Your Tweets Matter: How Social Media Sentiments Associate with COVID-19 Vaccination Rates in the US

Ana Aleksandric[1,2,] Mercy Jesuloluwa Obasanya MPH[1], Sarah Melcher[1], Shirin Nilizadeh PhD[2], Gabriela Mustata Wilson PhD[1]*

[1] The University of Texas at Arlington, Multi-Interprofessional Center for Health Informatics, Arlington, TX, USA

[2] The University of Texas at Arlington, Department of Computer Science and Engineering, Arlington, TX, USA

Abstract

Objective: The aims of the study were to examine the association between social media sentiments surrounding COVID-19 vaccination and the effects on vaccination rates in the United States (US), as well as other contributing factors to the COVID-19 vaccine hesitancy.

Method: The dataset used in this study consists of vaccine-related English tweets collected in real-time from January 4 - May 11, 2021, posted within the US, as well as health literacy (HL), social vulnerability index (SVI), and vaccination rates at the state level.

Results: The findings presented in this study demonstrate a significant correlation between the sentiments of the tweets and the vaccination rate in the US. The results also suggest a significant negative association between HL and SVI and that the state demographics correlate with both HL and SVI.

Discussion: Social media activity provides insights into public opinion about vaccinations and helps determine the required public health interventions to increase the vaccination rate in the US.

Conclusion: Health literacy, social vulnerability index, and monitoring of social media sentiments need to be considered in public health interventions as part of vaccination campaigns.

Keywords: COVID–19, Health Literacy, COVID–19 Vaccine Hesitancy, Social Vulnerability Index, Social Media, Social Determinants of Health

Abbreviations: Health Literacy (HL), Social Vulnerability Index (SVI), Social Determinants of Health (SDOH), United States (US)

Introduction

Vaccines remain one of the most significant advancements and achievements in public health for disease prevention and control and one of the most cost-effective and successful interventions to improve health outcomes. Unfortunately, vaccine hesitancy from the public is a serious threat to maintaining herd immunity and preventing outbreaks [1]. The delay in acceptance or refusal of vaccines despite the availability of vaccine services is influenced by complacency, convenience, and confidence [2] and may be fueled by health information obtained from various sources, including social media [1]. A national survey in 2020 suggested that the public's willingness to vaccinate against COVID-19 was low and may be insufficient to provide herd immunity [3]. In the US, as of January 2021, about 20% of the population remained hesitant to get the vaccine, and 31% say they will wait to see how it is working for others before getting the COVID-19 vaccine [4].

Although social media is a valuable tool for disseminating and receiving relevant health information for patients, clinicians, and scientists [5], it has also negatively impacted vaccination rates and public health promotion. The substantial spread of negative posts across different social media platforms, such as Twitter and Facebook on the vaccine's safety, has fueled the public's hesitation in getting vaccinated [6]. A relevant example is a study by Ahmed *et al.* that demonstrated that using Twitter and Facebook as sources of information related to the influenza virus has a significant inverse association with influenza vaccine uptake [6]. The anti-vaxxer propaganda has also seen a rise in the community on social media and has potentially magnified the hesitancy [7].

Since the connection between social media and vaccine hesitancy has already been established in previous studies [6], it is hypothesized that a similar pattern occurs related to COVID-19 vaccine uptake. During the COVID-19 pandemic and the country's lock-down, people had to rely on social media to keep social interactions and connections going as they could not do so in person. At the same time, misinformation regarding COVID - 19 emerged in other regions around the world [8]. Unfortunately, this exposed the public to unsubstantial rumors regarding protective measures against the spread of the virus and COVID-19 vaccines. In particular, studies suggest that the use of social media as a source of information about COVID-19 has been correlated with stronger beliefs in conspiracy theories, adverse information, and less-protective behaviors during the pandemic [5]. Some of the false information being spread is that 5G mobile networks were connected to the spread of the virus, that vaccine trial participants have died after taking a COVID-19 vaccine, and that the pandemic is a conspiracy or a bioweapon [9]. As a result, a decreased confidence in the vaccine's efficacy and willingness to take the vaccine once available have been observed. Therefore, it is crucial to understand the impact of social media posts on vaccination rates, which can help identify intervention areas and address misinformation and disinformation. It is also essential to investigate the association between vaccination rates, health literacy, and social determinants of health, which refers to conditions in the places where people live, learn, work, and play that affect a wide range of health and quality-of life-risks and outcomes [10]. These factors could impact both health and health care disparities outcomes [11] and potentially influence vaccination rates in the United States.

While there have been studies/reports on social media and vaccine hesitancy [1,3,5,9], no study has examined the association between sentiments (positive and negative) and vaccination rates using primary data collected from a social media platform. This study addresses this gap by obtaining sentiments of vaccine-related tweets on Twitter and conducting a multivariant regression analysis to examine the impact on vaccination rates. The study also examines the association between social determinants of health factors – health literacy and social vulnerability index, with vaccination rates and analyzes how social media sentiments correlate with COVID-19 vaccine hesitancy and vaccination rates in the United States. It also contributes to a better understanding of the impact of social media and its potential role in healthcare communication and identifies opportunities for interventions such as addressing miscommunications and increasing health literacy to improve health. It also demonstrates the importance of incorporating social determinants of health factors (health literacy and social vulnerability index) in public health interventions in the future.

Methods

Data Sources

The various data sources used in this study are listed in Table 1.

Table 1: **Data sources used in the analysis**

Data	Data Source
Tweets	CoVaxxy Dataset. http://arxiv.org/abs/2101.07694
Health Literacy	The University of North Carolina at Chapel Hill. http://healthliteracymap.unc.edu/#
Social Vulnerability Index	CDC https://www.atsdr.cdc.gov/placeandhealth/svi/index.html
Population per census block group	Policy Map https://www.policymap.com/newmaps#/
Population per state	Policy Map https://www.policymap.com/newmaps#/
Population Percentages	United States Census Bureau. https://www.census.gov/data/tables/time-series/demo/popest/2010s-state-detail.html

Study Design

This study focuses on one of the most popular microblogging and social networking service platforms, Twitter [12]. The dataset used herein comes from the CoVaxxy dataset [13], an extensive collection of vaccine-related tweets collected in real-time between January 4 - May 11, 2021, when the COVID-19 vaccine was approved and vaccination started. The primary goal of this research was to determine how tweets sentiments impacted the vaccination rates in the United

States and determine what other factors, such as health literacy (HL) and social vulnerability index (SVI), might influence vaccination rates. The study was also performed to find the association between HL and SVI, as well as between HL, SVI, and state demographics (percentages of White, Black, Hispanic, and Asian populations per state). As the assumption is that HL and SVI represent significant variables that should be taken into consideration when investigating the association between social media sentiments and vaccination rate, it is relevant to explore the relationship between them, as well as their relationships with demographic composition. Thus, examining such associations can help understand social media activity and its influence on the vaccination rate.

Study Population

The dataset was comprised of tweets from 51 individual states within the US, and from the collection of 85,100,935 vaccine-related tweets between January 4-May 11, 2021, 322,035 tweets were extracted with the specified location within the United States, belonging to 117,258 unique users. Twitter provides location as a bounding box of four points, each having longitude and latitude. However, for the purpose of this analysis, it was important to identify the state where the tweet was posted from. Therefore, the middle point of the bounding box (see Figure 1) was selected, and the Federal Communication Commission getArea API [14] was used to determine the state of every tweet. Not all tweets include such information; therefore, the tweets that received 'None' as a location were discarded from the dataset.
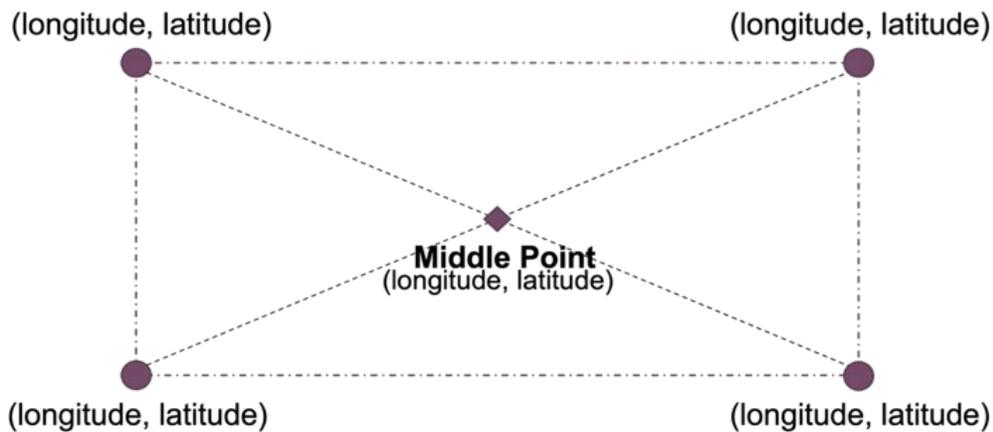

Figure 1: Tweet's bounding box location visualization; computing the middle point

One caveat is that bots can automatically generate tweets, which might not represent the opinion of people. Therefore, to remove any bias from the dataset, Botometer v4 [15] was used to remove any tweets belonging to the bot accounts. According to Botometer creators, between 9 and 15 percent of Twitter accounts are bot accounts, while 85 percent of the accounts are human [16]. Therefore, the 85th percentile of scores was computed on all the accounts scores in the dataset, being 0.69, which is then used as a threshold. Afterward, every tweet belonging to the account with a score higher than the threshold was removed from the dataset. This led to the removal of 58,263 tweets belonging to 17,597 users. Secondly, since the primary purpose of this research was to establish a connection between tweet sentiments and vaccination rates, all the tweets originating from the accounts representing organizations (e.g., pharmaceutical companies) were removed

using Humanizr, a tool that can distinguish between personal and organizational Twitter accounts considering different account features [17]. The filtered dataset included the total number of 243,202 tweets belonging to 95,292 unique users. Figure 2 highlights the process of data collection and the cleaning process.

Figure 2: Data Collection and Cleaning Process

Dependent and Independent Variables

Vaccination rate: The main dependent variable in this analysis is the vaccination rate, as the objective is to see if tweets' sentiments and social determinants of health impact the vaccination rate in the United States. The vaccination rate shown in Figure 3 is per hundred people on the last day of our data collection, May 11, 2021. The vaccination rate was extracted from COVID-19 vaccine data, collected daily in real-time by Indiana University's Observatory on Social Media (OSoMe) in collaboration with colleagues from Politecnico di Milano [13].
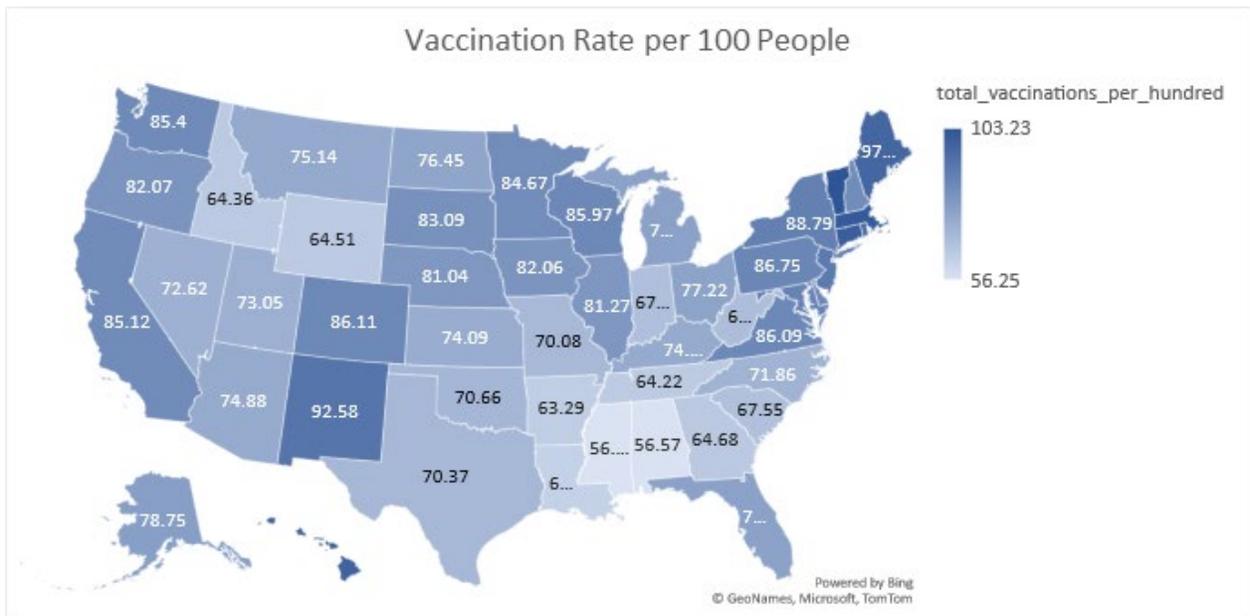

Figure 3: Vaccination rate in the United States per 100 people as of May 11, 2021 (figure generated with Microsoft Excel)

Tweets' sentiments: The sentiment of each tweet was extracted using VADER [18], a machine learning tool that allows for deriving sentiments of social media posts in multiple languages. VADER showed as the most accurate tool used for sentiment analysis on Twitter data [19] due to its ability to take the social media language common features into consideration. This tool accepts a tweet text as an input and returns a compound score in the range from –1 to 1, where scores closer to 1 are more positive, and scores closer to -1 are more negative.

Health literacy: Health literacy is the degree to which individuals have the ability to find, understand, and use information and services to inform health-related decisions and actions for themselves and others [20]. Health literacy was used in the analysis to provide a better understanding of how the ability of the community to receive and understand health information is impacting the public health interventions. The assumption is that vaccine hesitancy is lower in the communities with high health literacy and higher in the communities with lower health literacy. Several previous studies have been conducted to examine the connection between health literacy and vaccination hesitancy. One study in France explored the COVID - 19 vaccine and vaccine hesitancy in relation to fake news. This study suggests that there is an association between low health literacy levels and vaccination rates [21]. A different study conducted in the US found that individuals with low health literacy would be less willing to believe credible sources and less likely to receive a vaccine [22]. Finally, a systematic review of health literacy and vaccination rates indicated that there could not be a universal standard as many different factors come into play for each study [23].

Health literacy data was obtained from the University of North Carolina at Chapel Hill at the census block level [24]. A weighted average of health literacy indexes having census group population as weights was performed to calculate health literacy at the state level. The census group population was based on the 2010 census since health literacy data corresponds to 2014, which means health literacy data is calculated based on the 2010 census group data.

Social vulnerability index: The social vulnerability index refers to the potential negative effects on communities caused by external stresses on human health [25]. Such strains include natural or human-caused disasters or disease outbreaks. A low index means communities are especially at risk during public health emergencies because of socioeconomic status, household composition, minority status, or housing type and transportation [25]. The Center for Disease Control and Prevention (CDC) noted that areas with a high social vulnerability index had lower vaccination rates [26]. The Social Vulnerability Index (SVI) data used in this study originate from the CDC's SVI at the county level. Similarly, as in the case of HL, a weighted average of social vulnerability indexes was computed, having county population as weights to obtain SVI on the state level.

Given that population data were unavailable for the Oglala Lakota and Kusilvak, these counties were excluded from the calculation of weighted SVI per state. The latest SVI dataset available is from 2018, which means it was also calculated per 2010 census block groups. As specified earlier, the 2010 population data were used to calculate the weighted average to obtain SVI per state.

Control Variables

The control variables used in the statistical analysis of this study are at the state level and include the following: *population*, *the total number of tweets*, *number of unique users*, and *state race and ethnicity composition (%)*. For example, some states might have a higher population, and with that, it is very likely that more tweets will be collected in these states. These variables were included to avoid any bias that could appear due to the different numbers of tweets/sentiments in various states. State demographics include percentages of White, Black, Hispanic, and Asian populations per state, and they were also included to test if race and ethnicity play a role in the overall vaccination rate, HL, and SVI. Population and population percentages were calculated based on 2010 census group data since HL and SVI were also calculated per census 2010 data (2020 data was not available when HL and SVI were calculated).

Statistical Analysis

Various hypotheses were established to be tested by statistical analysis, such as:

Hypothesis 1: Positive tweets are associated with a higher vaccination rate at the state level.

Hypothesis 2: Social determinants of health play a significant role in vaccine rate.

Hypothesis 3: A higher social vulnerability index is associated with lower health literacy.

Hypothesis 4: Percentages of minority populations are lower in states where health literacy is higher.

Hypothesis 5: Percentage of minority populations is higher in states where the social vulnerability index is more elevated.

The analysis consists of different statistical tests to evaluate each hypothesis. To test the first two hypotheses, a multivariant linear regression model was used to predict vaccination rate as a dependent numerical variable and tweet sentiments, health literacy, and social vulnerability index as independent variables. All the control variables mentioned were included in the models. For each user in the dataset, an average sentiment was calculated for that user per week. For that particular user the following data was available: location at the state level from which tweets were generated, HL & SVI for that state, vaccination rate per state for that particular week, that state race and ethnicity composition (percentages of White, Black, Hispanic and Asian population), population, the total number of tweets, and the number of unique users. The third hypothesis was also tested using multivariant linear regression, with SVI as dependent and HL as independent variables. Finally, the last two hypotheses were also tested using two separate linear regression models, one containing HL as dependent variable, and the second one containing SVI as dependent variable. In both models, independent variables were state race and ethnicity composition (%).

Results

Association Between Vaccination Rate and Tweet Sentiments and Impact of SDoH on Vaccination Rate

The linear regression analysis showed a strong positive correlation between vaccination rate and tweets sentiments with the p-value < 0.05. This result supports **Hypothesis 1,** that positive tweets are associated with a higher vaccination rate. The results in Table 2 show that both health literacy (HL) and social vulnerability index (SVI) play a significant role, with HL yielding a strong positive correlation while SVI is yielding a negative correlation. This means that for each unit increase of health literacy, the vaccination rate increases by 0.458 (p-value < 0.001). Similarly, for each SVI unit increase, the vaccination rate decreases by 14.1 (p-value < 0.001). These results support **Hypothesis 2** that social determinants of health play a significant role in vaccination rates. Note that a similar analysis was performed using average tweet sentiments per user per day, and similar results were obtained.

Table 2: **Association between vaccination rates and selected independent variables**

INDEPENDENT VARIABLES	ESTIMATE	STD. ERROR	T-VALUE	P-VALUE
Sentiment	0.2734	0.1388	1.970	P < 0.05
Health Literacy	0.4578	0.0586	7.812	P < 0.001
Social Vulnerability Index	-14.0801	1.2823	-10.980	P < 0.001
White Population (%)	-9.4716	3.6687	-2.582	P < 0.05
Black Population (%)	3.1385	3.4969	0.897	P > 0.05
Hispanic Population (%)	27.1188	1.6105	16.838	P < 0.001
Asian Population (%)	31.0692	6.1772	5.030	P < 0.001
Population	0.0000	0.0000	4.384	P < 0.001
Number of Unique Users	-0.0042	0.0003	-16.480	P < 0.001
Total Number of Tweets	0.0013	0.0001	20.105	P < 0.001

Association Between HL and SVI

Another linear regression model was performed to test **Hypothesis 3**, having HL as explanatory variable and SVI as response variable. The results suggest that there is a significant negative correlation (P < 0.001) between HL and SVI. Thus, the social vulnerability index is lower in the states where health literacy is higher, suggesting that the areas with a higher ability to receive and understand health information are less vulnerable. This finding implies that public health interventions need to be performed in more vulnerable areas to educate communities and increase

the vaccination rates. The relationship between HL and SVI is presented in Figure 4, confirming **Hypothesis 3**.
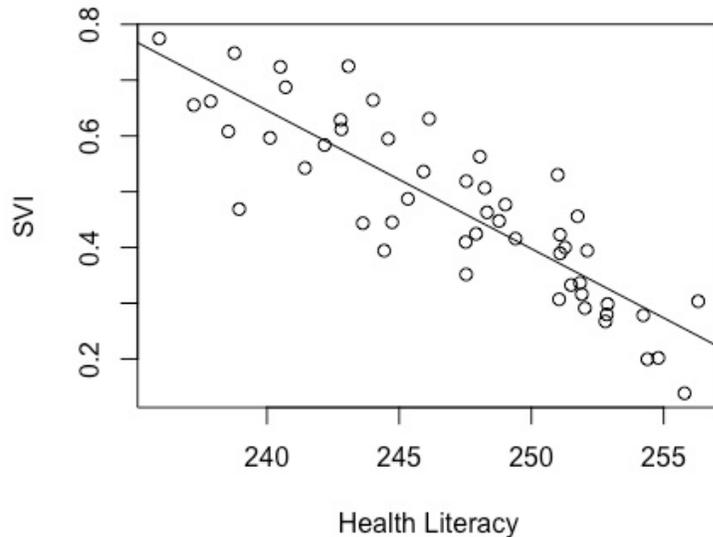

Figure 4: Linear relationship between SVI and HL

Association Between HL/SVI and State Race and Ethnicity Composition

Final linear regression models were used to test **Hypothesis 4** and **Hypothesis 5**, where dependent variables were HL in the first model, and SVI in the second model. The independent variables in both cases were state race and ethnicity composition (percentages of White, Black, Hispanic, and Asian population). The result confirms **Hypothesis 4** by indicating the negative significant association between HL and percentages of Black and Hispanic population per state. Meaning, the HL is lower in the states with a higher percentage of Hispanic ($p < 0.001$) and Black ($p < 0.05$) populations. Similarly, there is a significant negative relationship between SVI and White and Asian population percentage, and a significant positive correlation between SVI and Hispanic population percentage per state, confirming **Hypothesis 5**. This means that states with a higher social vulnerability index have a higher percentage of the Hispanic population ($p < 0.001$), while states with a lower SVI have a higher percentage of White ($p < 0.05$) and Asian ($p < 0.05$) populations. This suggests that public health organizations need to dedicate resources to improve health education and increase vaccine awareness in vulnerable areas where there is a higher percentage of minority populations. Results of the models are presented in Table 3.

Table 3: **Association between HL (above) and SVI (below) with state race and ethnicity composition**

INDEPENDENT VARIABLE (HL)	ESTIMATE	STD. ERROR	T-VALUE	P-VALUE
White Population (%)	13.477	9.231	1.460	P > 0.05
Black Population (%)	-20.556	8.634	- 2.381	P < 0.05
Hispanic Population (%)	- 31.423	3.673	- 8.555	P < 0.001
Asian Population (%)	- 3.839	16.750	- 0.229	P > 0.05
INDEPENDENT VARIABLE (SVI)	ESTIMATE	STD. ERROR	T-VALUE	P-VALUE
White Population (%)	- 0.9156	0.3830	- 2.390	P < 0.05
Black Population (%)	- 0.0199	0.3583	- 0.056	P > 0.05
Hispanic Population (%)	0.8704	0.1524	5.711	P < 0.001
Asian Population (%)	- 1.5058	0.6950	- 2.166	P < 0.05

Discussion

The purpose of this study was to examine the association between social media sentiments surrounding COVID-19 vaccination, the association between sentiments and vaccination rates in the US, and the importance of other social determinants of health factors that contribute to the COVID-19 vaccine hesitancy. The most important finding of this study from the public health perspective is the association between the state's vaccination rate and the tweets' sentiments. This indicates that social media could provide helpful information on vaccine acceptance, informing policymakers on what type of message would be beneficial for public health interventions on social media platforms. The use of social media by public health professionals could increase vaccine awareness and provide more detailed information about vaccination, contributing to the growth of vaccination rates. In addition, the analysis of social media activity could give early warnings about disease outbreaks, which public health organizations could use to prepare and guide their region/community-specific interventions proactively.

The potential future work would be identifying areas where social media sentiments are more negative than in others while analyzing the social vulnerability index, health literacy, demographics, and vaccination rates in these areas. Once those areas are determined, the next step would be to identify where health information can be distributed. One helpful tool to assist in this effort would be the Health Intelligence Atlas, a dashboard with multiple layers created by the research team that includes locations of public health agencies, libraries, places of worship, medically underserved areas and populations, etc.) [27]. Using a similar approach, it would be possible to identify vulnerable areas and perform public health interventions to increase the

vaccination rates in these areas. This would be beneficial because the data found in this study can be added to this dashboard which might be used for future public health emergencies.

The study described herein offers a unique perspective using the social vulnerability index and health literacy to understand better the association between vaccination rate, social media sentiments, and these other potential factors that might impact them. The main strength of this study is that it was conducted on one of the most popular social media platforms, Twitter, a platform that can represent the public's opinions on vaccination. Another strength is that both tweets and vaccination rates data were collected daily in real-time, giving insight into trends surrounding vaccination and providing value to the analysis described in the study. Finally, all datasets were extracted from credible sources such as the CDC and the University of North Carolina. With the significant results from the analysis, this study is not without limitations: (1) the dataset consists of English tweets only, which might not reflect the vaccine hesitancy that is being discussed by minorities speaking a different language (e.g., Spanish, Vietnamese, etc.); (2) even though all the posts are related to the vaccination, some tweets might not represent an actual opinion of the user about it. Despite the limitations mentioned above, this study shows practical applications and results that are critical to future public health interventions.

Conclusion

The primary purpose of this observational study was to investigate if any correlation exists between social media sentiments and the COVID-19 vaccination rates in the United States and identify other factors that impact the vaccination rate in the US. The regression analysis showed that social media sentiments are significantly associated with the US COVID-19 vaccination rate. In addition, health literacy and social vulnerability index play an essential role in the COVID-19 vaccination rate. Social media might be used as an effective tool to increase the overall acceptance of public health interventions, such as the COVID-19 vaccination. Acknowledging that there are limitations, the results shown are relevant for future interventions and should be considered. Public health professionals should incorporate social media listening tools to analyze and address the spreading of negative posts about health interventions on social media and improve health literacy in socially vulnerable areas. As shown in this study, higher health literacy and more positive social media sentiments are significant factors in increasing the vaccination rate in the US. Therefore, social media is a reliable resource for informing the population about emerging threats/events and interventions, which is crucial in reducing misinformation and disinformation and building trust in the public health system.

References

1. Puri N, Coomes EA, Haghbayan H, Gunaratne K. 2020. Social media and vaccine hesitancy: new updates for the era of COVID-19 and globalized infectious diseases. *Hum Vaccin Immunother.* 16(11), 2586-93. doi:https://doi.org/10.1080/21645515.2020.1780846. PubMed

2. MacDonald NE, & SAGE Working Group on Vaccine Hesitancy. 2015. Vaccine hesitancy: definition, scope and determinants. *Vaccine.* 33(34), 4161-64. Epub Apr 2015. doi:https://doi.org/10.1016/j.vaccine.2015.04.036. PubMed

3. Daly M, Jones A, Robinson E. 2021. Public trust and willingness to vaccinate against COVID-19 in the US from October 14, 2020, to March 29, 2021 [Internet]. *JAMA*. 325(23), 2397-99. doi:https://doi.org/10.1001/jama.2021.8246. PubMed

4. Hamel L, Kirzinger A, Lopes L, Kearney A, Sparks G, et al. KFF COVID-19 Vaccine Monitor: January 2021 - Vaccine Hesitancy. KFF [Internet]. 2021. Available at: https://www.kff.org/report-section/kff-covid-19-vaccine-monitor-january-2021-vaccine-hesitancy/

5. Cuello-Garcia C, Pérez-Gaxiola G, van Amelsvoort L. 2020. Social media can have an impact on how we manage and investigate the COVID-19 pandemic. *J Clin Epidemiol*. 127, 198-201. Epub Jun 2020. doi:https://doi.org/10.1016/j.jclinepi.2020.06.028. PubMed

6. Ahmed N, Quinn SC, Hancock GR, Freimuth VS, Jamison A. 2018. Social media use and influenza vaccine uptake among White and African American adults. *Vaccine*. 36(49), 7556-61. Epub Oct 2018. doi:https://doi.org/10.1016/j.vaccine.2018.10.049. PubMed

7. Wang Y, McKee M, Torbica A, Stuckler D. 2019. Systematic Literature Review on the Spread of Health-related Misinformation on Social Media. *Soc Sci Med*. 240, 112552. Epub Sep 2019. doi:https://doi.org/10.1016/j.socscimed.2019.112552. PubMed

8. Siwakoti S, Yadav K, Thange I, Bariletto N, Zanotti L, et al. Localized misinformation in a global pandemic: report on COVID-19 narratives around the world. *Empirical Study of Conflict*, Princeton University [Internet], 2021. pages 1-68. Available at: https://esoc.princeton.edu/publications/localized-misinformation-global-pandemic-report-covid-19-narratives-around-world

9. Loomba S, de Figueiredo A, Piatek SJ, de Graaf K, Larson HJ. 2021. Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA [Internet]. *Nat Hum Behav*. 5(3), 337-48. doi:https://doi.org/10.1038/s41562-021-01056-1. PubMed

10. Centers for Disease Control and Prevention[Internet]. Social Determinants of Health. c2021. Available at: https://www.cdc.gov/socialdeterminants/index.htm

11. Schillinger D. 2021. Social determinants, health literacy, and disparities: intersections and controversies [Internet]. *Health Lit Res Pract*. 5(3), e234-43. doi:https://doi.org/10.3928/24748307-20210712-01. PubMed

12. Twitter.com [Internet] c2021. Available from: https://twitter.com/

13. DeVerna MR, Pierri F, Truong BT, Bollenbacher J, Axelrod D, et al. CoVaxxy: A Collection of English-language Twitter Posts About COVID-19 *Vaccines*. ArXiv210107694 Cs [Internet]. 2021. Available at: http://arxiv.org/abs/2101.07694

14. Federal Communications Commission [Internet]. Area and Census Block. Accessed August 16, 2021. Available at:

https://anypoint.mulesoft.com/exchange/portals/fccdomain/55a35f65-7efa-42af-98af-41ae37d824b2/area-and-census-block/minor/1.0/console/method/%2333/

15. Sayyadiharikandeh M, Varol O, Yang K-C, Flammini A, Menczer F. Detection of novel social bots by ensembles of specialized classifiers. *Proc 29th ACM Int Conf Inf Knowl Manag* [Internet]. 2020. 2725-2732. doi:https://doi.org/10.1145/3340531.3412698

16. Varol O, Ferrara E, Davis C, Menczer F, Flammini A. Online human-bot interactions: detection, estimation, and characterization. In Proceedings of the international AAAI conference on web and social media (Vol. 11, No. 1). ArXiv: 1703.03107 [Internet]. 2017. Available at: https://arxiv.org/abs/1703.03107

17. McCorriston J, Jurgens D, Ruths D. Organizations are users too: characterizing and detecting the presence of organizations on Twitter. *Proc Int AAAI Conf Web Soc Media* [Internet]. 2021. 9(1):650-653. Available at: https://ojs.aaai.org/index.php/ICWSM/article/view/14672

18. Hutto C, Gilbert E. VADER: A parsimonious rule-based model for sentiment analysis of social media text. *Proc Int AAAI Conf Web Soc Media* [Internet]. 2014. 8(1):216-225. Available at: https://ojs.aaai.org/index.php/ICWSM/article/view/14550

19. Ribeiro FN, Araújo M, Gonçalves P, André Gonçalves M, Benevenuto F. 2016. SentiBench - a benchmark comparison of state-of-the-practice sentiment analysis methods [Internet]. EPJ Data Sci. 5(1). doi:https://doi.org/10.1140/epjds/s13688-016-0085-1.

20. Centers for Disease Control and Prevention [Internet] What is health literacy? Take action. Find out. c2021 Available at: https://www.cdc.gov/healthliteracy/learn/index.html

21. Montagni I, Ouazzani-Touhami K, Mebarki A, Texier N, Schück S, et al., & CONFINS group. 2021. Acceptance of a Covid-19 vaccine is associated with ability to detect fake news and health literacy. *J Public Health (Oxf)*. 43(4), 695-702. doi:https://doi.org/10.1093/pubmed/fdab028. PubMed

22. Kricorian K, Civen R, Equils O. 2022. COVID-19 vaccine hesitancy: misinformation and perceptions of vaccine safety. *Hum Vaccin Immunother.* 18(1). doi:https://doi.org/10.1080/21645515.2021.1950504. PubMed

23. Lorini C, Santomauro F, Donzellini M, et al. 2018. Health literacy and vaccination: A systematic review. *Hum Vaccin Immunother*. 14(2), 478-88. doi:https://doi.org/10.1080/21645515.2017.1392423. PubMed

24. National Health Literacy Mapping to Inform Health Care Policy. *Health Literacy Data Map*. University of North Carolina at Chapel Hill [Internet]. 2014. Available at: http://healthliteracymap.unc.edu/#

25. Centers for Disease Control and Prevention/ Agency for Toxic Substances and Disease Registry [Internet]. CDC/ATSDR's Social Vulnerability Index (SVI). 2021. Available at: https://www.atsdr.cdc.gov/placeandhealth/svi/index.html

26. Hughes M, Wang A, Grossman M, Pun E, Whiteman A, et al. 2021. County-Level COVID-19 vaccination coverage and social vulnerability — United states, december 14, 2020–March 1, 2021. US Department of Health and Human Services/Centers for Disease Control and Prevention [Internet]. 2021. 70(12). Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7993557/pdf/mm7012e1.pdf

27. Wilson GM, Ball MJ, Szczesny P, Haymann S, Polyak M, et al. 2021. Health Intelligence Atlas: A Core Tool for Public Health Intelligence. Appl Clin Inform. 12(04), 944-53. Epub Oct 2021. doi:https://doi.org/10.1055/s-0041-1735973. PubMed