Examining Public Sentiments and Attitudes Towards COVID-19 Vaccination: An Infoveillance Study using Twitter Posts

Ranganathan Chandrasekaran, Rashi Desai, Harsh Shah, Vivek Kumar, Evangelos Moustakas

Submitted to: JMIR Infodemiology
on: September 29, 2021

Table of Contents

Examining Public Sentiments and Attitudes Towards COVID-19 Vaccination: An Infoveillance Study using Twitter Posts

Ranganathan Chandrasekaran[1] PhD; Rashi Desai[2] MSc; Harsh Shah[2] MSc; Vivek Kumar[2] MSc; Evangelos Moustakas[3] PhD

[1]Department of Information & Decision Sciences University of Illinois at Chicago Chicago US
[2]Middlesex University Dubai AE

Corresponding Author:
Ranganathan Chandrasekaran PhD

Abstract

Background: A global rollout of vaccinations is currently underway to mitigate and protect people from the COVID-19 pandemic. Several individuals have been using social media platforms like Twitter as an outlet to express their feelings, concerns and opinions about COVID-19 vaccines and vaccination programs. This study examines COVID-19 vaccine related tweets from Jan 1, 2020 to April 30, 2021 to uncover the topics, themes and variations in sentiments of public twitter-users.

Objective: To examine key themes and topics from COVID-19 vaccine related English tweets posted by individuals, and to explore the trends and variations in public opinions and sentiments

Methods: We gathered and assessed a corpus of 2.94 million COVID-19 vaccine related tweets made by 1.2 million individuals. We used CoreX topic modelling to explore the themes and topics underlying the tweets, and used VADER sentiment analysis to compute sentiment scores and examine weekly trends. We also performed qualitative content analysis of the top three topics pertaining to COVID-19 vaccination.

Results: Topic modelling yielded 16 topics that were grouped into 6 broader themes underlying the COVID-19 vaccination tweets. The most tweeted topic about COVID-19 vaccination was related to vaccination policy - whether vaccines needed to be mandated or optional (13.94%), followed by vaccine hesitancy (12.63%), and post-vaccination symptoms and effects (10.44%). Average compound sentiment scores were negative throughout the 16 weeks for topics: post-vaccination symptoms and side effects, and hoax/conspiracy. However, consistent positive sentiment scores were observed for the topics: vaccination disclosure, vaccine efficacy, clinical trials and approvals, affordability, policy/regulation,distribution and shortage, travel, appointment and scheduling, vaccination sites, advocacy, opinion leaders and endorsement, and gratitude to healthcare workers. Reversal in sentiment scores in a few weeks were observed for vaccination eligibility and hesitancy.

Conclusions: : Identification of dominant themes, topics, sentiments and changing trends about COVID-19 vaccination can aid governments and healthcare agencies to frame appropriate vaccination programs, policies and rollouts.

(JMIR Preprints 29/09/2021:33909)
DOI: https://doi.org/10.2196/preprints.33909

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Examining Public Sentiments and Attitudes Towards COVID-19 Vaccination: An Infoveillance Study using Twitter Posts

Abstract

Background: A global rollout of vaccinations is currently underway to mitigate and protect people from the COVID-19 pandemic. Several individuals have been using social media platforms like Twitter as an outlet to express their feelings, concerns and opinions about COVID-19 vaccines and vaccination programs. This study examines COVID-19 vaccine related tweets from Jan 1, 2020 to April 30, 2021 to uncover the topics, themes and variations in sentiments of public twitter-users.

Objective: To examine key themes and topics from COVID-19 vaccine related English tweets posted by individuals, and to explore the trends and variations in public opinions and sentiments.

Methods: We gathered and assessed a corpus of 2.94 million COVID-19 vaccine related tweets made by 1.2 million individuals. We used CoreX topic modelling to explore the themes and topics underlying the tweets, and used VADER sentiment analysis to compute sentiment scores and examine weekly trends. We also performed qualitative content analysis of the top three topics pertaining to COVID-19 vaccination.

Results: Topic modelling yielded 16 topics that were grouped into 6 broader themes underlying the COVID-19 vaccination tweets. The most tweeted topic about COVID-19 vaccination was related to vaccination policy - whether vaccines needed to be mandated or optional (13.94%), followed by vaccine hesitancy (12.63%), and post-vaccination symptoms and effects (10.44%). Average compound sentiment scores were negative throughout the 16 weeks for topics: *post-vaccination symptoms and side effects*, and *hoax/conspiracy*. However, consistent positive sentiment scores were observed for the topics: *vaccination disclosure*, *vaccine efficacy*, *clinical trials and approvals*, *affordability*, policy/*regulation*,*distribution and shortage*, *travel*, *appointment and scheduling*, *vaccination sites*, *advocacy*, *opinion leaders and endorsement*, and *gratitude to healthcare workers*. Reversal in sentiment scores in a few weeks were observed for *vaccination eligibility* and *hesitancy*.

Conclusions: Identification of dominant themes, topics, sentiments and changing trends about COVID-19 vaccination can aid governments and healthcare agencies to frame appropriate vaccination programs, policies and rollouts.

Keywords: Coronavirus, Infoveillance, COVID-19, Vaccination, Social Media, Sentiment Analysis, Trends, Topic Modelling

Introduction

Since the outbreak of the COVID-19, caused by the SARS-CoV-2 virus in November 2019, the pandemic continues to pose a serious threat to the lives of millions of individuals around the globe. By June 2021, the virus had infected over 176 millions individuals, resulting in over 3.8 million deaths worldwide[1]. Impact of the pandemic on the world economy, well being and social norms of daily living have been profound. In light of the threats posed by this virus, scientists have been racing to understand the nature of the virus and discover potential treatment regimens and therapeutic mechanisms to deal with it. Though measures such as lockdowns, social distancing and wearing masks have been the primary measures to control the spread of the virus, effective vaccination is likely to constitute a definitive, long-term strategy that can contain the pandemic and help humankind return to normal life[2]. The foreseeable long-term solution to the COVID-19 pandemic is a globally rolled out, safe vaccination program covering substantial portions of the world population. Vaccines can provide both direct protection by minimizing susceptibility to the virus among uninfected, and indirect protection by reducing the virus spread among those infected [3]. Therefore, development and deployment of vaccines have become a central component in the global strategy to control and mitigate the spread of COVID-19 with several billions of dollars spent in R&D and development of the vaccines [4]. In December 2020, US regulatory authorities granted emergency and full authorization for vaccines developed by BioNTech and Pfizer, and Moderna and National Institutes of Health. In August 2021, US Food and Drug Administration provided approval for Pfizer vaccine. Other vaccines that have been granted approvals include the ones developed by University of Oxford and Astrazeneca, Johnson & Johnson, Sinopharm, Sputnik-V, Covaxin, among others. Close to 300 vaccines are currently in different phases of development to tackle the virus and its variants[5,6]. Governments across the world are devising strategies to quickly produce, procure and distribute vaccines to their citizens[7–9].

Social media platforms have become an important conduit and rich source of data for assessing public attitudes and behaviors during health emergencies. In light of the lockdowns and restrictions imposed due to COVID-19 pandemic, social media platforms have emerged as key forums for the public to express their opinions and experiences pertaining to the pandemic and vaccinations. Examination of social media data could reveal significant trends, patterns and changes, and can thus serve as a tool for health surveillance and monitoring the trends. This study builds upon the extant infoveillance research on the COVID-19 pandemic by focusing on the discourse pertaining to COVID-19 vaccinations in Twitter. We analyze over 2.94 million tweets from Jan 1, 2021 to April 30, 2021 to explore the trends, sentiments and key themes pertaining to COVID-19 vaccinations.

There is a growing interest in understanding public attitudes and opinions about COVID-19 vaccinations. A number of studies have employed surveys to examine public willingness, acceptance and hesitancy towards COVID-19 vaccines [10–16]. These studies have used responses from a hundred to few thousand respondents, often from a specific country or region. An alternate *infoveillance* approach using social media data has become a complementary, powerful mechanism to understand and explore public attitudes towards COVID-19 vaccination. A summary of studies using social media data to explore COVID-19 vaccines is provided in Table.1.

Table.1. Summary of key studies on COVID-19 vaccines using social media data

Source	Dataset	Time Period	Key Findings
Yin et al[17]	1.75 million Weibo messages from China	Jan - Oct 2020	Identified public opinions pertaining to pricing, Side effects, and inactivated vaccines
Hussain et al 2021 [18]	23,571 Facebook posts from UK and 144,864 from the USA. 40,268 tweets from the UK and 98,385 from the USA.	March 1 - Nov 22, 2020	Overall averaged positive, negative, and neutral sentiments were at 58%, 22%, and 17% in the UK, in contrast to 56%, 24%, and 18% in the USA. Public optimism regarding vaccine development, effectiveness, clinical trials, concerns over their safety, economic viability, and corporation control were identified.
Guntuku et al.2021 [19]	4 million tweets originating from 2957 US counties.	Dec 1, 2020 - Feb 28, 2021	Topics identified include side effects, conspiracy theories, trust issues in the US healthcare system in December 2020; mask wearing, herd immunity, natural infection, and concerns about nursing home residents and workers in Jan 2021; and access to black communities, vaccine appointments, family safety, and online misinformation campaigns in Feb 2021. Geographic variations on the topics across different counties were also identified.
Bonnevie et al 2021[20]	1,438,251 tweets 6,498 per day.	Anti-vaccine tweets from 15 Feb-2020 to 14 June 2020 as compared to those in Pre-COVID time 15 October 2019 to 14 Feb 2020.	Mentions of vaccine opposition increased by 79.9%. The themes identified were negative health impacts, pharmaceutical industry, policies and politics, vaccine ingredients, federal health authorities, research and clinical trials, religion, vaccine safety, disease prevalence, school and family.

Griffith et al. 2021[21]	3915 tweets about vaccine hesitancy from Canada	December 10, 2020, to December 23, 2020	Vaccine hesitancy was attributed to the following themes: concerns over safety, suspicion about political or economic forces driving the COVID-19 pandemic or vaccine development, a lack of knowledge about the vaccine, anti-vaccine or confusing messages from authority figures, and a lack of legal liability from vaccine companies
Hou et al. 2021[22]	7032 tweets and Weibo posts from five locations: New York, London, Mumbai, Sao Paulo, and Beijing.	June and July 2020	Beijing users (76.8%) had a higher vaccine acceptance rate as compared to the ones in New York (36.4%).Concerns expressed include: vaccine safety, distrust in governments and experts, widespread misinformation, vaccine production and supply, vaccine distribution and inequity.
Yousefinaghani et.al 2021[23]	4,552,652 tweets about COVID-19 vaccines	Jan 2020 - Jan 2021	Sentiment analysis revealed positive being the dominant polarity and having higher engagement. Themes among the positive sentiment tweets were: happiness and hope, support and religion. Themes among the negative sentiment tweets were: fear and frustration, disappointment, anger and politics.More discussion on vaccine rejection and hesitancy as compared to pro-vaccine.
Hu et al. 2021 ([24]	308,755 geo-coded tweets	March 1, 2020 to February 28, 2021	Identified three phases along the pandemic timeline and documented changes in public sentiments and emotions. An increase in positive sentiment, coupled with with a decrease in negative sentiment concerning vaccines were noted in most states. Major international or social events, and announcements by influential leaders or authorities are associated with changes in public opinions towards vaccines.
Lyu et al. 2021[25]	1,499,421 tweets	Mar 11 2020 - Jan 31 2021	16 topics under five broad themes were identified: opinions and emotions around vaccines and vaccination, knowledge around

			vaccines and vaccination, vaccines as a global issue, vaccine administration, and progress on vaccine development and authorization.
Eibensteiner et al. 2021 [26]	Poll of 3439 twitter users	February 12 and February 19, 2021	45.9% of twitter users felt the safety of COVID-19 vaccine to be adequate; Over 82.8% responded affirmatively about taking the vaccination.
Kwok et. Al 2021 [36]	31,100 English tweets from Australian Twitter users.	January and October 2020	3 topics were identified: attitudes toward COVID-19 and vaccination, infection control measures, misconceptions and complaints. Over two-thirds of the tweets had positive sentiment.

The extant studies have collectively helped us uncover some key public concerns and trends regarding vaccinations, vaccine advocacy and hesitancy. However, most of the existing studies have used data from early periods of the COVID-19 pandemic or initial phases of vaccination. Some of these studies have also not differentiated if the source of a tweet is an individual or an organization. Several thousands of tweets typically are made by news outlets, health agencies or other organizations. From an infoveillance perspective, it is critical to examine the social media discourses pertaining to COVID-19 vaccines by the common public, rather than by news agencies or other organizations. Building upon the emerging body of research, our study differs from prior ones in the following ways. First, we focus on tweets made between January and April 2021, capturing public attitudes during active periods of vaccinations in many countries. Second, we examine English language tweets from all over the world, without restricting it to a region or a country. Third, we focus on tweets made by individuals only, thus capturing public sentiments and concerns. Fourth, we use advanced text-mining and topic modelling techniques to unearth themes and topics underlying the twitter discourse on COVID-19 vaccinations.

Our research goals are to

(i) explore the themes and topics underlying social media discourse pertaining to COVID-19 vaccines.

(ii) uncover trends and temporal variations in sentiments underlying COVID-19 vaccine discourse in Twitter

Materials and Methods

Data Gathering

We used a Python scraper *snscrape* to collect historical tweets regarding COVID-19 vaccines and vaccination (available at https://github.com/JustAnotherArchivist/snscrape). Our search terms included a combination of "vaccine" and COVID-19 related terms ("covid", "coronavirus", "covid19", "covid-19", "ncov2019" and "SARS-CoV-2") to retrieve tweets published between January 1, 2021 and April 30, 2021. Snscrape and Getoldtweets are popular python libraries that have been used in several infoveillance studies to capture twitter data[23,27,28]. We ensured removal of retweets and duplicates so that the dataset contained only the original tweets made by the users.

Data Pre-Processing

We used a machine learning approach to separate tweets made by individuals and organizations. Following the approach outlined in [28], we developed a naive-bayes classifier to distinguish the twitter user as being an individual or an organization. The accuracy was 91.81%, providing confidence about the classifier that we used to segregate tweets made by individuals.

Our next step involved pre-processing and cleaning of tweets using a set of libraries in Python. Using *re*, *nltk*, and *sklearn* libraries, we removed punctuations, stop words, emojis and also lemmatized the text of tweets to prepare them for further processing

Topic Modelling and Sentiment Analysis

Topic modeling is an unsupervised machine learning method for identifying latent patterns of words in a large collection of documents. The most representative method for topic modeling is Latent Dirichlet allocation (LDA) which is a generative probabilistic method[29]. LDA does not assume any prior knowledge of topics and through appropriate tuning of parameters, one can explore different topic formations and clusters[30]. Many times, LDA can simply generate topics that can neither be meaningful nor effective.To overcome the restrictions and limitations of LDA, newer algorithms such as Correlation Explanation (CorEx) have been developed[31]. CorEx model, like LDA, does not make any assumptions about topics in underlying data. Further, CorEx identifies latent topics that are maximally informative about a collection of documents by examining how words are used in tweets and picks up on patterns to assess what the tweets convey. CorEx allows a researcher to iterate with different numbers of topics, review them and identify the optimal number of topics for further assessment. CoxEx has been effectively used in a number of health infovellance studies to uncover topics in Twitter data [32,33].

We used CorEx and iterated with a varying number of topics (eg: 5, 10,15,20,30 etc). The total correlation scores were compared across iterations to decide on the optimal number of topics produced. Next, the authors reviewed the results to infer appropriate topics on the basis of keywords. We also examined a set of randomly-chosen tweets for each topic to assess if those tweets were consistent with the topic. Through discussions, the authors then grouped the topics into broader themes. Further, we also computed the sentiment score for each tweet using the VADER (Valence Aware Dictionary and sEntiment Reasoner) tool in Python. VADER is a lexicon and rule-based sentiment analysis tool that is appropriate for social media texts such as Tweets [34]. VADER's

polarity score quantifies the sentiment of a tweet in the range -1(extreme negative) to 1(extreme positive).VADER's scoring method takes into account both the polarity and the intensity of emotion expressed in a tweet. The VADER output labels each tweet into one of the five sentiments namely overly positive (polarity score>=0.70), positive (polarity score between 0.01 & 0.70), neutral (polarity score between -0.01 to 0.01), negative (polarity score between -0.01 & -0.70) and overly negative (polarity score<=-0.70). We used the polarity score to classify the sentiment in the tweets.

In addition to topic modelling and sentiment analysis, we also performed qualitative analysis of tweets in each theme/topic to get further insights and temporal trends in the vaccine related tweets.

Results

Our data gathering resulted in an initial set of 3,707,187 tweets. We removed 762,657 tweets made by organizations. Consistent with our research goal of assessing public sentiments and attitudes, 2,944,530 tweets made by 1,210,225 twitter users were included in our analysis.

The trends in the number of tweets about COVID-19 vaccine from January to April 2021 are presented in Figure.1. All the weeks we examined had over 100,000 tweets, however a spike in the number was observed in the week of March 22 - 31, 2021. This was the week when the eligibility for receiving COVID-19 vaccines was changed to cover several individuals and groups with several US states opening up vaccination to larger sets of individuals.

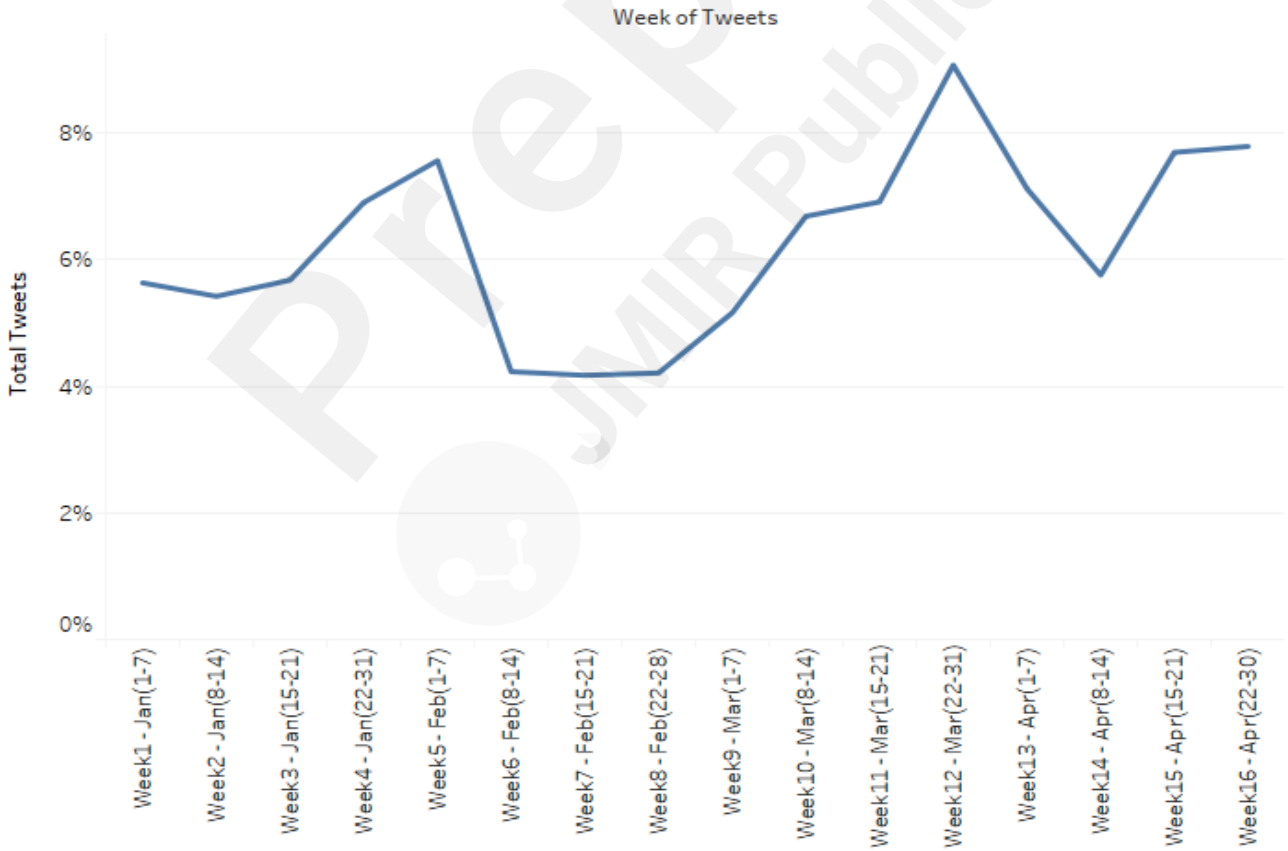


Figure.1 Proportion of COVID-19 Vaccine related tweets from January - April 2021

Themes and Topics

Our CoreX topic modelling resulted in sixteen topics (Table 2), which were further categorized into six broad themes viz. *vaccination experiences* (17.27%), *pharma industry - vaccine development, production and distribution* (15.71%), *vaccination policies* (21.42%), *vaccination rollout* (5.99%), *attitudes towards vaccination* (37.12%), and *gratitude towards healthcare workers* (2.49%). The topics and representative keywords are shown in Appendix A. The top three topics that were tweeted in the January - April 2021 time frame were: *regulatory issues - mandatory vs optional* (13.94%), *vaccine hesitancy* (12.63%), and *post-vaccination symptoms and side-effects* (10.44%).

Table.2. Topics and Broad Themes underlying COVID-19 Vaccine related Tweets

Themes	**Topics**	**Number of tweets (%)**			
Vaccination Experiences				508658	17.27%
	Vaccination disclosure	201102	6.83%		
	Post-vaccination symptoms and effects	307556	10.44%		
Pharma Industry: Vaccine Development, Production and Distribution				462529	15.71%
	Vaccine Efficacy	139280	4.73%		
	Clinical trials, approvals and suspensions	182673	6.20%		
	Vaccine distribution & shortage	140576	4.77%		
Vaccination Policies				630606	21.42%
	Vaccine Affordability	116205	3.95%		
	Regulation - Mandatory vs Optional	410466	13.94%		
	Travel	103935	3.53%		
Vaccination Rollout				176329	5.99%
	Vaccination appointment and scheduling	105586	3.59%		
	Vaccination sites	70743	2.40%		
Attitudes towards Vaccination				1093050	37.12%
	Vaccination eligibility and policies	76605	2.60%		
	Vaccination promotion and advocacy	264368	8.98%		
	Vaccination hesitancy	371843	12.63%		
	Opinion leaders and endorsement	172002	5.84%		

	Hoax/conspiracy	208232	7.07%		
Gratitude towards Healthcare Workers				73358	2.49%
Total				2944530	100.00%

Temporal Trends in Sentiments

We computed the sentiment scores of COVID-19 vaccination tweets and tracked the changes in them over the time period of our study. Results are presented in Figure.2. The proportion of positive or overly positive tweets was always greater than negative or overly negative ones in all the weeks that we examined. Overall, 41.62% tweets had a positive sentiment, 31.16% negative, and 27.22% had neutral sentiment scores.

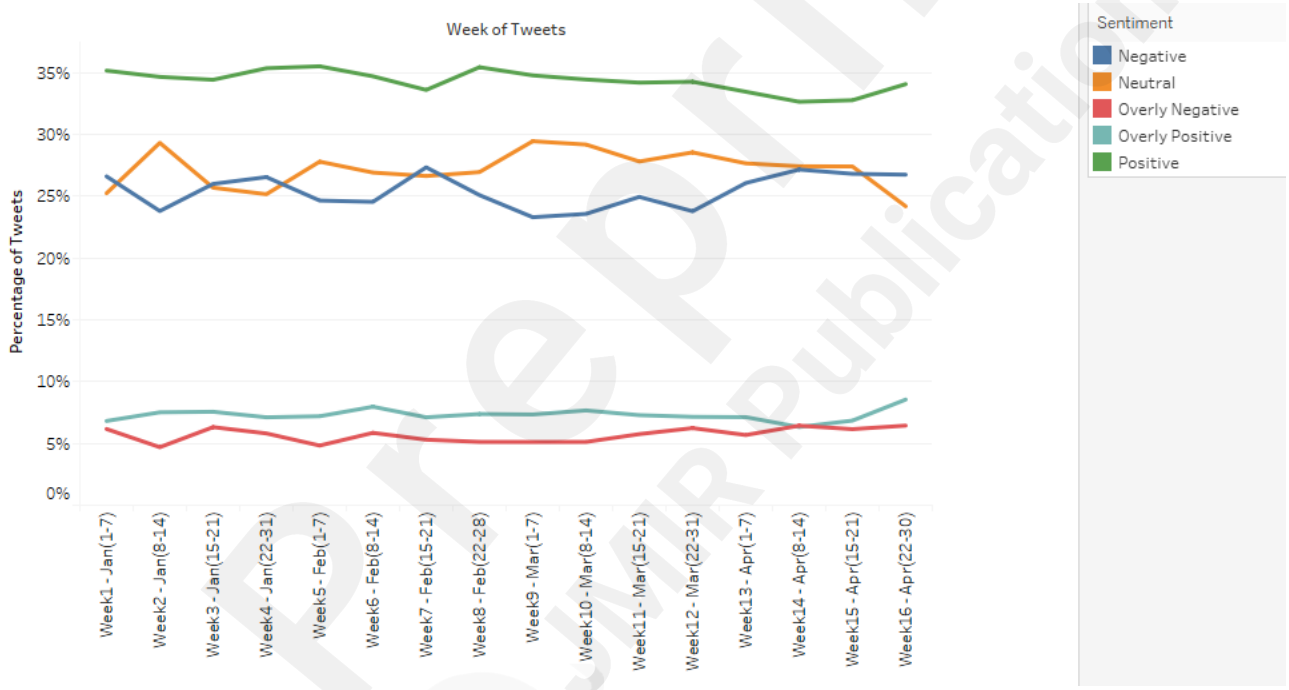


Figure 2. Proportion of Positive, Negative and Neutral Tweets about COVID-19 Vaccination

We further examined the trends in sentiments of sixteen topics over time. These results are presented in Appendix 2. A large proportion of tweets about *post-vaccination symptoms and side effects* (40-45%), and those about *conspiracy/hoax* (35-45 %) had negative or overly negative sentiments in all the weeks of our examination. In contrast, greater proportion of tweets about *vaccination disclosure* (35-40%), *vaccine efficacy* (45-55%), *clinical trials and approvals* (30-40%), *vaccine affordability* (35-35%), *vaccine regulation* (30-35%), *travel* (35-45%), *opinion leaders and endorsement* (30-50%), and *gratitude to healthcare workers* (30-45%) carried positive or overly positive sentiments throughout the time period of our research.

We also examined the trends in average sentiment score for each of the sixteen topics over the time period of examination, and plotted the average compound scores by topic and week. The results are

presented in Appendix 3. Average compound sentiment scores were found to be negative throughout the time period of our examination for the following themes: *post-vaccination symptoms and side effects*, and *hoax/conspiracy*. We found reversal of average sentiment scores from positive to negative during a few weeks for the topic of *vaccination policies*, and *vaccine hesitancy* For the rest of the topics, the average compound sentiment scores were consistently positive for all the weeks.

Qualitative Assessment of Tweet Content

To further examine the public sentiments and attitudes towards COVID-19 vaccines and vaccination rollouts, we qualitatively examined the tweets for the top three themes that emerged from our topic modelling assessment.

Public attitudes towards COVID vaccine regulation : About 14% of tweets about COVID-19 vaccination in the period focused on the issue of whether vaccines need to be made mandatory. Many tweeters argued for mandatory vaccination especially in places of work, schools, education institutions and for travel. For instance, tweets like "*Just like having a vaccination card to go to school, I feel businesses and all schools should make it mandatory to have Covid vaccine*", and "*Would you refuse to take the Covid vaccine; if it became compulsory to work?". "If, eventually, we need to show proof of vaccination to go to theatres, restaurants, sporting events etc. then no, it's not truly optional - by any reasonable measure that's coerced vaccination*". Tweeters also argued for making COVID-19 vaccines mandatory to healthcare workers. Several countries like France have introduced mandatory vaccination requirements for healthcare workers. Saudi Arabia announced that all of the employees in the public, private, and non-profit sectors must be vaccinated before they can return to work. Italy introduced a vaccination requirement for all their healthcare workers and pharmacists[35]. There were many tweets that supported mandatory vaccination: "*I support #MandatoryVaccination for nurses". "Let's keep pushing for #MandatoryVaccination of those who care for our most vulnerable", "Ridiculous that we're making vaccination optional for healthcare workers...vaccinate or GTFO"*. Tweeters opposed to mandatory vaccination opined about how such mandates can be extended to other areas, expressed displeasure as can be seen in the tweets: "*Its all part of the #mandatoryvaccination by coercion agenda. They are going to achieve it by: Divide and Rule -> getting the #vaccinated to blame the #unvaccinated. Threatening people with no sport events pubs etc. These narratives will grow and grow over the coming months. ", "What happens to #MyBodyMyChoice if we're forced into #mandatoryvaccination ? Next it will be #forced #abortion and #sterilization ?"*.

Vaccine hesitancy: About 12.63% of the tweets in our dataset were about vaccine hesitancy that highlighted the reluctance of a set of twitter users to receive COVID-19 vaccines. When we qualitatively examined these tweets, we found tweeters simply spelling out their stance to reject the vaccines, with many users highlighting reasons for not accepting vaccines. Promoting COVID-19 vaccines will need a clear understanding of the reasons the uptake of vaccines (particularly those against COVID-19) will require understanding whether people are willing to be vaccinated, the reasons why they are willing or unwilling to do so. We observed some common reasons cited by twitter users for their vaccine hesitancy. Some users expressed concerns on how quickly the vaccines were developed and wondered about safety. As an instance, a user tweeted "*I don't trust a vaccine that was developed in such a short period of time, when we cant even find one for so many other*

illnesses". Another user tweeted "*I don't trust that jab...its usually years before a vaccine is ready....too rushed..I dont trust it"*. There were others who expressed concerns about effectiveness of vaccines and if the vaccines can protect against newer strains of the virus. As a tweeter said, *"I'm not getting the vaccine. No one knows whats in it or the long term effects of it, or if it can stop new variants"*. From some other tweets, we observed public mistrust of the pharmaceutical industry, medical community and governments. As one user tweeted, "*I don't trust pharma and I won't be having any covid vaccine till it's been around for a while longer and the guinea-pigs have put it to good testing*". Other tweets read, "*I dont trust this vaccine, I dont trust the CDC, I dont trust free donuts from Krispy Kreme (LMFAO), i dont trust our government.*", "*Nope! Not getting the "vaccine". I dont trust the government nor companies who work with the government*.". Another user noted, "*I don't trust pharma and I won't be having any covid vaccine till it's been around for a while longer and the guinea-pigs have put it to good testing.*"

Post-vaccination symptoms and effects: Over 10% of tweets in our dataset were about users' sharing their experiences on symptoms and side-effects of COVID-19 vaccines. Moreover, the average compound sentiment for this topic remained negative throughout the four month period. Twitter users shared information about the dose and their experiences subsequent to vaccination. While some users reported little or no side effects ("*24 hours after my first jab of the Covid-19 vaccine, I have not observed any untoward effect from the vaccine*"), others provided more detailed information on side-effects and how they progressed over a period of time following the vaccination. One user said "*Had the jab at 11am yesterday and the chills & aches started at about 7pm last evening. Lots of Tylenol & fluids*." Another one reported, "*I received my 2nd covid shot yesterday morning. The biggest side effects were weakness and terrible dizziness*.." Another user mentioned, "*Day 2 post-vaccine was no cake walk. Fever, major aches, brain fog, sore everywhere. But man am I glad I got it* !". Mentions of side-effects were often accompanied by messages expressing elevated feelings about protection against the virus. As a tweeter remarked, "*I had side effects from the vaccine, but that 24 hours of chills and fever was worth it to keep myself, friends, family, and my community safe*". Similar remarks were made by another user: "*I would much rather take 48 hours of aches and chills from the second dose of the vaccine than risk gasping for my last breath in an ICU away from family*".

Discussion

A growing number of studies have used data from social media to explore and understand public concerns and attitudes about the COVID-19 pandemic. As governments around the world are trying to tackle the pandemic through mass vaccination, it is important to uncover public opinions and attitudes towards COVID-19 vaccines. We used a repository of approximately 3 million tweets from January 2021 until the last week of April 2021 to uncover the trends in sentiments of various themes and topics pertaining to COVID-19 vaccines . Through topic modelling, we found sixteen topics pertaining to COVID-19 vaccines that were grouped into six broad themes. Further, we examined sentiments associated with these topics and the changes in sentiments over the four month period.

A key finding from our study is that the regulation pertaining to COVID-19 vaccines was the most discussed issue by Twitter users. The number and proportion of tweets on this theme were greater than all the other topics. Proportion of tweets with positive sentiments about regulation of the

vaccination outweighed the proportion of negative and neutral tweets pertaining to this topic. We found vaccine hesitancy to be the second most discussed topic. We also observed negative sentiment scores for many weeks for this topic. Our qualitative analysis provided some preliminary insights into reasons behind vaccine hesitancy: shorter duration of vaccine development cycle, concerns about effectiveness of the vaccine in controlling the virus and its variants, and general mistrust about pharmaceutical and medical industry and governments. Another topic that was widely discussed was post-vaccination side-effects and symptoms. The average sentiment scores for this topic was negative throughout the time period we examined.

In order to control the COVID-19 pandemic, it is important that a substantial portion of the worldwide population acquire immunity through vaccination. Policymakers and public health officials are increasingly focusing on ways to boost and accelerate vaccine uptake. Vaccination campaigns are being designed to address misinformation and public concerns regarding the vaccines. In addition, several efforts are being made to increase vaccine supply, introduce incentive mechanisms for encouraging vaccine uptakes, and enhance public education and outreach programs. However, our findings indicate that vaccine mandates and vaccine hesitancy continue to dominate the minds of the general public, as can be seen from their posts on social media. It is important to take their attitudes into account while framing and designing vaccination campaigns and programs.

It should also be noted that most COVID-19 vaccines have been approved for emergency use and authorization, rather than through a regular licensing route. As more vaccines that are currently authorized for emergency use get regular approval and licenses by authorities like FDA, the issue of vaccine mandates are likely to gain more prominence. More employers and authorities could enforce vaccine mandates. Schools and educational institutions in many parts of the world have started mandating COVID-19 vaccines. Further, vaccination is also a requirement for most international travel. It is more likely to become a requirement for even domestic travel in several countries. A complementary approach to mandating COVID-19 vaccines is creation of trust and favorable attitudes towards vaccines in the minds of the public. Mass outreach and education programs, and incentives for vaccination can go a long way in accelerating the vaccination uptake. Further, endorsement by leaders and celebrities, and experience sharing by peer individuals could also help alleviate concerns regarding vaccines.

Limitations and Future Work

This study used tweets from January 1 to April 31, 2021. Vaccination efforts accelerated in several parts of the world in Summer 2021, and those have not been captured by our study. Another limitation is that we covered only tweets posted in the English language. Due to the nature of data we gathered, we did not explore any geographical disparities in the tweets and this could be a fruitful extension to our work. Another extension of our work would be to examine emotions expressed in tweets pertaining to COVID-19 vaccinations. Another important limitation of our study is that we have captured only the attitudes and opinions of Twitter users, who have a presence in social media. A larger set of population who do not have a presence in Twitter has not been covered by our study.

Conclusion

With variants of COVID-19 virus creating multiple waves of pandemic in several countries, it is

important to accelerate the rate of vaccinations and improve the uptake. As COVD-19 vaccination efforts move forward, it will be important to continue to monitor public opinions regarding vaccine mandates, vaccine hesitancy and vaccination uptake. Some individuals and groups are likely to continue to oppose vaccines, whereas there may be many others who could be convinced by appropriate education and outreach programs. While mandates by governments, or employers could be contested on legal grounds, appropriate exemptions will need to be made for people with certain health conditions or special situations. Infoveillance based on social media data can provide rich insights for policy makers and health officials to frame appropriate policies and programs for COVID-19 vaccination.

References

1. Coronavirus disease (COVID-19) – World Health Organization [Internet]. [cited 2021 Jun 17]. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019

2. Cot C, Cacciapaglia G, Islind AS, Óskarsdóttir M, Sannino F. Impact of US vaccination strategy on COVID-19 wave dynamics. Sci Rep [Internet] 2021 May 26;11(1):10960. PMID:34040088

3. Paltiel AD, Schwartz JL, Zheng A, Walensky RP. Clinical outcomes of A COVID-19 vaccine: Implementation over efficacy. Health Aff [Internet] Health Affairs (Project Hope); 2021 Jan;40(1):42–52. PMID:33211536

4. Bloom DE, Cadarette D, Ferranna M, Hyer RN, Tortorice DL. How new models of vaccine development for COVID-19 have helped address an epic public health crisis. Health Aff [Internet] Health Affairs (Project Hope); 2021 Mar;40(3):410–418. PMID:33539191

5. Shrotri M, Swinnen T, Kampmann B, Parker EPK. An interactive website tracking COVID-19 vaccine development. Lancet Glob Health [Internet] thelancet.com; 2021 May;9(5):e590–e592. PMID:33667404

6. Koch T, Fathi A, Addo MM. The COVID-19 Vaccine Landscape [Internet]. Advances in Experimental Medicine and Biology. 2021. p. 549–573. [doi: 10.1007/978-3-030-63761-3_31]

7. Pagliusi S, Jarrett S, Hayman B, Kreysa U, Prasad SD, Reers M, Hong Thai P, Wu K, Zhang YT, Baek YO, Kumar A, Evtushenko A, Jadhav S, Meng W, Dat DT, Huang W, Desai S. Emerging manufacturers engagements in the COVID -19 vaccine research, development and supply. Vaccine [Internet] Elsevier BV; 2020 Jul 22;38(34):5418–5423. PMID:32600908

8. Yamey G, Schäferhoff M, Hatchett R, Pate M, Zhao F, McDade KK. Ensuring global access to COVID-19 vaccines. Lancet [Internet] thelancet.com; 2020 May 2;395(10234):1405–1406. PMID:32243778

9. Wouters OJ, Shadlen KC, Salcher-Konrad M, Pollard AJ, Larson HJ, Teerawattananon Y, Jit M. Challenges in ensuring global access to COVID-19 vaccines: production, affordability, allocation, and deployment. Lancet [Internet] Elsevier; 2021 Mar 13;397(10278):1023–1034. PMID:33587887

10. Wang K, Wong EL-Y, Ho K-F, Cheung AW-L, Yau PS-Y, Dong D, Wong SY-S, Yeoh E-K. Change of Willingness to Accept COVID-19 Vaccine and Reasons of Vaccine Hesitancy of Working People at Different Waves of Local Epidemic in Hong Kong, China: Repeated Cross-Sectional Surveys. Vaccines (Basel) [Internet] 2021 Jan 18;9(1). PMID:33477725

11. Lazarus JV, Ratzan SC, Palayew A, Gostin LO, Larson HJ, Rabin K, Kimball S, El-Mohandes A. A global survey of potential acceptance of a COVID-19 vaccine. Nat Med [Internet] 2021 Feb;27(2):225–228. PMID:33082575

12. Wong MCS, Wong ELY, Huang J, Cheung AWL, Law K, Chong MKC, Ng RWY, Lai CKC, Boon SS, Lau JTF, Chen Z, Chan PKS. Acceptance of the COVID-19 vaccine based on the health belief model: A population-based survey in Hong Kong. Vaccine [Internet] 2021 Feb 12;39(7):1148–1156. PMID:33461834

13. Edwards B, Biddle N, Gray M, Sollis K. COVID-19 vaccine hesitancy and resistance: Correlates in a nationally representative longitudinal survey of the Australian population. PLoS One [Internet] 2021 Mar 24;16(3):e0248892. PMID:33760836

14. Al-Mohaithef M, Padhi BK. Determinants of COVID-19 Vaccine Acceptance in Saudi Arabia: A Web-Based National Survey. J Multidiscip Healthc [Internet] 2020 Nov 20;13:1657–1663. PMID:33262600

15. Lin Y, Hu Z, Zhao Q, Alias H, Danaee M, Wong LP. Understanding COVID-19 vaccine demand and hesitancy: A nationwide online survey in China. PLoS Negl Trop Dis [Internet] 2020 Dec;14(12):e0008961. PMID:33332359

16. Szilagyi PG, Thomas K, Shah MD, Vizueta N, Cui Y, Vangala S, Kapteyn A. National trends in the US public's likelihood of getting a COVID-19 vaccine-April 1 to December 8, 2020. JAMA [Internet] American Medical Association (AMA); 2020 Dec 29;325(4):396. PMID:33372943

17. Yin F, Wu Z, Xia X, Ji M, Wang Y, Hu Z. Unfolding the Determinants of COVID-19 Vaccine Acceptance in China. J Med Internet Res [Internet] 2021 Jan 15;23(1):e26089. PMID:33400682

18. Hussain A, Tahir A, Hussain Z, Sheikh Z, Gogate M, Dashtipour K, Ali A, Sheikh A. Artificial Intelligence–Enabled Analysis of Public Attitudes on Facebook and Twitter Toward COVID-19 Vaccines in the United Kingdom and the United States: Observational Study (Preprint) [Internet]. [doi: 10.2196/preprints.26627]

19. Guntuku SC, Buttenheim AM, Sherman G, Merchant RM. Twitter discourse reveals geographical and temporal variation in concerns about COVID-19 vaccines in the United States. Vaccine [Internet] 2021 Jun 9; [doi: 10.1016/j.vaccine.2021.06.014]

20. Bonnevie E, Gallegos-Jeffrey A, Goldbarg J, Byrd B, Smyser J. Quantifying the rise of vaccine opposition on Twitter during the COVID-19 pandemic. J Commun Healthc [Internet] Taylor & Francis; 2021 Jan 2;14(1):12–19. [doi: 10.1080/17538068.2020.1858222]

21. Griffith J, Marani H, Monkman H. COVID-19 Vaccine Hesitancy in Canada: Content Analysis of Tweets Using the Theoretical Domains Framework. J Med Internet Res [Internet] 2021 Apr 13;23(4):e26874. PMID:33769946

22. Hou Z, Tong Y, Du F, Lu L, Zhao S, Yu K, Piatek SJ, Larson HJ, Lin L. Assessing COVID-19 Vaccine Hesitancy, Confidence, and Public Engagement: A Global Social Listening Study. J Med Internet Res [Internet] 2021 Jun 11;23(6):e27632. PMID:34061757

23. Yousefinaghani S, Dara R, Mubareka S, Papadopoulos A, Sharif S. An analysis of COVID-19 vaccine sentiments and opinions on Twitter. Int J Infect Dis [Internet] 2021 May 27;108:256–262. PMID:34052407

24. Hu T, Wang S, Luo W, Zhang M, Huang X, Yan Y, Liu R, Ly K, Kacker V, She B, Li Z. Revealing Public Opinion Towards COVID-19 Vaccines With Twitter Data in the United States: Spatiotemporal Perspective. J Med Internet Res [Internet] 2021 Sep 10;23(9):e30854. PMID:34346888

25. Lyu JC, Han EL, Luli GK. Topics and Sentiments in COVID-19 Vaccine-related Discussion on Twitter. J Med Internet Res [Internet] 2021 Jun 10; PMID:34115608

26. Eibensteiner F, Ritschl V, Nawaz FA, Fazel SS, Tsagkaris C, Kulnik ST, Crutzen R, Klager E, Völkl-Kernstock S, Schaden E, Kletecka-Pulker M, Willschke H, Atanasov AG. People's Willingness to Vaccinate Against COVID-19 Despite Their Safety Concerns: Twitter Poll Analysis. J Med Internet Res [Internet] 2021 Apr 29;23(4):e28973. PMID:33872185

27. Ridhwan KM, Hargreaves CA. Leveraging Twitter Data to Understand Public Sentiment for the COVID-19 Outbreak in Singapore. International Journal of Information Management Data Insights [Internet] 2021 Jun 11;100021. [doi: 10.1016/j.jjimei.2021.100021]

28. Chandrasekaran R, Mehta V, Valkunde T, Moustakas E. Topics, Trends, and Sentiments of Tweets About the COVID-19 Pandemic: Temporal Infoveillance Study. J Med Internet Res [Internet] 2020 Oct 23;22(10):e22624. PMID:33006937

29. Blei DM, Ng AY, Jordan MI. Latent Dirichlet Allocation. J Mach Learn Res [Internet] jmlr.org; 2003 [cited 2020 Jul 7];3(Jan):993–1022. Available from: http://www.jmlr.org/papers/v3/blei03a

30. Jelodar H, Wang Y, Yuan C, Feng X, Jiang X, Li Y, Zhao L. Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. Multimed Tools Appl [Internet] Springer; 2019 Jun 1;78(11):15169–15211. [doi: 10.1007/s11042-018-6894-4]

31. Gallagher RJ, Reing K, Kale D, Ver Steeg G. Anchored Correlation Explanation: Topic Modeling with Minimal Domain Knowledge. Transactions of the Association for Computational Linguistics [Internet] MIT Press; 2017 Dec 1;5:529–542. [doi: 10.1162/tacl_a_00078]

32. Zhou S, Zhao Y, Bian J, Haynos AF, Zhang R. Exploring Eating Disorder Topics on Twitter: Machine Learning Approach. JMIR Med Inform [Internet] 2020 Oct 30;8(10):e18273. PMID:33124997

33. Rizvi RF, Wang Y, Nguyen T, Vasilakes J, Bian J, He Z, Zhang R. Analyzing Social Media Data to Understand Consumer Information Needs on Dietary Supplements. Stud Health Technol Inform [Internet] ncbi.nlm.nih.gov; 2019 Aug 21;264:323–327. PMID:31437938

34. Gilbert CHE, Hutto E. Vader: A parsimonious rule-based model for sentiment analysis

of social media text. Eighth International Conference on Weblogs and Social Media (ICWSM-14) Available at (20/04/16) http://comp social gatech edu/papers/icwsm14 vader hutto pdf [Internet] eegilbert.org; 2014. p. 82.

35. Stokel-Walker C. Covid-19: The countries that have mandatory vaccination for health workers. BMJ [Internet] 2021 Jun 29;373:n1645. PMID:34187932

36. Kwok SWH, Vadde SK, Wang G: Tweet Topics and Sentiments Relating to COVID-19 Vaccination Among Australian Twitter Users: Machine Learning Analysis. J Med Internet Res 2021;23(5):e26953

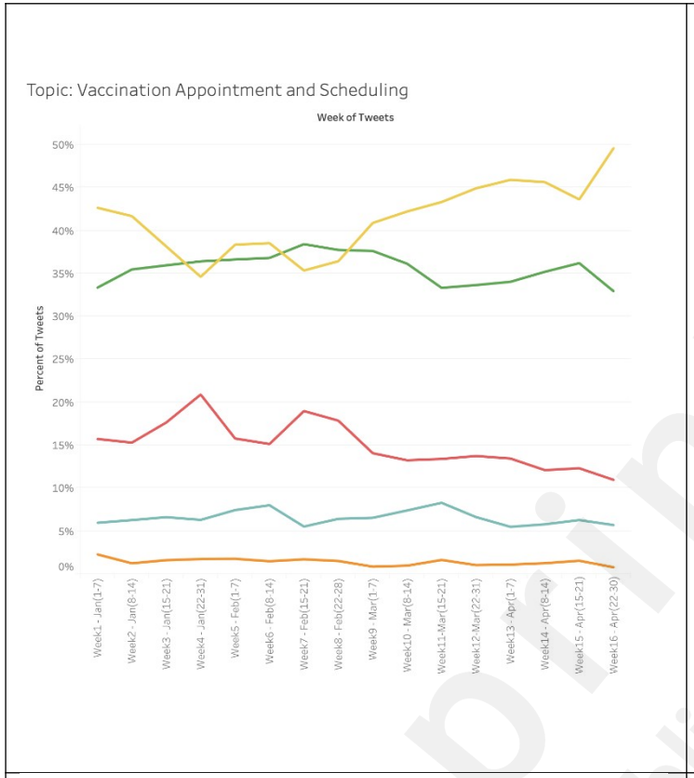
Appendix 1: Themes, Topics and Associated Keywords

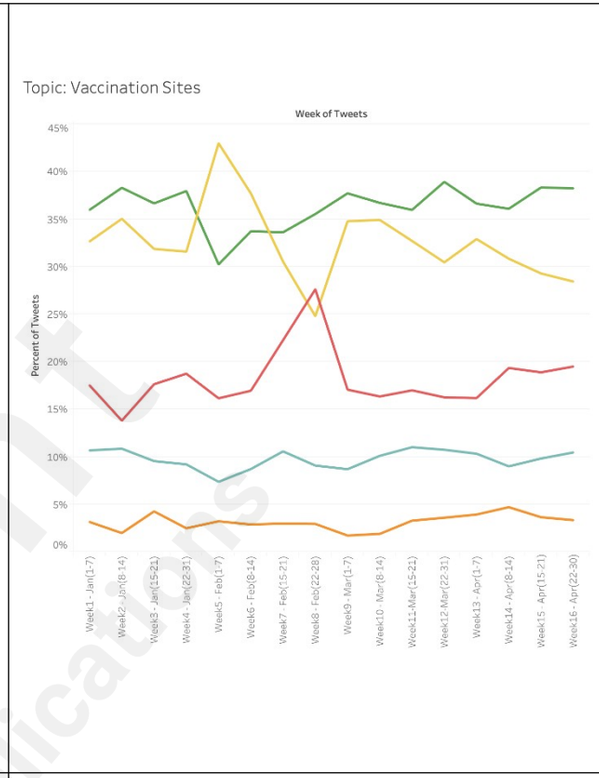
Themes and Topics	Top Keywords
1. Vaccination Experiences	
1.1 Vaccination disclosure	shot, jabbed, got, vaccinated, fullyvaccinated, first dose, second dose, today
1.2 Post-vaccination symptoms and effects	symptom, fever, chills, pain, fatigue, headache, nausea, vomiting, sore, swollen, tired, sleepy, sick, dizziness, numbness, reaction, ache, flu
2. Pharma industry: Vaccine Development and Production	
2.1 Vaccine efficacy	efficacy, effective, compared, efficient, survive, accurate, protects, immune, immunity
2.2 Clinical trials, approvals and suspensions	approval, clinical, trial, rollout, release, study, experiment, emergency, authorized, lawful, regulatory, breakthrough, suspend, license, ban, temporary.
2.3 Vaccine distribution and shortage	distribution, production, logistics, manufacture, export, import, million doses, billion doses, transport, backlog, shortage, scarcity, stock, limited, insufficient, deficit
3. National Policies on Vaccination	
3.1 Vaccine Affordability	free, costly, taxpayer, capped, fee, profits, paid, expensive, gouge
3.2 Regulation - Mandatory vs Optional	optional, required, necessary, protocol, enforce, regulation, certificate, compulsory, obligatory, forced, mandate
3.3 Travel	passport, travel, airline, plane, airfare, visa, flight, train, airport, trip, hotel, vacation, border, sealed, fly, vaccine passport
4. Vaccination Rollout	
4.1 Vaccination Appointment and Scheduling	book, schedule, website, email, walk-in, appointment, register, hotline, reschedule, cancellation, drive-thru, sign-up
4.2 Vaccination Sites	location, miles, drive, county, hospital, far, site, clinic, pharmacy, centers, community
5. Attitudes towards Vaccination	
5.1 Vaccination eligibility and policies	eligible, resident, senior, old, elderly, age, policy, under, young, eligibility
5.2 Vaccination promotion and	getvaccinated, readytovaccinate, vaccineswork,

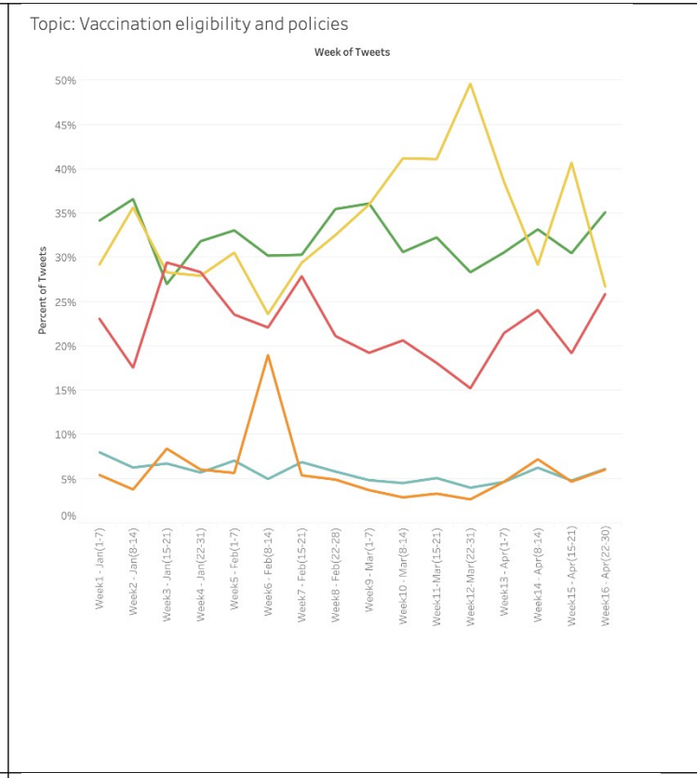
advocacy	vaccinessavelives, saves, advocacy, thisisourshot, urge, vaccinate, vaccinatethem, vaccineforall, encourage
5.3 Vaccination hesitancy	antivaccine, vaccinekills, novaccination, nomandatoryvaccine, justsayno, novaccineforme, boycottvaccine, saynotovaccines, boycott, danger, sayno, nogood, novaccine, refuse, hesitant, vaccinehesitancy
5.4 Opinion leaders and endorsement	Trump, biden, modi, trudeau, joe, desantis, realdonaldtrump, olaf scholz, cuomo, fauci, ministry, president, congress, republican, democrat, potus, celebrity, actor, sports, governor, government, minister, mayor, secretary, leader
5.5 Hoax/conspiracy	fake, steal, hoax, conspiracy, misinformation, scam, fraudulent, forged, rumour, myth
6. Gratitude to healthcare workers	Grateful, healthcare, doctor, nurse, frontline, worker, volunteer, military, physician, staff, thank you

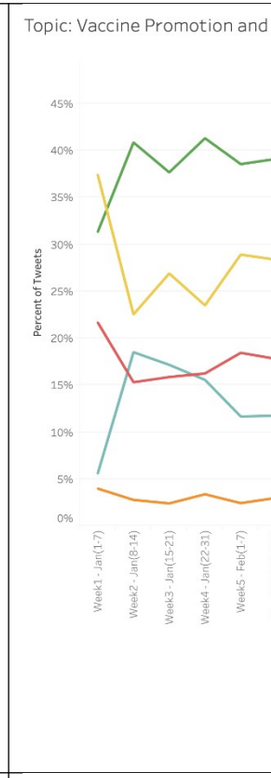
Appendix 2 : Trends in the proportions of positive, neutral, and negative tweets by topic

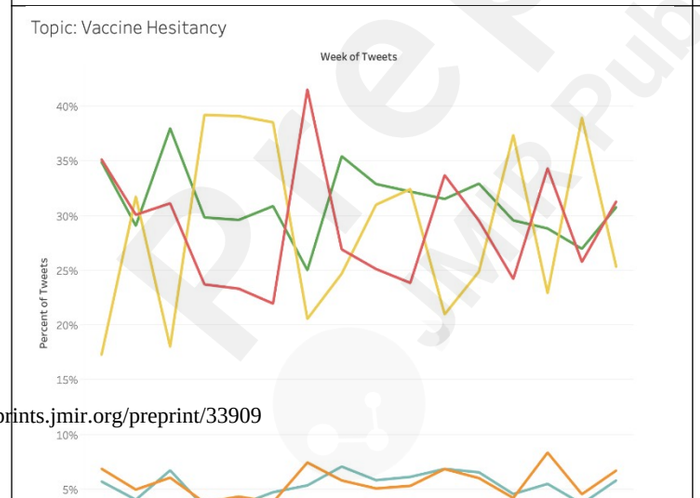
Color legend: green=positive, red=negative, yellow=neutral, teal = overly positive, orange =

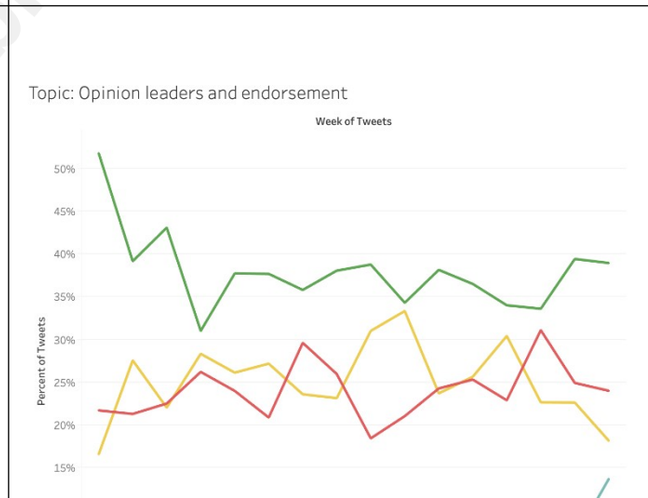


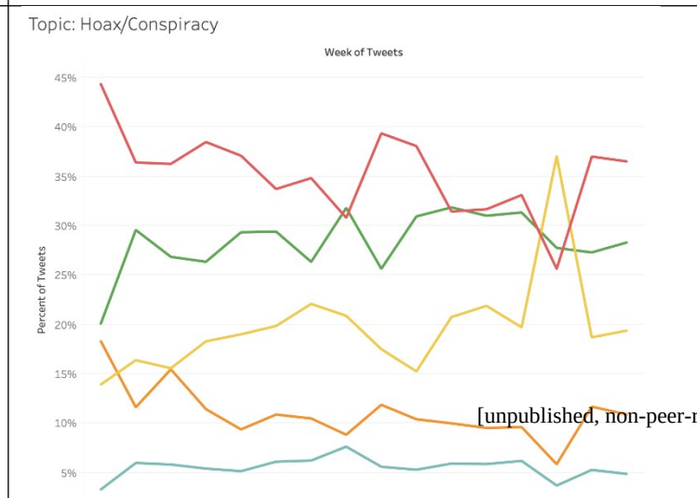


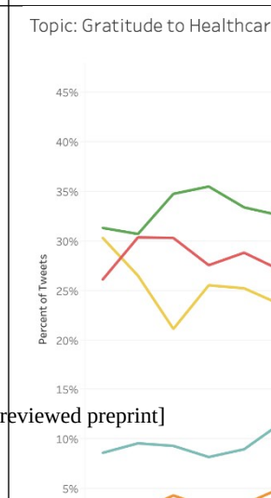




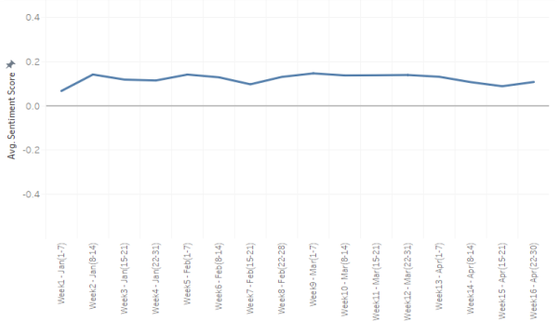




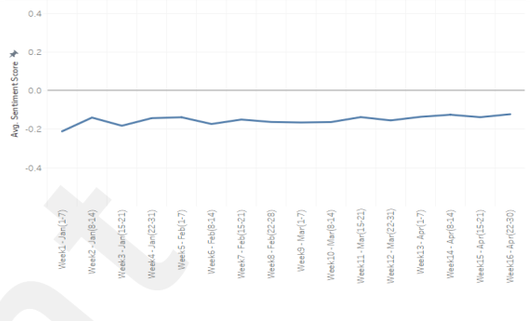




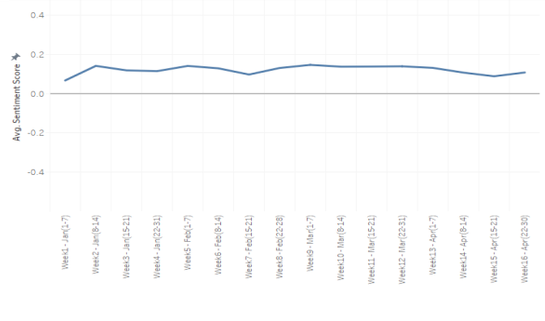
Appendix 3 : Trends in Average Sentiment Score by Topic



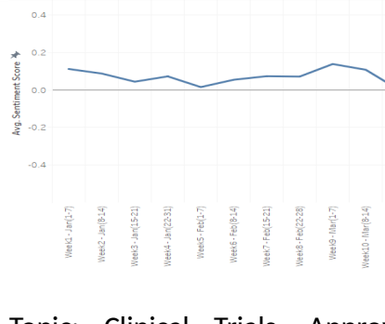
Topic: Vaccine Disclosure (n=201,102)



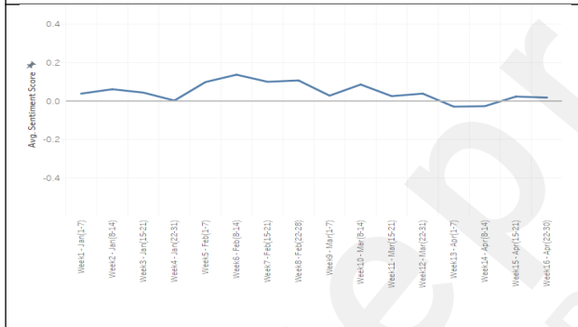
Topic: Post Vaccination Symptoms and Effects (n=307,556)



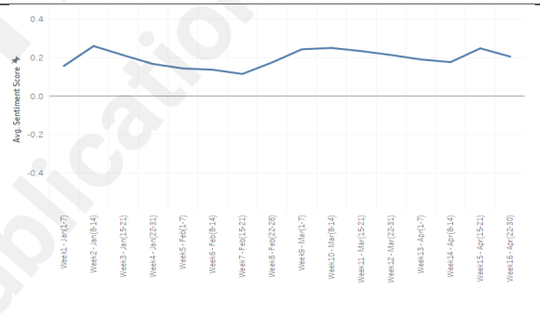
Topic: Vaccine Efficacy (n=139,280)



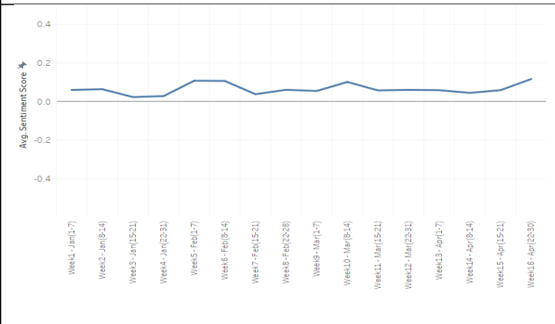
Topic: Clinical Trials, Approv (n=182,673)



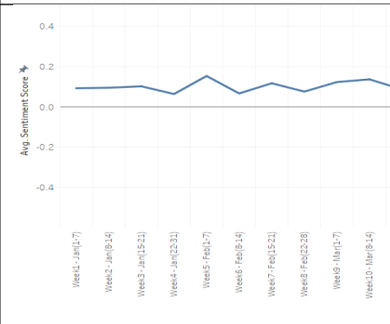
Topic: Vaccine Distribution & Shortage (n=140,576)



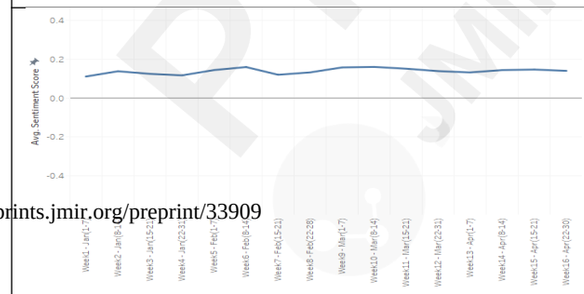
Topic: Affordability of Vaccines (n=116,205)

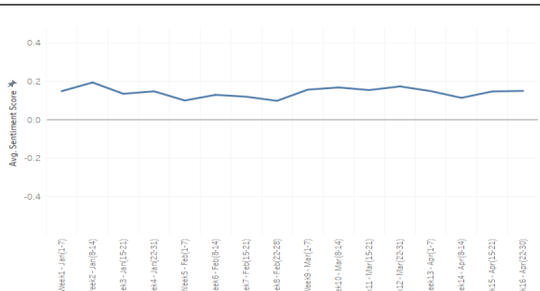


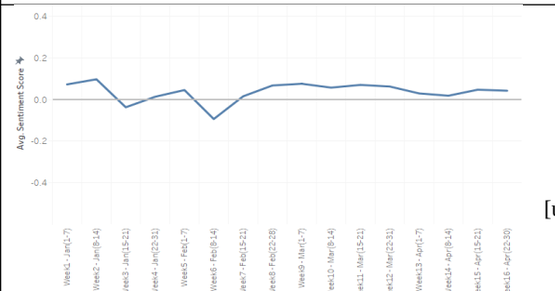
Topic: Regulation - Mandatory vs Optional (n=410,466)

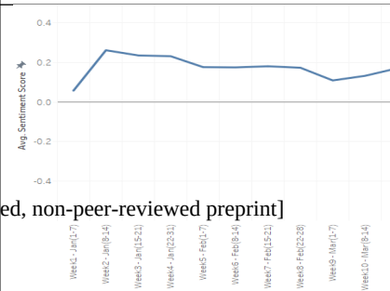


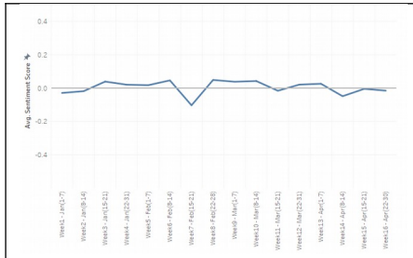
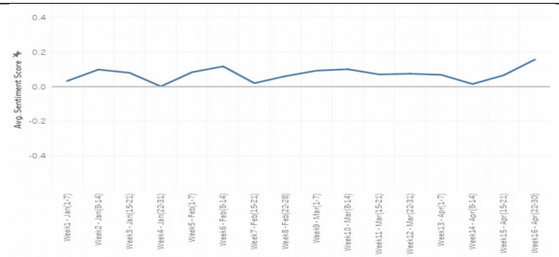
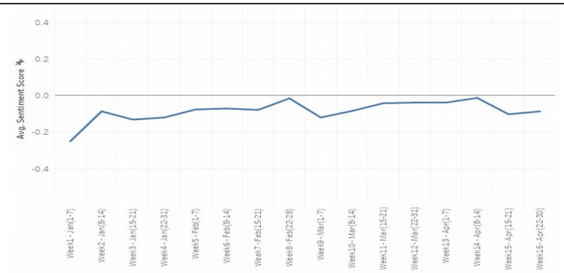
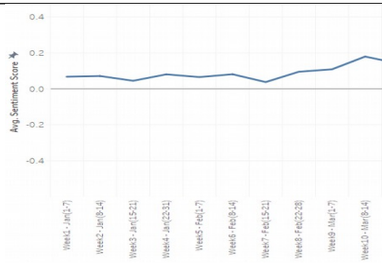
Topic: Travel (n=103,935)









 Topic: Vaccine Hesitancy (n=371,843)	 Topic: Opinion Leaders & Endorsements (n=172,002)	 Topic: Hoax/Conspiracy (n=208,232)	 Topic: Gratitude to Healthcare W