.2196/19276](https://doi.org/10.2196/19276)] [Medline: [32421686](https://pubmed.ncbi.nlm.nih.gov/32421686/)]
16. El-Gayar O, Nasralah T, Elnoshokaty A. Wearable Devices for Health and Wellbeing: Design Insights from Twitter. In: Proceedings of the 52nd Hawaii International Conference on System Sciences. 2019 Jan 08 Presented at: Hawaii International Conference on System Sciences 2019; January 7-10, 2020; Maui, Hawaii, USA. [doi: [10.24251/hicss.2019.467](https://doi.org/10.24251/hicss.2019.467)]
17. Runge KK, Yeo SK, Cacciatore M, Scheufele DA, Brossard D, Xenos M, et al. Tweeting nano: how public discourses about nanotechnology develop in social media environments. *J Nanopart Res* 2013 Jan 4;15(1):1381. [doi: [10.1007/s11051-012-1381-8](https://doi.org/10.1007/s11051-012-1381-8)]
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. Nasralah T, El-Gayar O, Wang Y. What Social Media Can Tell Us About Opioid Addicts: Twitter Data Case Analysis. In: Proceedings of the 25th Americas Conference on Information Systems (AMCIS '19). 2019 Jul 12 Presented at: 25th Americas Conference on Information Systems (AMCIS '19); August 15-17; Cancun, Mexico p. 1-10.
20. Cavazos-Rehg PA, Sowles SJ, Krauss MJ, Agbonavbare V, Gruzca R, Bierut L. A content analysis of tweets about high-potency marijuana. *Drug Alcohol Depend* 2016 Sep 01;166:100-108 [FREE Full text] [doi: [10.1016/j.drugalcdep.2016.06.034](https://doi.org/10.1016/j.drugalcdep.2016.06.034)] [Medline: [27402550](https://pubmed.ncbi.nlm.nih.gov/27402550/)]
21. Daniulaityte R, Nahhas RW, Wijeratne S, Carlson RG, Lamy FR, Martins SS, et al. "Time for dabs": Analyzing Twitter data on marijuana concentrates across the U.S. *Drug Alcohol Depend* 2015 Oct 01;155:307-311 [FREE Full text] [doi: [10.1016/j.drugalcdep.2015.07.1199](https://doi.org/10.1016/j.drugalcdep.2015.07.1199)] [Medline: [26338481](https://pubmed.ncbi.nlm.nih.gov/26338481/)]
22. Dai H, Hao J. Mining social media data on marijuana use for post traumatic stress disorder. *Computers in Human Behavior* 2017 May;70:282-290. [doi: [10.1016/j.chb.2016.12.064](https://doi.org/10.1016/j.chb.2016.12.064)]
23. Wahbeh A, Al-Ramahi M, El-Gayar O, Nasralah T. Health risks of e-cigarettes: Analysis of Twitter data using topic mining. Twenty-sixth Americas Conference on Information Systems (AMCIS) 2020 Proceedings. 16 2020 [FREE Full text]

24. Raymond J. The great mask debate: A debate that shouldn't be a debate at all. *WMJ* 2020;119(4):229-239. [Medline: [33428832](#)]
25. Blei D, Ng A, Jordan M. Latent Dirichlet allocation. *J Mach Learn Res* 2003 Jan;3:993-1022.
26. Hopkins D, King G. A method of automated nonparametric content analysis for social science. *Am J Pol Sci* 2010;54(1):2010-2047. [doi: [10.1111/j.1540-5907.2009.00428.x](#)]
27. Lewiston-Auburn. Maine's Twitter users among top for anti-mask feelings. *sunjournal.com*. 2020. URL: <https://www.sunjournal.com/2020/07/10/maines-twitter-users-among-top-for-anti-mask-feelings/> [accessed 2020-07-10]
28. Hashtags.org - Hashtag Analytics. URL: <https://www.hashtags.org/> [accessed 2021-03-29]
29. Hashtagify. URL: <https://hashtagify.me/> [accessed 2021-03-29]
30. Covid-19 Data in Motion. Johns Hopkins University & Medicine - Coronavirus Resource Center.: Johns Hopkins University URL: <https://coronavirus.jhu.edu/> [accessed 2021-03-22]
31. Moreno MA, Goniu N, Moreno PS, Diekema D. Ethics of social media research: common concerns and practical considerations. *Cyberpsychol Behav Soc Netw* 2013 Sep;16(9):708-713 [FREE Full text] [doi: [10.1089/cyber.2012.0334](#)] [Medline: [23679571](#)]
32. Krotov V, Silva L, editors. Legality and Ethics of Web Scraping. In: Twenty-fourth Americas Conference on Information Systems (AMCIS). 2018 Presented at: Twenty-fourth Americas Conference on Information Systems; August 16-18, 2018; New Orleans, USA URL: <https://aisel.aisnet.org/amcis2018/DataScience/Presentations/17>
33. Le Q, Mikolov T. Distributed Representations of Sentences and Documents. In: Proceedings of the 31st International Conference on Machine Learning. 2014 Presented at: 31st International Conference on Machine Learning; 2014; Beijing, China p. 1188-1196.
34. Wahbeh A, Al-Ramahi M, Noteboom C, Nasralah T. Discovering Patient Portal Features Critical to User Satisfaction: A Systematic Analysis. In: Proceedings of the 52nd Hawaii International Conference on System Sciences. Discovering Patient Portal Features Critical to User Satisfaction: Proceedings of the 52nd Hawaii International Conference on System Sciences; 2019 Presented at: 52nd Hawaii International Conference on System Sciences; 08-11 Jan 2019; Maui, Hawaii, USA.
35. Al-Ramahi MA, Liu J, El-Gayar OF. Discovering design principles for health behavioral change support systems: a text mining approach. *ACM Trans. Manage. Inf. Syst.* 2017 Aug 24;8(2-3):1-24. [doi: [10.1145/3055534](#)]
36. Al-Ramahi M, Noteboom C. Mining user-generated content of mobile patient portal: Dimensions of user experience. *Trans. Soc. Comput.* 2020 Aug 05;3(3):1-24. [doi: [10.1145/3394831](#)]
37. Noteboom C, Al-Ramahi M. What are the Gaps in Mobile Patient Portal? Mining Users Feedback using Topic Modeling. In: Proceedings of the 51st Hawaii International Conference on System Sciences. 2018 Presented at: Hawaii International Conference on System Sciences 2018 (HICSS-51); 3-6 January 2018; Hilton Waikoloa Village, Hawaii p. 2018 URL: https://aisel.aisnet.org/hicss-51/cl/collaborations_in_healthcare/3/
38. Al-Ramahi M, Noteboom C. A Systematic Analysis of Patient Portals Adoption, Acceptance and Usage: The Trajectory for Triple Aim? In: Proceedings of the 51st Hawaii International Conference on System Sciences. 2018 Jan 03 Presented at: Hawaii International Conference on System Sciences 2018 (HICSS-51); January 3-6, 2018; Waikoloa Village, Hawaii, USA URL: <https://scholarspace.manoa.hawaii.edu/handle/10125/49994> [doi: [10.24251/hicss.2018.107](#)]
39. Mudelsee M. Estimating Pearson's correlation coefficient with bootstrap confidence interval from serially dependent time series. *Mathematical Geology* 2003 Aug;35(6):651-665. [doi: [10.1023/b:matg.0000002982.52104.02](#)]
40. Haugh LD. Checking the independence of two covariance-stationary time series: a univariate residual cross-correlation approach. *Journal of the American Statistical Association* 1976 Jun;71(354):378-385. [doi: [10.1080/01621459.1976.10480353](#)]
41. Kim A, Andrew S, Froio J. These are the states requiring people to wear masks when out in public. *CNN*. 2020 Aug 17. URL: <https://www.cnn.com/2020/06/19/us/states-face-mask-coronavirus-trnd/index.html> [accessed 2021-03-23]
42. A Timeline of COVID-19 Developments in 2020. *AJMC*. 2021 Jan 01. URL: <https://www.ajmc.com/view/a-timeline-of-covid19-developments-in-2020> [accessed 2021-03-23]
43. Tavernise S. It may be weeks before we know if Thanksgiving travel fed a virus surge. *The New York Times*. 2020 Nov 30. URL: <https://www.nytimes.com/2020/11/30/world/thanksgiving-coronavirus-us-spread.html> [accessed 2021-03-23]
44. Lyu W, Wehby GL. Community use of face masks and COVID-19: evidence from a natural experiment of state mandates in the US. *Health Aff (Millwood)* 2020 Aug 01;39(8):1419-1425. [doi: [10.1377/hlthaff.2020.00818](#)] [Medline: [32543923](#)]
45. Scerri M, Grech V. WITHDRAWN: To wear or not to wear? Adherence to face mask use during the COVID-19 and Spanish influenza pandemics. *Early Hum Dev* 2020 Nov 12:105253 [FREE Full text] [doi: [10.1016/j.earlhumdev.2020.105253](#)] [Medline: [33221028](#)]
46. Betsch C, Korn L, Sprengholz P, Felgendreff L, Eitze S, Schmid P, et al. Social and behavioral consequences of mask policies during the COVID-19 pandemic. *Proc Natl Acad Sci U S A* 2020 Sep 08;117(36):21851-21853 [FREE Full text] [doi: [10.1073/pnas.2011674117](#)] [Medline: [32820078](#)]
47. Lenhart B. The Peoples' Constitution: COVID-19 versus Freedom. *Loudoun Now*. 2020 May 07. URL: <https://loudounnow.com/2020/05/07/the-peoples-constitution-covid-19-versus-freedom/> [accessed 2021-03-23]
48. Shu K, Sliva A, Wang S, Tang J, Liu H. Fake news detection on social media: A data mining perspective. *SIGKDD Explor Newsl* 2017 Sep;19(1):22-36. [doi: [10.1145/3137597.3137600](#)]

49. Palm R. Conspiracy Theories About The Origins Of COVID-19 Outweigh Science's Influence, Researchers Say.: Georgia State University News; 2020 Oct 29. URL: <https://news.gsu.edu/2020/10/29/covid-19-conspiracy-theory-exposure-risa-palm/> [accessed 2021-03-23]
50. Considerations for Wearing Masks Help Slow the Spread of COVID-19. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/cloth-face-cover-guidance.html> [accessed 2021-03-22]
51. Esposito S, Principi N. To mask or not to mask children to overcome COVID-19. *Eur J Pediatr* 2020 Aug 9;179(8):1267-1270 [FREE Full text] [doi: [10.1007/s00431-020-03674-9](https://doi.org/10.1007/s00431-020-03674-9)] [Medline: [32388722](https://pubmed.ncbi.nlm.nih.gov/32388722/)]
52. Pfattheicher S, Nockur L, Böhm R, Sassenrath C, Petersen M. The emotional path to action: Empathy promotes physical distancing and wearing of face masks during the COVID-19 pandemic. *Psychol Sci* 2020 Nov;31(11):1363-1373. [doi: [10.1177/0956797620964422](https://doi.org/10.1177/0956797620964422)] [Medline: [32993455](https://pubmed.ncbi.nlm.nih.gov/32993455/)]
53. Moreno R. COVID-19: The Fearmongering Must End. *FRCAction*. 2020 Oct 28. URL: <https://www.frcaction.org/updatearticle/20201028/covid-fearmongering> [accessed 2021-03-23]
54. Moutier C. Suicide prevention in the COVID-19 era: Transforming threat into opportunity. *JAMA Psychiatry* 2020 Oct 16. [doi: [10.1001/jamapsychiatry.2020.3746](https://doi.org/10.1001/jamapsychiatry.2020.3746)] [Medline: [33064124](https://pubmed.ncbi.nlm.nih.gov/33064124/)]
55. Xiong J, Lipsitz O, Nasri F, Lui LM, Gill H, Phan L, et al. Impact of COVID-19 pandemic on mental health in the general population: A systematic review. *J Affect Disord* 2020 Dec 01;277:55-64 [FREE Full text] [doi: [10.1016/j.jad.2020.08.001](https://doi.org/10.1016/j.jad.2020.08.001)] [Medline: [32799105](https://pubmed.ncbi.nlm.nih.gov/32799105/)]
56. Coping with Stress. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/coronavirus/2019-ncov/daily-life-coping/managing-stress-anxiety.html> [accessed 2021-03-23]
57. Ashwanden C. The false promise of herd immunity for COVID-19. *Nature* 2020 Nov;587(7832):26-28. [doi: [10.1038/d41586-020-02948-4](https://doi.org/10.1038/d41586-020-02948-4)] [Medline: [33087872](https://pubmed.ncbi.nlm.nih.gov/33087872/)]
58. Wojcik S, Hughes A. Sizing Up Twitter Users. Pew Research Center. 2019 Apr 24. URL: <https://www.pewresearch.org/internet/2019/04/24/sizing-up-twitter-users/> [accessed 2021-03-29]

Abbreviations

LDA: latent Dirichlet allocation

TLCC: time-lagged cross-correlation

Edited by T Sanchez; submitted 25.12.20; peer-reviewed by L Laestadius, L Shen, SH Puah, T Ramanather; comments to author 20.01.21; revised version received 03.02.21; accepted 03.03.21; published 05.04.21

Please cite as:

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;7(4):e26780

URL: <https://publichealth.jmir.org/2021/4/e26780>

doi: [10.2196/26780](https://doi.org/10.2196/26780)

PMID: [33720841](https://pubmed.ncbi.nlm.nih.gov/33720841/)

©Mohammad Al-Ramahi, Ahmed Elnoshokaty, Omar El-Gayar, Tareq Nasralah, Abdullah Wahbeh. Originally published in *JMIR Public Health and Surveillance* (<http://publichealth.jmir.org>), 05.04.2021. 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 Public Health and Surveillance*, is properly cited. The complete bibliographic information, a link to the original publication on <http://publichealth.jmir.org>, as well as this copyright and license information must be included.