Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

___Lanyue Zhang_____ ____04/30/2021_____

Date

Social Media in Public Health Surveillance

By

Lanyue Zhang

Master of Public Health

Hubert Department of Global Health

_____

Scott JN McNabb, PhD, MS

Thesis Advisor

Social Media in Public Health Surveillance

By

Lanyue Zhang
Bachelor of Medical English and Health Communication
Tianjin Medical University
2018

Thesis Advisor: Scott JN McNabb, PhD, MS

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in the Hubert Department of Global Health
2021

Abstract

Social Media in Public Health Surveillance

By Lanyue Zhang

Background

Infectious diseases are significant threats throughout the world. In recent years, social media has shown exciting promise to send early warnings to public health authorities to prevent and control the spread of infectious diseases. Most studies about social media and public health surveillance (PHS) discussed predictive models based on internet search queries or online trends but using social media for PHS is rare.

Objective

The aim of this study was to perform a literature review to identify and evaluate the use of social media for infectious disease PHS. Besides, this study also compared and contrasted PHS of different social media platforms and identified advantages and weakness.

Methods

Systematic search of published literature in PubMed and included the articles from 2012 to Mar 2021. Citations were exported to Covidence™ system to perform the title and abstract screening and full-text review. Fifteen articles were identified in the end.

Results

The fifteen selected articles focused on five social media platforms and included seven infectious diseases. This study found that the posts in social media platforms could be early warning system to support PHS, since the real-time reporting on social media will circulate faster than traditional notification. Besides, based on interaction function, the public health workers could promote health education through social media, and explore emerging disease symptoms and prevention methods.

Conclusions

Social media could be a great supplement for traditional PHS. Developing visual social media platforms also need to be considered. Besides, we also need pay attention to weakness of social media in PHS and improve them in future research.

Limitations

Only 15 articles were identified in this review, and most of them focused on text-based posts in social media for PHS, only one article focused on visual pictures-based posts. Research on social media for PHS is still a new and growing field, we need more research focused on different social media platforms to develop PHS.

Recommendations

Developing visual social media platforms could support PHS for people who have lower education levels to promote health education. Public health authorities could use social media platforms to post more positive and active information about vaccination to influence public opinion.

Key words: Health Surveillance; Public Health Surveillance; Social Media; Internet; Infectious Disease; Disease Prevention; Public Health Emergency; Sharing Information.

Social Media in Public Health Surveillance

By

Lanyue Zhang
Bachelor of Medical English and Health Communication
Tianjin Medical University
2018

Thesis Advisor: Scott JN McNabb, PhD, MS

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in the Hubert Department of Global Health
2021

Acknowledgments

My sincerest thanks to my supervisor, Dr. Scott JN McNabb, for his encouragement, supports and guidance throughout the thesis process as well as my master program. I would like to give my deepest gratitude to his thoughtful suggestions and professional help. I feel so lucky that I could have such a great professor, who teaches me and help me rebuild confidence to face the challenges, he is the professor that has influenced me most in Rollins.

Thanks also to my APE advisor, Dr. James V. Lavery for providing me the opportunity to complete my summer practicum and gave me a new ethical perspective on public health crisis. Thanks to my ADAP, Theresa Nash, for her great patience and kindness, she helped me a lot to adapt new environment when I was a first-year student and supports me through these two years. Thanks to research librarian Shenita Peterson for her help with developing systematic literature review and search terms. Thanks also to Dr. Carol Haley, Dr. Xiaochun Wang and Dr. Brian Dixon for their patience and expert suggestion for our cooperation.

I would also like to thank my family, especially my parents for their love and support. Thanks for my mom and dad who always encouraged me during the two years, and thanks for my cousin Haijing and her husband Wenchao for their care and affection.

Finally, I would like to thank my wonderful friends, Xiaoyue and her brother, Meng, Feifei, Jiahui, my Co-TA Andrea, Edward and his wife, Hong's family, I consider myself lucky for our paths to have intersected in Atlanta.

Contents

Chapter 1. Introduction

Based on the World Health Organization (WHO) definition, public health surveillance (PHS) is the "continuous, systematic collection, analysis, and interpretation of health-related data needed for the planning, implementation, and evaluation of public health practice".[1] As the best weapon to avert epidemics. [2] PHS has improved in the past decades and increasingly required real-time reporting to alert relevant public health and healthcare systems. [3]

In recent years, social media has also shown exciting promise to send early warnings to public health authorities to prevent and control the spread of infectious diseases. Especially during the COVID-19 pandemic, the question has arisen: how should we use social media for more effective and efficient PHS? This study reviewed PHS systems that used social media. What approaches were effective to fill gaps and what future actions should be taken to address needs that remaining unmet?

1.1 Definition of Terms

The key terms used in this systematic literature review are below.

PHS is the ongoing systematic collection, analysis, and interpretation of health data from defined populations for use in planning, implementing, and evaluating public health programs. [4]

Social media are "web-based services that allow individuals, communities, and organizations to collaborate, connect, interact, and build community by enabling them to create, co-create, modifies, share, and engage with user-generated content that is easily accessible". [5] Social media is the platform based on communication and internet technologies that allow people to interact and communicate with each other, share personal experiences and locations, and create or exchange information in networks.

Infectious disease is defined as a communicable disease that can spread, directly or indirectly, from one person to another, and "can be grouped in three categories: diseases which cause high levels of mortality; diseases which place on populations heavy burdens of disability; and diseases which owing to the rapid and unexpected nature of their spread can have serious global repercussions". [6]

Emergency preparedness and response is a critical part of the disaster management cycle. It includes any systematic response to an unexpected or dangerous occurrence. [7] The goal of emergency response is to mitigate the impact of the event on a target population.

1.2 Background and Significance

Infectious diseases are significant threats throughout the world. During the COVID-19 pandemic, we realized the importance of improving public health and establishing early warning systems. Social media played a significant role in informing and forming the hearts and minds of individuals and society. [8] At the beginning of the social media evolution, it

was a place for communication among users to create, generate, and exchange content and information, plus initiate voluntary participation and communication messages. [9] Now, social media in public health is recognized and valued. People share their personal healthcare experiences, exchange health information, and express their emotions on various social media platforms.

Using social media for PHS could be a mechanism to add value, since real-time reporting and public participation are assets. Social media could provide the early warning signals to public health authorities. [3]

1.3 Purpose Statement

Recently, studies have discussed using social media data for PHS; most discussed predictive models based on internet search queries or online trends (e.g., Google Flu Trends and Google Dengue Trends). [10, 11] However, using social media for PHS is rare. The main goal of this study was to perform a literature review to identify and evaluate the use of social media for PHS. The specific objectives were to …

1. evaluate PHS in different social media platforms.
2. compare and contrast PHS of different social media platforms.
3. identify advantages and weakness of using social media for PHS.
4. recommend how to approach real-time reporting for infectious disease outbreak.

Chapter 2. Methods

This study was based on analysis of secondary data. The systematic literature review consisted of publications, articles, and papers accessible on the internet. We used a keyword search of PubMed for ("Health Surveillance") AND ("Social Media" OR "Internet" OR "Mass Media" OR "Multimedia" OR "Social Web"). The search was limited to studies conducted between 2012 and Mar 2021.

After the search was conducted, citations were exported to Covidence™ system to perform the title and abstract screening and full-text review. To be included in the systematic review, the literature had to be …

1. written in English.
2. available through the Emory library, or found in other catalog systems.
3. discussed a specific social media platform.
4. discussed infectious disease surveillance.

Exclusion criteria included …

1. not written in English, or only the tile and abstract in English.
2. published after 2012 and before Mar 2021.

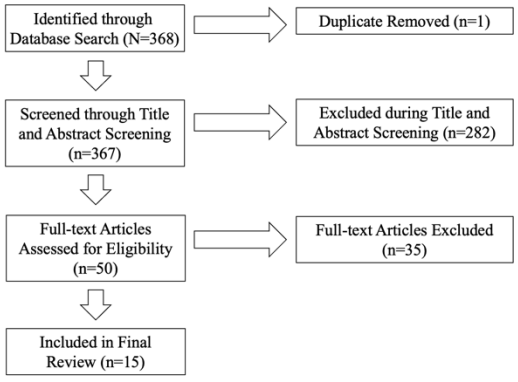
3. not found the full text in the Emory library or cannot be found in any other catalog system.

4. did not evaluate or discuss infectious disease or did not discuss a specific social media platform.

5. article is the systematic literature review.

After the title and abstract screening, full text was obtained for all articles that met the inclusion criteria (Fig 1). All articles were carefully read to ensure their adherence to the specified inclusion criteria. After carefully review, a list of core themes within the literature were identified: social media; health surveillance; and infectious disease. Finally, there were 368 articles retrieved from the databases. After title and abstract screening, 282 irrelevant articles were excluded. A total of 35 articles were excluded after the full text review.

Fig 1. Social Media and Public Health Surveillance Articles Reviewed and Chosen though Covidence™, 2021



Fifteen articles were identified in the published literature.

Chapter 3. Results

Of the 15 articles reviewed, eight focused on Twitter™, three on Weibo®, two on WeChat®, one on Baidu Tieba®, and one on Instagram™ (Table 1).

Table 1. Manuscripts Extracted during Systematic Literature Review of Social Media and Public Health Surveillance, 2021

Author(s)	Title	Year	Pandemic	Social Media Platform
Dunn, *et* al.	Associations Between Exposure to and Expression of Negative Opinions About Human Papillomavirus Vaccines on Social Media: An Observational Study.	2015	HPV Vaccines	Twitter™
Zhou, *et* al.	Using social connection information to improve opinion mining: Identifying negative sentiment about HPV vaccines on Twitter.	2015	HPV Vaccines	Twitter™
Guo, *et* al.	Mining twitter to explore the emergence of COVID-19 symptoms.	2020	COVID-19	Twitter™
Mackey, *et* al.	Machine Learning to Detect Self-Reporting of Symptoms, Testing Access, and Recovery Associated With COVID-19 on Twitter: Retrospective Big Data Infoveillance Study.	2020	COVID-19	Twitter™
Allen, *et* al.	Applying GIS and Machine Learning Methods to Twitter Data for Multiscale Surveillance of Influenza.	2016	Influenza	Twitter™
Wang, *et* al.	Regional Level Influenza Study with Geo-Tagged Twitter Data.	2016	Influenza	Twitter™
Odlum & Yoon	What can we learn about the Ebola outbreak from tweets?	2015	Ebola	Twitter™
Nsoesie, *et* al.	Social Media as a Sentinel for Disease Surveillance: What Does Sociodemographic Status Have to Do with It?	2016	Dengue	Twitter™
Fung & Wong	Efficient use of social media during the avian influenza A(H7N9) emergency response.	2013	Avian Influenza	Weibo®
Zhang, *et* al.	Leveraging social networking sites for disease surveillance and public sensing: the case of the 2013 avian influenza A(H7N9) outbreak in China.	2015	Avian Influenza	Weibo®
Shen, *et* al.	Using Reports of Symptoms and Diagnoses on Social Media to Predict COVID-19 Case Counts in Mainland China: Observational Infoveillance Study.	2020	COVID-19	Weibo®
Wang, *et* al.	A New System for Surveillance and Digital Contact Tracing for COVID-19: Spatiotemporal Reporting Over Network and GPS.	2020	COVID-19	WeChat®
Wang, *et* al.	Using WeChat, a Chinese Social Media App, for Early Detection of the COVID-19 Outbreak in December 2019: Retrospective Study.	2020	COVID-19	WeChat®
Dong, *et* al.	HIV-related posts from a Chinese internet discussion forum: An exploratory study.	2019	HIV	Baidu Tieba®
Nobles, *et* al.	#HIV: Alignment of HIV-Related Visual Content on Instagram with Public Health Priorities in the US.	2020	HIV	Instagram™

3.1 Twitter™

The eight articles focused on Twitter™ included two HPV vaccines, two COVID-19 symptoms, two Influenza, one Ebola, and one Dengue outbreak.

HPV Vaccines

The two articles on HPV vaccine discussed negative opinions about HPV vaccines on Twitter™. [12, 13] In these studies, authors collected tweets that included terms related to HPV vaccine and the social connections among users posting or reposting the tweets. Due to the huge number of collected tweets, the author used a supervised machine learning approach to classify the tweets, and then trained algorithms and recognized similar patterns in the remaining tweets. [12] These data show that more users posted negative tweets after exposure to negative tweets (37.78% [2780/7361]. This compared to 10.92% [1234/11,296]) of users who were exposed to positive or neutral tweets: (relative risk= 3.46 (95% CI=3.25-3.67, P<.001). Users more often exposed to negative opinions were more likely to subsequently post negative opinions. [12]

Zhou, *et* al. used machine learning methods to train classifiers using the collected tweet data. They divided the six months of data into two distinct but contiguous 3-month periods and randomly sampled tweets for use in the classifier training and testing. They demonstrated the difference in temporal degradation across classifiers constructed from content and social

connection features. Their results showed that public health workers could use information about the Twitter™ users that people follow online to predict their opinions, since that the information about who users follow, rather than who follows them, could be more useful for predicting the direction of their expressed ideas. [13]

COVID-19 Symptoms

The COVID-19 pandemic has emphasized the importance of identifying and detecting infectious diseases. Timely identification of COVID-19 symptoms and effective testing could reduce the spread and improve disease prevention. The two articles based on Twitter™ explored the emergence of COVID-19 symptoms and people's testing needs. [14, 15]

The study of Guo, *et* al. (2020) was based on unique tweets posted between March 30, 2020 and April 19, 2020. It explored > 30 COVID-19 symptoms. The study collected tweets related to COVID-19 and used search terms for COVID-19 synonyms and COVID-19 symptoms (fever, dry cough, and shortness of breath) suggested by CDC in Mar 2020. [14, 16] After extracted physical symptom-related terms, the author used R software to calculate the frequency of each symptom mentioned within the tweets and used Microsoft Excel™ to generate graphs illustrating symptom trends mentioned over time. A total of 36 symptoms were extracted. The three most frequently mentioned symptoms were cough, fever, and difficulty breathing. The other mentioned symptoms (i.e., chills, muscle pain, headache, sore

throat, loos of taste and loss of smell) were added in COVID-19- related symptoms suggestion by CDC in late April. [14, 16]

The study of Mackey, *et* al. (2020) used unsupervised machine learning to characterize self-reporting symptoms, experiences with testing, and mentions of recovery related to COVID-19 based on tweets in March 2020. They collected tweets that contained terms which related with COVID-19 symptoms and found that the co-occurrence of tweets for these themes was significant for users reporting symptoms with a lack of testing and with a discussion of recovery. The author showed that the value of this study is in its innovative method using data mining in combination with modeling to screen a large volume of unstructured data to detect and characterize potential underreported cases of COVID-19. [15]

Influenza

Two articles explored using Twitter™ for Influenza PHS. Wang, *et* al. developed a prototype framework to automatically collect, analyze, and model geo-tagged flu tweets from real-time Twitter ™ data streams. They extracted flu tweets from real-time data streams and tagged each with geographic locations. They proposed a partial differential equation based on a mathematical model to predict flu tweets counts in a specific area during a certain week. This was the first paper that studied the correlation at the regional level; this PHS could be an early warning before CDC released official statistics. [17] Allen, *et* al. (2016) developed search tools that queried the Twitter™ search API for Twitter™

users who posted tweets based on spatial constraints and keyword filters. A web map was created to display the locations related to each keyword. [18]

Ebola

The article *What can we learn about the Ebola outbreak from Tweets* [19] showed the importance of Twitter™ as an early warning system to support PHS during the Ebola outbreak in 2013. The author used tweets Corpus to collect the tweets' content, time stamps, and geographical locations; they investigated geographic spread of Ebola information and discussed public perceptions of Ebola for content analysis. The study found an increase in the frequency of Ebola-related tweets in the days leading up to the official news alert even in areas that adoption and use of Twitter™ were limited by resources. [19]

Dengue

Nsoesie, *et* al. showed the limitation of social media as a sentinel for PHS from a sociodemographic perspective. [20] The study used machine learning methods to explore spatio-temporal dengue event reporting trends on Twitter™ related with confirmed cases, and quantified associations with sociodemographic factors in three states in Brazil. The results showed that poorer-educated people, males, and people > 60 years of age are less likely to tweet about dengue. Therefore, they proposed social media might not be an adequate supplement to traditional PHS. [20] They suggested that public health

professionals should consider population, representation by location, age, education level and racial/ethnic backgrounds when the research use social media data.

3.2 Weibo®

The three articles focused on Weibo® included two avian influenza (H7N9) and one COVID-19.

Avian Influenza

Both articles discussed social media for PHS in the avian influenza (H7N9) outbreak and emphasized the importance of social media accounts of health authorities. [21, 22] Zhang, *et* al. established two data sets: one included a line list of confirmed cases from conventional public health information channels. The other included Weibo® posts containing the term "N7N9". They used SPSS for quantitative analysis comparing timeliness of reporting of new cases by different information channels. They performed qualitative analysis of Weibo® users with the timeliest posts on new cases. The study showed that the volume of social media discussions corresponded to the number of reported cases. When new cases appeared, the number of Weibo® posts increased at the same time. This phenomenon might be considered as the public's awareness of disease outbreaks increased. They mentioned province/municipality-based news agencies were the source for Weibo® users on new H7N9 cases. [21] These news agencies posted news on their Weibo® account immediately when they received new information from provincial/municipal health authorities. By

comparison, the official website of National Health Commission lagged behind because the additional time taken to collate information from all provincial/municipal health authorities around the country for the Chinese central government's daily update. [22]

COVID-19

Shen, *et* al. developed a supervised machine learning classifier to predict COVID-19 case counts by using reports of symptoms and diagnosis on social media. The author announced that this infoveillance study employs the largest, most comprehensive and fine-grained social media data to predict COVID-19 cases in mainland China. [23] This research identified "sick posts" in which users reported their own or other people's symptoms and diagnoses related to COVID-19 on Weibo®-based user pool.

The user pool was built by the author and included 250 million people. The author retrieved around 15 million posts from Nov 1, 2019 – Mar 31, 2020 and claimed this study showed that these sick posts significantly predicted daily cases up to 14 days before official statistics. It also highlighted the importance of using rigorous procedures and understanding information sharing behavior to obtain quality disease signals. [23] Based on a subset of geotagged posts for Hubei Province, the author mentioned that Hubei province was the epicenter of the outbreak. Lots of residents used social media like Weibo® to seek help for testing or medical care because of the extreme shortage of both testing and medical resources during the early stage of the outbreak. However, social media help-seeking

activities were uncommon in other areas in China, since the testing and health resources were much more adequate in other cities or provinces. The authors showed the predictive power of social media data; that it may vary across geographic areas, with different medical resources preparing levels, and at different stages of the outbreak. The authors emphasized that social media surveillance will improve the allocation of medical resources and could be performed online during the pandemic.

3.3 WeChat®

Two articles focused on the COVID-19 pandemic and used WeChat® for PHS. [24, 25] The study by Wang, *et* al. developed a mini-program within WeChat® app to collect data. This mini-program analyzed data from all WeChat® users and traced close contacts of patients based on GPS and geospatial artificial intelligence (AI) technologies. The author mentioned this WeChat® mini-program could also capture volunteer-interacting data from other mobile apps (e.g., location data from Uber history or payment history). Therefore, the spatiotemporal data in the cloud could support health professionals perform PHS for both diagnosed patients and healthy individuals [24]. Wang, *et* al. focused on WeChat® to do a retrospective study for early detection of COVID-19 outbreaks in the early stage in China. The WeChat® Index based on data that shows the frequency of specific words in user posts, subscriptions, and searches. The author proposed that increasing WeChat® Index results for a specific keyword before the day of the official outbreak announcement could be

considered a potential candidate outbreak sign. The research verified this hypothesis; data from WeChat® could have enabled the detection of the COVID-19 outbreak in 2019 about 14 days earlier than the official announcement in China. [25] The author also claimed that keywords about symptoms (e.g., “shortness of breath”, “dyspnea” and “diarrhea”) were not meaningful for early detection compared with the specific keyword “Feidian (SARS)”. SARS impressed Chinese people; it had potential meaning like “horrible infectious disease” and “need to pay attention”. This keyword could be used with more effect than others.

3.4 Baidu Tieba®

Dong, *et* al. used word clouds and coding schemes with negative binomial regression to compare HIV- and tuberculosis-related posts in Baidu Tieba®. They evaluated the needs of people living with HIV/ADIS and their online habits and verified geolocations from posts of Baidu Tieba® reported cases of HIV infection. [26] Authors collected data from the HIV forum and found more than one-third of HIV-related posts were seeking advice, including over 40% of posts were related to medical topics. However, these kinds of posts received fewest replies. Besides, the author claimed that the number of HIV-related social support requests was almost more than triple as many as the number of posts providing social support, while the relatively similar proportions of support requests and support provision were observed in the tuberculosis forum. [26] In addition, the research showed the HIV-related forum users usually sought advice and social support at 10-11 AM and 9-10 PM. The

author suggested that social support providers need to be available during those time periods.

3.5 Instagram™

Nobles, *et* al. used Instagram™ to do PHS for HIV prevalence. The author mentioned that the majority of at-scale analyses of HIV-related posts on social media are text-based posts. So this research was based on users' pictures or photos on Instagram™, the famous visual social media platform. [27] The authors collected posts with the hashtag "HIV" and captured posters' self-labeled posts as related to HIV. Then they analyzed the visual content of the images in the posts. Firstly, the automated image recognition software was used to assign textual tags to content in the images. Secondly, the authors identified 30 topics present in the images, and the top 10 textual tags associated with each topic and five random images associated with each topic were examined to manually assign a label to each topic. Finally, the authors resulted a total of 10 themes present in the images. In the end, they showed that except the most common hashtag "#HIV" and its co-occurring hashtags (such as "#AIDS" or "#hivpositive"), the second most frequent cluster of hashtags was for sexual and gender minorities, such like "#gay" or "#LGBT". However, they also mentioned that the specific HIV prevention strategies or methods were rarely mentioned in posts of Instagram™.

Chapter 4. Discussion

4.1 PHS of Social Media

The COVID-19 pandemic was discussed in five articles in three social media platforms, followed by HIV in two social media platforms.

COVID-19 in Twitter™, WeChat®, and Weibo®

To do PHS for testing or emergency symptoms, two articles focused on COVID-19 symptoms based on Twitter™. [14, 15] However, on Chinese social media platforms, Weibo® and WeChat®, the articles mentioned that the majority of COVID-19 posts reflected public awareness of COVID-19 instead of specific symptoms. [23] Even the term "Feidian (SARS)" in collected posts was more meaningful than other symptoms like "shortness of breath" in the early stage of the disease outbreak. [25] This difference may show the users' different culture backgrounds in different social media. The outbreak of SARS in 2003 impressed Chinese people; it has the potential meaning like "horrible infectious disease" and "need to pay attention". This keyword could be used more prevalent than other symptom words in Chinese social media platforms.

HIV in Baidu Tieba® and Instagram™

Both articles about HIV surveillance on social media talked about the "stigma" and "discrimination" of HIV patients. [26, 27] The authors mentioned that HIV patients or their

family members would like to post their diagnosis, experiences, and requirements in anonymous social media platforms. Besides, all the authors mentioned that the number of posts about seeking help or social support were more than posts about specific prevention strategies or support provision in Baidu Tieba® or Instagram™. However, the biggest difference between these two PHS methods is the research of Dong, *et* al. focused on text-based posts in Baidu Tieba®, while the research of Nobles, *et* al. paid attention to visual content in Instagram™. The study of Nobles, *et* al. helps us have a better understanding of PHS in social media platforms based on both text and visual content.

4.2 Advantages and Weakness of Social Media for PHS

Advantages

Real-time reporting

Almost all articles mentioned that real-time reporting being the advantage of using social media for PHS. Posts on social media will circulate faster than traditional notification. Based on social media, public health workers could identify early warning through collecting and analyzing disease-related posts and explore emerging disease symptoms and prevention methods. Social media could be a great supplement for traditional PHS.

Interaction among users

All social media platforms have communicating functions such as “Like”, “Thumbs-up”, “Forward”, “Comments”, “Follow” or “Subscribe”. These functions allow users to communicate with each other; public health professionals could post their suggestions.

For example, during the avian influenza (H7N9) outbreak Chinese health experts and authorities held many real-time, questions-and-answer sessions on Weibo®. [22] Public health workers could also collect opinions about vaccination and disease control or prevention strategies. For instance, users who were more often exposed to negative opinions were more likely to subsequently post negative opinions. [12] These trends should lead public health officials to promote health education and post positive posts about vaccine on social media.

Weaknesses

Rumors and unrelated information

It is difficult to remove all background noise of collected data from social media posts. Rumors and unrelated information should be identified; this is labor-intensive work.

Personal information security

Public health ethics identifies the potential conflict between the needs of public health and personal privacy rights. Personal location, context of posts and other information will be available when public health professionals collect data. Additional crowdsourced epidemiologic information on infected cases may also be explored through user posts on

social media. For example, updates on patient health conditions, exposure history and family member contacts challenge personal rights. [22]

Distraction of public concern

Public concern may be distracted by other social crises. If so, the concern or search trend of disease-related key words will decrease and disturb PHS by researchers. For instance, the author mentioned that the Weibo® users' interest in avian influenza plummeted briefly on 21 and 24 Apr 2013 because the public's concern shifted to the Sichuan earthquake in China.

4.3 Limitations

The first limitation is that only 15 articles were identified meeting the inclusion criteria of this review. This was because research on social media for PHS is still a new and growing field. This systematic literature review only focused on English articles; this study was conducted by one individual relying on the Emory library system. There were some articles that may have been relevant but were unable to be retrieved because of limited resources. Additionally, most research focused on text-based posts in social media platforms for PHS; only one article focused on visual pictures-based posts. Based on Nsoesie, *et* al.'s study, the lower education population and people older than 60 are less likely to post tweets on Twitter™. [20] This implied that some portions of the population, such as individuals with little or no basic education may lack access and the knowledge to read the posting content in

text-based tweets or other posts. Therefore, more picture-based or video-based posts need to be evaluated. As the most famous short-form, mobile video application, TikTok™, also hosted popular videos during the COVID-19 pandemic, including users' Vlog about being infected, and users' experience about quarantine and testing. However, only a few articles discussed the effect of TikTok™ on PHS during the outbreak. We need more research focused on different social media platforms to develop PHS.

4.4 Recommendations

After examining articles in this systematic literature review, several recommendations for future PHS can be made. First, as a supplement to traditional PHS, social media PHS should be applied. Developing visual social media platforms could support PHS for people who have lower education levels to promote health education.

Specific foci within PHS-like vaccination coverage could help vaccine preventable disease activities among communities. Public health departments could take action to help and protect people who live in communities or areas with lower vaccination coverage. [28] User opinions about vaccination on social media platforms should be monitored and guided by public health professionals. The two articles about HPV vaccine opinions on Twitter™ showed users who were more likely to subsequently post negative opinions about vaccination when they often exposed by negative opinions. [12] Information about who users follow could be more useful for predicting the direction of users' expressed ideas. [13]

Public health authorities could use social media platforms to post more positive and active information about vaccination to influence public opinion.

References:

1. WHO, *Public Health Surveillance.* 2021.
2. Choi, B.C.K., *The Past, Present, and Future of Public Health Surveillance.* Scientifica, 2012. **2012**: p. 875253.
3. Al-Surimi, K., et al., *The Potential of Social Media and Internet-Based Data in Preventing and Fighting Infectious Diseases: From Internet to Twitter.* Adv Exp Med Biol, 2017. **972**: p. 131-139.
4. Thacker, S.B. and R.L. Berkelman, *Public health surveillance in the United States.* Epidemiol Rev, 1988. **10**: p. 164-90.
5. McCay-Peet, L. and A. Quan-Haase, *A Model of Social Media Engagement: User Profiles, Gratifications, and Experiences*. 2016. p. 199-217.
6. WHO, *Infectious diseases.* 2021.
7. WHO, *Emergency Response.* 2002.
8. Zolkepli, I.A. and Y. Kamarulzaman, *Social media adoption: The role of media needs and innovation characteristics.* Computers in Human Behavior, 2015. **43**: p. 189-209.
9. Kaplan, A.M. and M. Haenlein, *Users of the world, unite! The challenges and opportunities of Social Media.* Business Horizons, 2010. **53**(1): p. 59-68.
10. Xu, Q., et al., *Forecasting influenza in Hong Kong with Google search queries and statistical model fusion.* PLOS ONE, 2017. **12**(5): p. e0176690.
11. Strauss, R.A., et al., *Google dengue trends: An indicator of epidemic behavior. The Venezuelan Case.* International Journal of Medical Informatics, 2017. **104**: p. 26-30.
12. Dunn, A.G., et al., *Associations Between Exposure to and Expression of Negative Opinions About Human Papillomavirus Vaccines on Social Media: An Observational Study.* J Med Internet Res, 2015. **17**(6): p. e144.
13. Zhou, X., et al., *Using social connection information to improve opinion mining: Identifying negative sentiment about HPV vaccines on Twitter.* Stud Health Technol Inform, 2015. **216**: p. 761-5.
14. Guo, J.W., et al., *Mining twitter to explore the emergence of COVID-19 symptoms.* Public Health Nurs, 2020. **37**(6): p. 934-940.
15. Mackey, T., et al., *Machine Learning to Detect Self-Reporting of Symptoms, Testing Access, and Recovery Associated With COVID-19 on Twitter: Retrospective Big Data Infoveillance Study.* JMIR Public Health Surveill, 2020. **6**(2): p. e19509.
16. CDC, *Symptoms of COVID-19.* 2020.
17. Wang, F., et al., *Regional Level Influenza Study with Geo-Tagged Twitter Data.* J Med Syst, 2016. **40**(8): p. 189.

18. Allen, C., et al., *Applying GIS and Machine Learning Methods to Twitter Data for Multiscale Surveillance of Influenza.* PLoS One, 2016. **11**(7): p. e0157734.
19. Odlum, M. and S. Yoon, *What can we learn about the Ebola outbreak from tweets?* Am J Infect Control, 2015. **43**(6): p. 563-71.
20. Nsoesie, E.O., et al., *Social Media as a Sentinel for Disease Surveillance: What Does Sociodemographic Status Have to Do with It?* PLoS Curr, 2016. **8**.
21. Fung, I.C. and K. Wong, *Efficient use of social media during the avian influenza A(H7N9) emergency response.* Western Pac Surveill Response J, 2013. **4**(4): p. 1-3.
22. Zhang, E.X., et al., *Leveraging social networking sites for disease surveillance and public sensing: the case of the 2013 avian influenza A(H7N9) outbreak in China.* Western Pac Surveill Response J, 2015. **6**(2): p. 66-72.
23. Shen, C., et al., *Using Reports of Symptoms and Diagnoses on Social Media to Predict COVID-19 Case Counts in Mainland China: Observational Infoveillance Study.* J Med Internet Res, 2020. **22**(5): p. e19421.
24. Wang, S., S. Ding, and L. Xiong, *A New System for Surveillance and Digital Contact Tracing for COVID-19: Spatiotemporal Reporting Over Network and GPS.* JMIR Mhealth Uhealth, 2020. **8**(6): p. e19457.
25. Wang, W., et al., *Using WeChat, a Chinese Social Media App, for Early Detection of the COVID-19 Outbreak in December 2019: Retrospective Study.* JMIR Mhealth Uhealth, 2020. **8**(10): p. e19589.
26. Dong, Y., et al., *HIV-related posts from a Chinese internet discussion forum: An exploratory study.* PLoS One, 2019. **14**(2): p. e0213066.
27. Nobles, A.L., et al., *#HIV: Alignment of HIV-Related Visual Content on Instagram with Public Health Priorities in the US.* AIDS Behav, 2020. **24**(7): p. 2045-2053.
28. CDC, *What is Vaccination Coverage and Why is it Important?* 2016.