

Why do people oppose mask wearing? A comprehensive analysis of US tweets during the COVID-19 pandemic

Lu He^{*1}, Changyang He^{*2}, Tera L Reynolds¹, Qiushi Bai³, Yicong Huang³, Chen Li³,
Kai Zheng^{1,4}, Yunan Chen¹

**Denotes equal contribution*

1. Department of Informatics, Donald Bren School of Information and Computer Science, University of California, Irvine, Irvine, California, United States
2. Department of Computer Science and Engineering, Hong Kong University of Science and Technology, Hong Kong SAR, China
3. Department of Computer Science, Donald Bren School of Information and Computer Science, University of California, Irvine, Irvine, California, United States
4. Department of Emergency Medicine, School of Medicine, University of California, Irvine, Irvine, California, United States

Corresponding author:

Yunan Chen, PhD, MBBS
6099 Bren Hall, Irvine
CA 92697-3440
United States
yunanc@ics.uci.edu
(949) 824-0959

Keywords: Social Media [L01.178.751], Natural Language Processing [L01.224.050.375.580], Machine Learning [G17.035.250.500], Public Health [H02.403.720], Health Communication [L01.143.350], Coronavirus [B04.820.504.540.150], Masks [E07.325.877.500], Personal Protective Equipment [E07.700.560]

Word Count: 5,544

© The Author(s) 2021. Published by Oxford University Press on behalf of the American Medical Informatics Association. All rights reserved. For permissions, please email: journals.permissions@oup.com

ABSTRACT

Objective: Facial masks are an essential personal protective measure to fight the COVID-19 pandemic. However, the mask adoption rate in the US is still less than optimal. This study aims to understand the beliefs held by individuals who oppose the use of facial masks, and the evidence that they use to support these beliefs, to inform the development of targeted public health communication strategies.

Materials and Methods: We analyzed a total of 771,268 US-based tweets between January to October 2020. We developed machine-learning classifiers to identify and categorize relevant tweets, followed by a qualitative content analysis of a subset of the tweets to understand the rationale of those opposed mask wearing.

Results: We identified 267,152 tweets that contained personal opinions about wearing facial masks to prevent the spread of COVID-19. While the majority of the tweets supported mask wearing, the proportion of anti-mask tweets stayed constant at about 10% level throughout the study period. Common reasons for opposition included physical discomfort and negative effects, lack of effectiveness, and being unnecessary or inappropriate for certain people or under certain circumstances. The opposing tweets were significantly less likely to cite external sources of information such as public health agencies' websites to support the arguments.

Discussion and Conclusion: Combining machine learning and qualitative content analysis is an effective strategy for identifying public attitudes toward mask wearing and the reasons for opposition. The results may inform better communication strategies to improve the public perception of wearing masks and, in particular, to specifically address common anti-mask beliefs.

INTRODUCTION

The COVID-19 pandemic has caused significant morbidity and mortality across the globe. On December 28, 2020, there were 441,861 new confirmed cases worldwide, with 145,959 of these being in the US.^{1,2} While scientific research has advanced our knowledge of the disease, new therapeutic treatments have been developed and vaccines are now approved and available, widespread adoption of individual protective behaviors, such as wearing personal protective equipment and practicing social distancing, remains crucial to reducing the spread of COVID-19.² Despite a handful of studies questioning the effectiveness of community mask wearing,^{3,4} it has been shown that facial masks, even if homemade, can lead to a decrease in mortality by more than 20% if worn by more than 80% of community members.⁵

The benefits of facial masks can only be realized when most people wear them, which is known as universal mask use.⁶ While several countries (e.g., Singapore, South Korea, China) have achieved this goal,⁷ promoting widespread mask wearing in the US has encountered substantial obstacles. Besides cultural norms dictating that only the sick wear masks, the anti-mask opinion held by some government officials, and shifting positions by US and international public health authorities, have added confusion to the debate. In particular, the US Centers for Disease Control and Prevention (CDC) initially recommended against public mask wearing on February 27, 2020.⁸ However, on April 3, the CDC reversed its position to instead recommend universal mask use.⁹ Then, on April 6, the World Health Organization (WHO) issued an advisory stating that healthy individuals do not need to wear masks, directly contradicting the CDC's recommendation.¹⁰ Further, there has been significant variation across state and county health departments on mask-wearing policies, and the debate on whether or not there should be a national mask mandate in the US remains unsettled.¹¹

While surveys by the New York Times, the Pew Research Center, and the CDC reported a relatively high self-reported mask use rate in the US, 59%, 65%, and 74.1%, respectively, the actual adoption rate is questionable. For example, in the same survey conducted by the Pew Research Center, only 44% of the participants reported that members in their communities were actually wearing masks all or most of the time when in public, suggesting that social desirability bias may be affecting self-reported rates.⁹ Further, all currently available surveys used simple yes-no questions on mask wearing without soliciting the rationales behind the opposing opinions;^{7,9,12} only a handful of news outlets and advocacy groups have provided excerpts from the public or conjectured on why some people refused to wear masks.^{13–16} Finally, most of the existing surveys were carried out at sporadic times for cross-sectional analysis in limited geographic areas. Continuous monitoring of public perception across the country is rare, leaving a knowledge vacuum of understanding how the public attitudes toward mask wearing have evolved over time since the beginning of the pandemic.^{7,9,12}

To address these gaps, we analyzed a large Twitter dataset collected in the US from January to October 2020 to answer the following research questions:

1. (a) What is the general public's attitude toward mask wearing in the US?
(b) How has the general public's attitude changed over time as the pandemic progressed?
2. Among those expressing an anti-mask opinion, what are their concerns or justifications?
3. What is the external source of information shared to support the pro- or anti-mask arguments?

Based on the results of available conventional surveys, we hypothesize that the general public's attitude toward facial masking expressed through tweets would be generally positive, even

though unfavorable viewpoints would not be uncommon. Further, this attitude would have shifted over time as a result of changing CDC guidelines and local mask-wearing policies and how such policies are enforced. We also hypothesize that anti-mask tweets would be less likely to cite external sources of information especially from public health authorities.

MATERIALS AND METHODS

We used a computer-aided qualitative analysis approach that combines machine learning and qualitative content analysis. To answer RQ1 (*changing attitude*), we trained a machine learning classifier to label personal opinions regarding mask wearing and to examine its evolution over time. To answer RQ2 (*concerns or justifications for opposition*), we conducted an in-depth qualitative content analysis of a random set of anti-mask tweets to examine the common beliefs held by those who opposed mask wearing, as well as the reasoning behind such beliefs. To answer RQ3 (*sharing of external information*), we quantitatively analyzed the external evidence cited in the tweets, e.g., by calculating the proportion from public health authorities such as WHO and the CDC. The overall analytical flow is exhibited in Figure 1. To protect the privacy of individuals, we paraphrased all tweets presented in the following sections, instead of directly quoting the original tweets.

[Insert Figure 1 about here]

Data Collection and Pre-Processing

Retrieving tweets

As this study concerns the public attitudes toward mask wearing in the US, only those geocoded tweets falling into the coordinates of $\{-173.847656, 17.644022, -65.390625, 70.377854\}$, which approximately represents the continental US, were included. Geocodes are the main source of information that researchers could use to determine user locations. While user profiles are available through the Twitter API, we did not opt to use them for inferring geolocations because users' locations specified at the point of registration may not match their locations when relevant tweets were posted. Therefore, only the geolocations attached to tweets meeting our criteria were used. Based on the estimate provided by Twitter Inc., approximately 1–2% of Twitter users opt to allow their geolocation information to be tracked.¹⁷

The data collection of this study covered a ten-month period between January 1 and November 1, 2020. It was conducted by leveraging the official Twitter API version 1.1, which continuously retrieves all tweets meeting the search criteria from a stream of general tweets that are located in the US provided by Twitter.¹⁸ The keywords used are detailed in Table 1, which were developed based on a manual review of sample tweets in addition to examining the search terms employed in prior research.¹⁹

Table 1. Search keywords

<p>“mask” OR “face cover” OR “cloth cover” OR “face cloth” OR “mouth cover” OR “nose cover” OR “facial cover” OR “nose cloth” OR “eye cloth” OR “mouth cloth”</p>

Further filtering

Social media data retrieved through keywords search may contain a substantial amount of irrelevant content, e.g., “face cloth” may refer to facial washcloth for makeup removal.²⁰ To

remove such noise, we manually analyzed a random set of tweets to examine the characteristics of irrelevant posts (a total of 200 tweets were reviewed; saturation was achieved after coding about 70 tweets). The results show that the majority of such tweets fall into the following two categories: (1) those related to mask manufacturing, product advertisements, or referred to other meanings of the word “mask” (e.g., “*The government has been masking the fact that the it is a failure*”); and (2) those pertinent to mask wearing but did not express any personal attitude (e.g., tweets that merely shared a URL with no personal opinions explicitly stated), or the attitude is difficult to discern from the content of the tweet (e.g., “*Should you wear a mask?? #COVID #facemask #comfortmask. Read this blog: URL*”).

To remove such tweets from further analyses, we developed two models using supervised machine learning. Model 1 (*Mask Wearing Classifier*) is a text classifier for determining whether a tweet is related to mask wearing in the COVID-19 context; and Model 2 (*Opinion Expression Classifier*) determines whether a tweet contained personal opinions. To train these models, we annotated a total of 1,000 tweets through the following two steps. First, two authors (LH and CH) separately coded a random sample of 200 tweets to calibrate the annotation. The inter-rater agreement ratio was 0.93; differences were resolved in consensus development meetings. Then, the same two authors independently coded an additional random set of 800 tweets. To prepare for text classification, all tweets were preprocessed by, e.g., lowercasing, removal of punctuations and hashtags, and stemming. Next, for each model, words were represented as 100-dimensional vectors trained by the word2vec algorithm.²¹ To address the issue of imbalance between relevant and irrelevant tweets in the training data, we used the oversampling strategy as specified by Hilario et al.²² Then, we tested several commonly used machine-learning models including Support Vector Machine (SVM), XGBoost, and Long Short-term Memory network

(LSTM). The best-performing one, LSTM, was selected for further analyses, which achieved the highest F1-score under ten-fold cross validation. LSTM, due to its ability to account for sequential information and order dependencies, is particularly suited for handling data such as time series and natural language. Extant literature has demonstrated the predictive power of LSTM on NLP tasks such as document classification and sentiment analysis.²³ We added a dropout layer (rate=0.3) to avoid overfitting, a common technique used in the literature that randomly removes units of neural networks.²⁴

To identify tweets posted by social bots (i.e., programmed Twitter accounts that generate posts automatically²⁵), we first applied Botometer,^{26,27} a well-established social bot detection tool that has been widely used by researchers and organizations such as the Pew Research Center.²⁸ Following Rauchfleisch and Kaiser's recommendation, we used a random sample of 500 distinct users from our dataset to validate the tool's performance.²⁹ Two authors (LH and CH) manually reviewed these users' profiles (e.g., profile picture, description, account creation time, number of followers), tweeting history (e.g., number of tweets, re-tweets, and likes), and interactions with other users (e.g., commenting on others' tweets), in order to determine if a user was a social bot or not. The inter-rater reliability is 100% when calibrating based on 100 users' data; none of these users were determined as social bots. We then annotated the remaining 400 users; none of them were determined as social bots either. On the contrary, Botometer labeled 29 (5.8%) users as social bots. However, based on manual review, many of these users were simply hyperactive tweeters. Their online activities did exhibit the normal behavior of human users, e.g., the content they posted did not appear to be automatically authored and they participated in active interactions with other Twitter users. Removing these users could thus result in systematic biases in our analysis of the data.

Data Analyses

RQ1a: Attitude

To classify public opinions toward mask wearing, we first applied the sentiment analysis approach which has been commonly used in the literature to study attitudes expressed in social media data.³⁰ We tested four off-the-shelf sentiment analysis tools that have been most commonly used: Valence Aware Dictionary for sEntiment Reasoning (VADER),³¹ TextBlob,³² Stanford Natural Language Processing (NLP),³³ and Linguistic Inquiry and Word Count (LIWC).^{30,34,35} We manually annotated the sentiment of 500 random tweets and compared the results to the outputs of these tools. We found that none of these off-the-shelf tools were able to produce accurate attitude classifications, at least not in our study context. The results produced by VADER, TextBlob, Stanford NLP, and LIWC achieved low F1-scores of 59%, 57.4%, 58.6%, and 51.9%, respectively. This is likely because of domain transferability issues, i.e., such tools are often trained with text corpora from non-healthcare domains such as movie reviews. Also, in our study, positivity of the sentiment is often not consistent with the mask wearing attitude expressed. For example, users tend to use strong negative tones such as “*Fuck. Mask On!*” to encourage others to wear masks, the sentiment of which was labeled by the off-the-shelf sentiment analysis as negative, even though the tweet was in fact in favor of mask wearing.

Thus, we developed a specialized machine-learning model (Model 3, *Attitude Classifier*) instead, using the same approach adopted in Model 1 and Model 2. This produced an LSTM text classifier for attitudes, trained based on 500 annotated tweets.

RQ1b: Evolution of attitude

Further, to investigate the temporal trend of public attitudes toward mask wearing, we grouped tweets by week and calculated the percentage of tweets expressing support for, or opposition to, mask wearing on a week-to-week basis. We also analyzed word frequencies based on term frequency-inverse document frequency (TF-IDF) to assess changes in commonly discussed topics related to mask wearing over time.

RQ2: Concerns or justifications for opposition

We conducted a manual qualitative content analysis on a random set of tweets posted by distinct users in order to answer the second research question, “among those expressing an anti-mask opinion, what are their concerns or justifications.” We did not opt to use a computational approach for this research question because the concerns expressed in tweets were heterogeneous and subtle that were challenging for the machine to classify. For example, the following tweet:

“Where I live in Los Angeles – You can’t get groceries without a mask... In reality, very few are sick in CA. 250 out of 40 million have died, most with preexisting health issues” implies that coronavirus does not cause many casualties and masks are not necessary; hence the attitude expressed in this tweet was opposing mask wearing. This level of natural language understanding is difficult for currently available lexicon-based tools or machine-learning models to disentangle, especially on short texts such as tweets.³⁶ We therefore decided to qualitatively analyze such concerns and justifications for opposing mask wearing. We used grounded theory to code the data.³⁷ Using this method, we first randomly selected 100 tweets for open coding to generate a set of initial codes (e.g., perceived physical harm and discomfort). Differences were resolved through consensus development research meetings, which produced a final set of codes for

coding the rest of the data. During the open coding, saturation was achieved after coding approximately 70 tweets.

Based on the finalized codebook (provided in Supplementary File 1), two authors (LH and CH) independently coded 100 randomly selected tweets to calibrate coding. The inter-rater reliability was 0.87. Then, each of them separately coded an additional set of 200 randomly selected tweets. Thus, in total, 500 tweets were coded and analyzed to answer this research question.

RQ3: Sharing of external information

In this analysis, we extracted all external URLs embedded in the tweets, e.g., websites, images, or videos. We then calculated and compared the proportion of pro- and anti-mask tweets that cited external sources of information. Then, we analyzed the nature of such external information, for example, whether the information originating from public health authorities such as WHO, the CDC, and state- or county-level health departments (based on the URLs, e.g., “cdc,” “who,” “.gov,” “.int”) was cited differently between the pro- and anti-mask groups.

We further did a drill-down analysis through manually reviewing a random set of 100 anti-mask tweets that contained external links. The objective was to specifically investigate what types of external information was used to support the anti-mask attitude. To do this, we read each of these tweets and followed the external links to review and analyze the source information cited (e.g., news articles, journal papers, and videos and images).

RESULTS

A total of 771,268 tweets met our inclusion criteria. Model 1 (*Personal Mask Wearing Classifier*) achieved an F1-score of 84.48% under 10-fold cross validation. After applying this model, 463,369 tweets that were not relevant to mask wearing were removed. The remaining 307,899 tweets were then analyzed using Model 2 (*Opinion Expression Classifier*), the purpose of which was to exclude tweets that did not express a personal opinion, or the opinion was difficult to discern. Model 2 achieved an F1 score of 86.38% under 10-fold cross validation. This model further removed 40,747 irrelevant tweets, leaving a total of 267,152 tweets used in the subsequent Model 3 (*Attitude Classifier*) analysis. Model 3 achieved an F1 score of 90.16% under 10-fold cross validation.

Descriptive Analysis

Figure 2 exhibits the temporal trend of relevant tweets from January 1 to November 1, 2020. Several distinct phases can be observed. Phase I started on February 27, around the time when a statement was issued that the CDC “does not currently recommend the use of face masks,”⁹ and ended around March 30, representing a gradual increase in the volume of relevant Twitter discussions. Phase II lasted till around June 1, representing continued public interests in the subject with some fluctuations in tweet volume. A sharp increase of the number of relevant tweets followed, starting around June 2 and lasting approximately one month till July 6 (Phase III), which may be associated with public reactions toward facial mask mandates by large states such as California. This phase was followed by another dramatic increase of relevant tweet volume (Phase IV), leading toward a second peak around July 13, before returning to the Phase II level, around July 20. The last phase, Phase V, showed a downward trend of relevant tweet volume until the end of October, when the data collection for this study stopped.

[Insert Figure 2 about here]

Table 2 reports the results of a word frequency analysis for examining the evolution of commonly discussed topics across these five distinct phases. Initially, in Phases I (2/27–3/30), the discussions focused on symptoms of COVID-19 and its mechanisms of transmission. In Phases II (3/31–6/1), frequently appearing topics included daily experience coping with the pandemic and the reactions to the recommendation of wearing facial masks by the CDC. Frustration and distress can be observed in Phase III (6/2–7/6), expressing strong sentiments toward mask wearing. This could be associated with many circumstantial factors such as isolation, mask mandates, and a series of protests that broke out across the US. The facial mask discussions continued in Phase IV (7/7–7/20), and in Phase V (7/21–10/31), moved on to focus on school re-opening and the US general election.

Table 2. Evolution of frequently discussed topics over time

Phase	Frequently observed words	Topic
Phase I (2/27–3/30)	symptoms, coughing, sneeze	Symptom
	droplets, mouth, touch, skin, airborne, respiratory, spreading	Mechanism of transmission
	protection, sanitizer, hand, washing, soap, water, sleep, eye, face, wash, clean, mask, scarf	Best practices for personal protection
Phase II (3/31–6/1)	walking, breath, grocery, outside, seeing, wear, people, place, public, stores, shopping, car, talking, shop	Daily experience during the pandemic
	CDC, cloth, mask, bandana, mask, wearing, apart, home, covering, homemade, protect, quarantine	CDC recommendation on facial mask wearing

Phase III (6/2–7/6)	fucking, dumb, stupid, wrong, bad, fuck, hate, damn, hard, selfish, lives, risk	Distress with strong sentiments
Phase IV (7/7–7/20)	people, asthma, mask, mandate, science, enough, children, folks, mandatory, life, rights, never, shut, refuse	Continued discussion on a variety of topics
Phase V (7/21–10/31)	kids, family, wear, school, safe, first, day	School reopening
	vote, trump, tested, distancing	US general election

RQ1a: Attitude

Overall, during the 10-month period, 87.56% of the relevant tweets (233,931) expressed a supportive attitude toward mask wearing whereas 12.44% (33,221) opposed it; the latter is shown as the orange line in Figure 2. This finding confirms our hypothesis that pro-mask tweets would outnumber anti-mask tweets; although the latter were not uncommon. Note that in Figure 2, we only plotted data from February 27 onward as the number of relevant tweets at the initial stage of the pandemic was very small.

RQ1b: Evolution of attitude

As Figure 2 shows, initially, when the CDC recommended against the use of facial masks, the proportion of opposing tweets was around 20%. This was followed by a gradual decrease and around the time when the CDC began to recommend mask wearing (April 3), the proportion of opposing tweets had dropped to about 10% level. Between February and October, there were a number of bursts in the volume of opposing tweets, all of which appear to be associated with state-level mask mandates. The overall opposing rate nonetheless remained around 10% to 15% throughout the study period.

RQ2: Concerns and justifications for opposition

The qualitative content analysis of a sample of opposing tweets revealed six major categories of concerns or justifications for opposing facial masks: (1) physical discomfort or negative effects (30.6%), (2) lack of effectiveness (27.4%), (3) unnecessary or inappropriate for certain people or under certain circumstances (17%), (4) political beliefs (12.2%), (5) lack of mask-wearing culture (9.6%), and (6) coronavirus not a serious threat (3.2%). Table 3 provides more details on each of these categories.

Table 3. Major categories of concerns or justifications for opposing mask wearing (examples paraphrased to protect confidentiality)

Category	Description	Example	Proportion (N=500)
Physical discomfort or negative effects	Perception or experience of discomfort or negative effects as a result of mask wearing such as rash, acne, shortness of breath, or fainting; or beliefs that wearing a mask would cause damage to the immune system.	<i>“It is mandatory to wear a mask at work at my very physical job will cause restrictions to airflow, making it tough to breathe. Is CDC correct or city?!!!! WTF”</i>	30.6%
Lack of effectiveness	Beliefs that mask wearing is not effective as it claims to be, or is not always effective (e.g., if not properly worn), or there are other better alternatives.	<i>“Non-medical face mask made of clothes are not useful for COVID-19 – NAFDAC warns URL”</i>	27.4%
Unnecessary or inappropriate for certain people or under certain circumstances	Beliefs that healthy individuals, children, and/or those with certain health conditions should not wear masks, or that masks are not necessary outdoors or when social distancing is	<i>“If you are not sick, you don’t need to wear a mask.... people are so dumb.”</i> <i>“People who drive wearing a mask alone are the same ones riding their bikes on a</i>	17%

	practiced.	<i>walk/run path. ?”</i>	
Political beliefs	Beliefs that mandatory mask-wearing policies infringe upon personal liberty, or that those mask mandates are politicized and are manipulation tactics by certain politicians and special interest groups	<p><i>“That’s because we don’t want to be forced to wear a mask!!!”</i></p> <p><i>“You’re stupid if you believe that wearing a mask will protect you from COVID-19. ? Anybody would know that this is all about manipulation control and ultimately a part of the deep state genocide! ?”</i></p>	12.2%
Lack of mask-wearing culture	The negative connotations associated with mask wearing such as being odd-looking, “unAmerican,” criminal resembling, or reflective of panic and fear.	<p><i>“I have always been made uncomfortable by someone wearing a surgical mask. Masks make me feel uneasy. This alone is enough to keep me inside. Well done, CDC.”</i></p>	9.6%
Coronavirus not a serious threat	Coronavirus is not a serious threat, or not as serious as what the government suggests, and thus widespread mask wearing is an overreaction.	<p><i>“During flu season you run into many people who have been exposed to the flu without knowing it and in turn expose you and we are not wearing a mask all flu season! Coronavirus is not a new virus, it is just a new strand!”</i></p>	3.2%

As shown in Table 3, the most common concerns expressed in anti-mask tweets are related to “*physical discomfort and negative effects*” (30.6%). These included difficulties in breathing and causing sweaty face and foggy glasses or skin ailments (e.g., rash and acne). A substantial proportion of the opposing tweets also argued that facial masks were not effective in preventing the spread of coronavirus (27.4%), because the particles carrying the virus were too small and ordinary cloth coverings would not be able to stop them from penetrating through. The next notable category is beliefs that facial masks were “*unnecessary or inappropriate for certain*

people or in certain situations” (17%), believing that only those infected with the virus needed to wear masks, or that masks were only useful in settings where social distancing was not possible. Additional reasons for the anti-mask attitude included “*political beliefs*” (12.2%), postulating that mask mandates were unconstitutional that infringed upon one’s personal liberty, and were primarily politicians or the government’s attempt to control the thinking and behavior of the people; “*lack of mask-wearing culture*” (9.6%), e.g., wearing a mask is associated with panic and fear; and “*coronavirus not a serious threat*” (3.2%): which were beliefs that the threat of coronavirus had been intentionally overstated.

RQ3: Sharing of external information

About one-fifth of the relevant tweets (17.69%) included external information (e.g., websites, images, or videos) to support the pro- or anti-mask arguments. Commonly used sources of external information included other tweets (31,619), Instagram images (10,742), YouTube videos (705), the CDC website (189), New York Times (134), and CNN.com (134). Table 4 shows a comparison between the pro- and the anti-mask group. Those opposed mask wearing appear to be much less likely to include external information in their tweets, compared to those supporting. This difference is statistically significant ($p < 0.05$). Further, the opposing group was statistically less likely to use external information from public health authorities such as the CDC, WHO, and other state- or county-level health departments ($p < 0.05$). These findings confirm our hypothesis that anti-mask tweets would be less likely to cite external sources of information especially from public health authorities.

Table 4. Comparison of use of external information among pro- vs. anti-mask tweets

Source	Pro-mask tweets	Anti-mask tweets
Other tweets	12.96%	3.88%
Instagram	4.48%	0.75%
YouTube	0.28%	0.16%
CDC website	0.073%	0.054%
New York Times	0.055%	0.039%
Websites of local public health agencies (e.g., coronavirus.ohio.gov)	0.093%	0.03%
<i>Information sourced from any public health authority</i>	0.178%	0.093%
<i>Total</i>	19.35%	5.98%

Based on the qualitative content analysis that we conducted on a subset of the relevant tweets, we performed a drill-down analysis of the external supporting evidence cited in tweets that opposed mask wearing. The results show that the opposing tweets often cited user-created YouTube videos arguing against mask wearing (e.g., “COVID-19 MASK REFURBISHMENT,”³⁸) and news articles from conservative news media (e.g., Newsmax). While opposing tweets were less likely to cite information disseminated by public health authorities, a few papers published in prestigious medical journals were prominently featured, such as the perspective article “Universal Masking in Hospitals in the Covid-19 Era” published on May 21 in the New England Journal of Medicine stating that, “*We know that wearing a mask outside health care facilities offers little, if any, protection from infection.*”³⁹ Further, the opposing tweets commonly referenced health experts within their social circles to support the anti-mask arguments, e.g., “*A family member of mine is a dr of infectious diseases. Surgical*

mask is 8-10 microns a n95 mask is 4 microns filter size the covid virius is .6 - 1 micron meaning the mask filter is 4-10x bigger then the virius is Covid can enter the body through your eyes and ears But we dont cover those why?"

DISCUSSION

In this study, we analyzed public opinions expressed on Twitter regarding whether to or not to wear facial masks to help to prevent the spread of coronavirus. We studied a total of 771,268 tweets collected from January 1 to November 1, 2020 using an analytical strategy that combined qualitative content analysis for understanding opinions expressed in subtle human language and machine learning for scalability. The results show that while the overall volume of mask-related tweets fluctuated according to real-world events (e.g., WHO/CDC recommendations and mask mandates), the proportion of anti-mask tweets stayed constant at approximately 10%. The top three reasons for opposing public mask wearing were physical discomfort and negative effects, lack of effectiveness, and being unnecessary or inappropriate for certain people or under certain circumstances. The results also show that anti-mask tweets were significantly less likely to use external sources of information to support the arguments, particularly information from public health authorities such as WHO and the CDC.

While there has been a large body of literature analyzing social media data to understand public opinions toward controversial health-related issues, many studies simply applied off-the-shelf sentiment analyzers by equaling the sentiment of an expression to the attitude expressed in the expression,^{40,41} which may lead to incorrect interpretation of the data. Such studies may also suffer from poor domain transferability of the existing sentiment analysis tools, as most of them were developed in non-health domains (e.g., movie reviews).³⁰ In this paper, instead of using the

off-the-shelf tools, we developed a comprehensive computational pipeline that included multiple machine-learning models trained on human-annotated data. These models helped us improve the relevance of data retrieved by keywords search by excluding tweets that were not related to the mask-wearing behavior or did not contain personal opinions. These models also helped us achieve more accurate classification of pro- or anti-mask attitude expressed.

To the best of our knowledge, no studies to date have used social media data to investigate the public's attitude toward mask wearing, except for one medRxiv paper that used unsupervised topic modeling to look at the general discussion trends related to use of facial masks.¹⁹ Conventional surveys available only reported on respondents' pro- vs. anti-mask stances.^{7,9,12} In contrast, our study was able to understand the reasoning of those who were against mask wearing. Further, our study was able to capture the evolution of attitudes over time to reveal the public's reactions to the constantly changing public health recommendations and local, regional, and national mask mandating policies. Understanding such longitudinal trends can be cost-prohibitive to achieve using the conventional survey method. Despite the advantages, use of social media data such as tweets has several limitations. First, the only way to reliably identify tweets posted by users in the US is to use geocodes, which may introduce self-selection bias as not all Twitter users would opt to turn on geo-tracking. Further, the dominance of active/vocal users on Twitter, and on any other social media platforms more generally, may introduce additional self-selection biases in the data.

The results of this study provide several implications for researchers, public health practitioners, and policy makers. First, methodologically, we noticed that how to properly retrieve tweets relevant to mask wearing in the context of the pandemic requires careful consideration. Initially, we followed the common practice used in prior research (e.g., Sanders et

al.¹⁹) in searching for such tweets by including COVID-19 related keywords in addition to mask-related keywords. We found that doing this would result in a loss of nearly 50% of the relevant data, due to the fact that many tweets relevant to COVID-19 did not explicitly use any word related to COVID-19 because of its prevalence in public discourse. Further, as mentioned earlier, we found that commonly used off-the-shelf sentiment analyzers (e.g., VADER, LIWC, Stanford NLP, and TextBlob) failed to produce accurate sentiment classifications, or the sentiments identified were not in accordance with the attitudes expressed. Therefore, we suggest future studies thoroughly compare existing computational tools and if needed, train specialized, domain-specific tools to ensure the validity of study results.

Second, our data analysis revealed several distinct phases of mask-related discussions on Twitter, which closely aligned with real-world events such as shifting recommendations from public health authorities, mask mandates issued at the state level, and other contemporary events such as school reopening and the US general election. This finding confirms previous studies^{41,42} and suggests that social media can be used as a reliable source of information for continuously monitoring the changing public attitude in response to major social, political, and public health events. This may provide implications into designing and implementing a social media-based real-time dashboard to monitor and track public opinions toward important health-related issues, so that health communication strategies could be more targeted and thus effective.

Third, our analysis also revealed common reasons underlying the anti-mask opinions, e.g., perceived physical damage and discomfort. While there was information addressing some concerns on public health authority websites⁴³ and disseminated through traditional news media,⁴⁴ it appears that such information had not become highly visible through social media. Therefore, public health authorities may consider finding more creative and engaging ways to

provide public education and combat with misinformation on social media platforms, e.g., by creating entertaining YouTube or Tik Tok short videos to show how to make facial masks and how to properly wear them.⁴⁵

Further, some users opposed mask wearing because they believed that it was not an effective measure for preventing the spread of the virus and it was unnecessary for healthy individuals. These stances were indeed supported by authoritative bodies such as the CDC and the US Surgeon General at the early stage of the pandemic. However, many of these tweets were posted in June and July, yet the CDC had reversed its position on public mask wearing in April. This finding indicates that certain segments of the US population might not be aware of, or refuse to believe in, new recommendations from the CDC, and that earlier, contradicting recommendations might have a long-lasting impact. This means that public health experts and officials must clearly communicate about scientific uncertainty, reasons for specific recommendations, and the possibility that recommendations could change as more evidence emerges. In addition, public health experts should continue to engage the public so that new scientific findings, conclusions, and recommendations are immediately delivered to all members of the public .

Lastly, it is also interesting to note that while political beliefs are thought to be the key reason for anti-mask opinions,⁷ based on our analysis, it was not the most often cited reason in the anti-mask tweets. Instead, such tweets emphasized physical discomfort and negative effects and lack of effectiveness. However, based on our data, we are unable to determine if these reasons had ties to political beliefs that might not be made explicit in these tweets.

The findings of this study provide several insights into developing better public health communication strategies to convey the benefits of wearing facial masks, using other protective

measures, as well as combating with misinformation. First, transparency is paramount in public health communication on sensitive and/or controversial issues.⁴⁶ This is particularly true for politically charged debates such as mask wearing. As our data show, misinformation about the lack of effectiveness of using facial masks was widespread, and many opposing opinions reflected strong political beliefs that mandatory mask wearing had been used as a tool by certain political groups to control and manipulate the public. For example, recently, the state officials of California refused to disclose the data and reasoning for the state's decision to lift the stay-at-home order, stating that "they rely on a very complex set of measurements that would confuse and potentially mislead the public if they were made public."⁴⁷ This lack of transparency, intentionally or unintentionally, will likely lead to loss of public trust and incubate conspiracy theories. Second, given the complexity, the efficacy of community mask wearing on preventing the spread of SARS-CoV-2 is difficult to definitively prove using scientific methods. As a result, the general public tends to cherry pick the results aligned with their beliefs, or they use cautionary language commonly used in study limitation sections, e.g., "the results of this study may not be generalizable," as proof of lack of evidence. As our data show, some sentences from a handful of studies published in high-impact journals (e.g., the New England Journal of Medicine) were taken out of context and shared widely among those holding anti-mask beliefs. Therefore, in trying times such as this, public health officials need to put an extraordinary effort in educating the public how to properly interpret the findings reported in scientific studies,⁴⁸ in addition to proactively addressing misinformation that may result from inconclusive research findings. Third, policy changes are often inevitable as new scientific evidence emerges. The rationale of such changes must be well articulated to the public. Public health officials should also not be hesitant to admit the mistakes that they might have made at the early stage of this

unprecedented global pandemic due to the lack of information and uncertainties in decision making. Lastly, public health agencies should develop a means to constantly monitor the public's opinions, particularly those circulated through social media, in order to timely adjust their communication strategies in response to viral spread of misinformation, misinterpretation, or misbeliefs.

Future work should develop more effective machine learning classifiers to facilitate opinion mining using social media data, tweets in particular, which are often short and informal, so that automatic and continuous extraction and monitoring of public opinions are possible. In addition, future work should include more diverse social media platforms representing different types of user groups and different interaction modalities; and use qualitative approaches such as interviews and focus groups to obtain a more in-depth understanding of why certain segments of the population have a strong attitude against mask wearing, rather than relying solely on their publicly available social media posts.

CONCLUSION

Public mask wearing, while believed to be an essential personal protection measure to contain the COVID-19 pandemic, has provoked significant controversies in the US. Through an analysis of a large Twitter dataset using a combination of qualitative content analysis and machine-learning approaches, this study classified the public's attitude toward mask wearing and the evolution of this attitude over time. The results show that while most tweets were pro-mask, opposing opinions were not uncommon, and the proportion stayed rather constant throughout the pandemic to date. Common reasons for the anti-mask attitude included physical discomfort and negative effects, lack of effectiveness, and being unnecessary or inappropriate for certain people

or under certain circumstances. Based on these findings, we recommend public health agencies improve their communication strategies to better convey to the public the benefits of mask wearing and combat with misinformation. Such strategies may include increased transparency in data and reasoning, being not afraid of admitting mistakes that might have been made at the early stage of the pandemic due to the lack of information, and educating the public on how to properly interpret inconclusive or conflicting findings from scientific studies.

Figure 1: Method Flowchart

Figure 2: Temporal change of tweet volume and attitude

FUNDING STATEMENT

This work was supported in part by the National Center for Research Resources and the National Center for Advancing Translational Sciences (NCATS), National Institutes of Health (NIH), United States, through Grant UL1TR001414. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

COMPETING INTERESTS STATEMENT

The authors have no competing interests to declare.

CONTRIBUTORSHIP STATEMENT

LH, CH, and YC designed the study. LH annotated data, designed coding scheme, and drafted the manuscript. CH annotated data, built machine-learning models, and drafted a significant portion of the manuscript. TR resolved disagreement during qualitative analysis, assisted in literature review, and revised the manuscript. KZ contributed to the study design and revised the manuscript. YH, QB, and CL collected the data and provided feedback to the manuscript. All authors have read and approved the final draft of the manuscript.

DATA AVAILABILITY STATEMENT

The data analyzed in this paper were retrieved using Twitter API and can be made available upon request, provided that proper data use agreements are established with the research team and with Twitter, Inc (San Francisco, California, United States).

REFERENCE

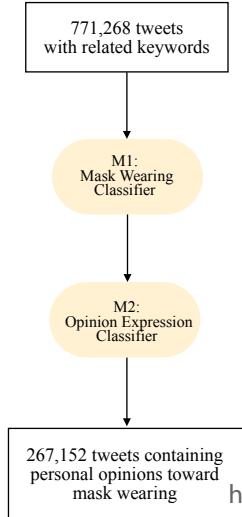
1. CDC. COVID-19 Cases, Deaths, and Trends in the US | CDC COVID Data Tracker. Centers for Disease Control and Prevention. Published March 28, 2020. Accessed December 23, 2020. <https://covid.cdc.gov/covid-data-tracker>
2. WHO Coronavirus Disease (COVID-19) Dashboard. Accessed December 28, 2020. <https://covid19.who.int>
3. Bundgaard H, Bundgaard JS, Raaschou-Pedersen DET, et al. Effectiveness of Adding a Mask Recommendation to Other Public Health Measures to Prevent SARS-CoV-2 Infection in Danish Mask Wearers. *Ann Intern Med*. Published online November 18, 2020. doi:10.7326/M20-6817
4. National Academies of Sciences E. *Rapid Expert Consultation on the Effectiveness of Fabric Masks for the COVID-19 Pandemic (April 8, 2020)*.; 2020. doi:10.17226/25776
5. Eikenberry SE, Mancuso M, Iboi E, et al. To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic. *Infectious Disease Modelling*. 2020;5:293-308. doi:10.1016/j.idm.2020.04.001
6. Sunjaya AP, Jenkins C. Rationale for universal face masks in public against COVID-19. *Respirology*. Published online April 30, 2020. doi:10.1111/resp.13834
7. Katz J, Sanger-Katz M, Quealy K. A Detailed Map of Who Is Wearing Masks in the U.S. *The New York Times*. <https://www.nytimes.com/interactive/2020/07/17/upshot/coronavirus-face-mask-map.html>. Published July 17, 2020. Accessed September 6, 2020.
8. Transcript for CDC Telebriefing: CDC Update on Novel Coronavirus | CDC Online Newsroom | CDC. Published February 12, 2020. Accessed December 1, 2020. <https://www.cdc.gov/media/releases/2020/t0212-cdc-telebriefing-transcript.html>
9. COVID-19: Considerations for Wearing Masks | CDC. Accessed September 28, 2020. <https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/cloth-face-cover-guidance.html>
10. Organization WH. Advice on the use of masks in the context of COVID-19: interim guidance, 6 April 2020. Published online 2020. Accessed November 9, 2020. <https://apps.who.int/iris/handle/10665/331693>
11. August 14 JWR, Economics 2020. Biden's National Mask Mandate is Absurd and Despotic. Discovery Institute. Published August 14, 2020. Accessed December 6, 2020. <https://www.discovery.org/a/bidens-mask-mandate/>
12. Pew Research Center. Most Americans say they regularly wore a mask in stores in the past month; fewer see others doing it. Pew Research Center. Accessed August 2, 2020. <https://www.pewresearch.org/fact-tank/2020/06/23/most-americans-say-they-regularly-wore-a-mask-in-stores-in-the-past-month-fewer-see-others-doing-it/>
13. Why some Americans won't wear face masks, in their own words. Accessed August 2, 2020. <http://www.advisory.com/daily-briefing/2020/06/19/mask-wearing>

14. Why Do Some People Refuse to Wear a Face Mask in Public? | Health.com. Accessed August 2, 2020. <https://www.health.com/condition/infectious-diseases/coronavirus/face-mask-refuse-to-wear-one-but-why>
15. Why Some People Choose Not to Wear a Mask. Office for Science and Society. Accessed November 7, 2020. <https://www.mcgill.ca/oss/article/covid-19-health/why-some-people-choose-not-wear-mask>
16. Here's why some people are not wearing masks during the coronavirus crisis. NBC News. Accessed November 7, 2020. <https://www.nbcnews.com/news/us-news/here-s-why-some-people-are-not-wearing-masks-during-n1200701>
17. Tweet geospatial metadata. Accessed February 1, 2021. <https://developer.twitter.com/en/docs/tutorials/tweet-geo-metadata>
18. Twitter. Twitter API Standard v1.1. Accessed December 31, 2020. <https://developer.twitter.com/en/docs/twitter-api/v1>
19. Sanders A, White R, Severson L, et al. Unmasking the conversation on masks: Natural language processing for topical sentiment analysis of COVID-19 Twitter discourse. *medRxiv*. Published online September 1, 2020:2020.08.28.20183863. doi:10.1101/2020.08.28.20183863
20. Kim Y, Huang J, Emery S. Garbage in, Garbage Out: Data Collection, Quality Assessment and Reporting Standards for Social Media Data Use in Health Research, Infodemiology and Digital Disease Detection. *Journal of Medical Internet Research*. 2016;18(2):e41. doi:10.2196/jmir.4738
21. Mikolov T, Sutskever I, Chen K, Corrado GS, Dean J. Distributed Representations of Words and Phrases and their Compositionality. In: Burges CJC, Bottou L, Welling M, Ghahramani Z, Weinberger KQ, eds. *Advances in Neural Information Processing Systems* 26. Curran Associates, Inc.; 2013:3111-3119. <http://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality.pdf>
22. Hilario AF, López SG, Galar M, Prati RC, Krawczyk B, Herrera F. *Learning from Imbalanced Data Sets*. Springer International Publishing; 2018. doi:10.1007/978-3-319-98074-4
23. Wang Y, Huang M, zhu xiaoyan, Zhao L. Attention-based LSTM for Aspect-level Sentiment Classification. In: *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics; 2016:606-615. doi:10.18653/v1/D16-1058
24. Gal Y, Ghahramani Z. A Theoretically Grounded Application of Dropout in Recurrent Neural Networks. *Advances in Neural Information Processing Systems*. 2016;29:1019-1027.
25. Ferrara E, Varol O, Davis C, Menczer F, Flammini A. The rise of social bots. *Commun ACM*. 2016;59(7):96-104. doi:10.1145/2818717
26. GitHub - IUNetSci/botometer-python: A Python API for Botometer by OSoMe. Accessed February 6, 2021. <https://github.com/IUNetSci/botometer-python>
27. Yang K-C, Varol O, Davis CA, Ferrara E, Flammini A, Menczer F. Arming the public with artificial intelligence to counter social bots. *Human Behavior and Emerging Technologies*. 2019;1(1):48-61. doi:<https://doi.org/10.1002/hbe2.115>

28. Pew Research Center. How we identified bots on Twitter. Pew Research Center. Accessed February 24, 2021. <https://www.pewresearch.org/fact-tank/2018/04/19/qa-how-pew-research-center-identified-bots-on-twitter/>
29. Rauchfleisch A, Kaiser J. The False positive problem of automatic bot detection in social science research. *PLOS ONE*. 2020;15(10):e0241045. doi:10.1371/journal.pone.0241045
30. He L, Yin T, Hu Z, Chen Y, Hanauer DA, Zheng K. Developing a standardized protocol for computational sentiment analysis research using health-related social media data. *Journal of the American Medical Informatics Association*. 2020;(ocaa298). doi:10.1093/jamia/ocaa298
31. Hutto CJ, Gilbert E. VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. In: *Eighth International AAAI Conference on Weblogs and Social Media*. ; 2014. Accessed July 15, 2019. <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8109>
32. TextBlob: Simplified Text Processing — TextBlob 0.15.2 documentation. Accessed March 30, 2019. <https://textblob.readthedocs.io/en/dev/index.html>
33. Manning C, Surdeanu M, Bauer J, Finkel J, Bethard S, McClosky D. The Stanford CoreNLP Natural Language Processing Toolkit. In: *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*. Association for Computational Linguistics; 2014:55-60. doi:10.3115/v1/P14-5010
34. Tausczik YR, Pennebaker JW. The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology*. 2010;29(1):24-54. doi:10.1177/0261927X09351676
35. He L, Zheng K. How Do General-Purpose Sentiment Analyzers Perform when Applied to Health-Related Online Social Media Data? *Stud Health Technol Inform*. 2019;264:1208-1212. doi:10.3233/SHTI190418
36. Tang J, Meng Z, Nguyen X, Mei Q, Zhang M. Understanding the Limiting Factors of Topic Modeling via Posterior Contraction Analysis. *ICML14*':9.
37. Corbin J, Strauss A. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. SAGE Publications; 2014.
38. COVID-19 MASK REFURBISHMENT - YouTube. Accessed December 22, 2020. <https://www.youtube.com/watch?v=LrXLqXydVlo>
39. Universal Masking in Hospitals in the Covid-19 Era | NEJM. Accessed December 3, 2020. <https://www.nejm.org/doi/full/10.1056/nejmp2006372>
40. Davis MA, Zheng K, Liu Y, Levy H. Public Response to Obamacare on Twitter. *Journal of Medical Internet Research*. 2017;19(5):e167. doi:10.2196/jmir.6946
41. Du J, Xu J, Song H-Y, Tao C. Leveraging machine learning-based approaches to assess human papillomavirus vaccination sentiment trends with Twitter data. *BMC Med Inform Decis Mak*. 2017;17(Suppl 2). doi:10.1186/s12911-017-0469-6

42. Gui X, Wang Y, Kou Y, et al. Understanding the Patterns of Health Information Dissemination on Social Media during the Zika Outbreak. *AMIA Annual Symposium Proceedings*. 2017;2017:820.
43. Your Guide to Masks | CDC. Accessed February 13, 2021. <https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/about-face-coverings.html>
44. Why Some People Don't Wear Masks. NPR.org. Accessed February 13, 2021. <https://www.npr.org/2020/07/01/886299211/why-some-people-dont-wear-masks>
45. Watch: Dentist on TikTok shows how to make face mask fit better. Accessed February 12, 2021. <https://www.insider.com/doctor-tiktok-shows-how-to-make-face-mask-fit-better-2020-7>
46. French PE. Enhancing the Legitimacy of Local Government Pandemic Influenza Planning through Transparency and Public Engagement. *Public Administration Review*. 2011;71(2):253-264. doi:<https://doi.org/10.1111/j.1540-6210.2011.02336.x>
47. News ABC. It's a secret: California keeps key virus data from public. ABC News. Accessed February 13, 2021. <https://abcnews.go.com/Technology/wireStory/secret-california-key-virus-data-public-75432129>
48. Gruber-Miller S. Fact check: New England journal article taken out of context, didn't bash face masks. USA TODAY. Accessed February 25, 2021. <https://www.usatoday.com/story/news/factcheck/2020/07/22/fact-check-new-england-journal-medicine-article-face-masks-coronavirus-covid-19-spread/5454384002/>

Data Collection and Pre-processing



Data Analysis for Research Questions

