

Assessing COVID-19 vaccine hesitancy, confidence and public engagement: a global social listening study

Zhiyuan Hou, Yixin Tong, Fanxing Du, Linyao Lu, Sihong Zhao, Kexin Yu, Simon J. Piatek, Heidi Larson, Leesa Lin

Submitted to: Journal of Medical Internet Research
on: February 01, 2021

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript..... 5

Supplementary Files..... 17

 Figures 18

 Figure 1..... 19

 Figure 2..... 20

 Figure 3..... 21

 Multimedia Appendixes 22

 Multimedia Appendix 1..... 23

Assessing COVID-19 vaccine hesitancy, confidence and public engagement: a global social listening study

Zhiyuan Hou^{1*} PhD; Yixin Tong^{1*}; Fanxing Du¹ MPH; Linyao Lu¹; Sihong Zhao¹; Kexin Yu¹; Simon J. Piatek²; Heidi Larson² PhD; Leesa Lin² PhD

¹School of Public Health, Global Health Institute, NHC Key Laboratory of Health Technology Assessment Fudan University Shanghai CN

²Department of Infectious Disease Epidemiology London School of Hygiene & Tropical Medicine London GB

*these authors contributed equally

Corresponding Author:

Zhiyuan Hou PhD

School of Public Health, Global Health Institute, NHC Key Laboratory of Health Technology Assessment

Fudan University

130 Dong'an Road

Shanghai

CN

Abstract

Background: Monitoring public confidence and hesitancy is crucial for COVID-19 vaccine rollout. Social media listening can not only monitor public attitudes on COVID-19 vaccine in real-time, but also assess the dissemination and public engagement of these opinions.

Objective: To assess global hesitancy, confidence and public engagement towards COVID-19 vaccination.

Methods: We collected and manually coded 12886 social media posts mentioning COVID-19 vaccine from five global metropolises with high COVID-19 burden between June and July, 2020. After assessment, 7032 posts were included in analysis. We manually double-coded these posts using a coding framework developed according to the WHO's Confidence, Complacency, and Convenience model of vaccine hesitancy, and conducted engagement analysis to investigate public communication about COVID-19 vaccine on social media.

Results: Among social media users, 36.4% (571/1568) in New York, 51.3% (738/1440) in London, 67.3% (144/214) in Sao Paulo, 69.8% (726/1040) in Mumbai, and 76.8% (2128/2770) in Beijing indicated that they intended to accept COVID-19 vaccination. With high perceived risk of getting COVID-19, more tweeters in New York and London expressed lack of confidence in vaccine safety, distrust in governments and experts, and widespread misinformation and rumors. Tweeters from Mumbai, Sao Paulo and Beijing worried more about vaccine production and supply, whereas tweeters from New York and London had more concern on vaccine distribution and inequity. Negative tweets expressing lack of vaccine confidence and misinformation or rumors, had more followers and attracted more public engagement online.

Conclusions: COVID-19 vaccine hesitancy is prevalent across the world, and negative tweets attract higher engagement on social media. It is urgent to develop an effective vaccine campaign that boosts public confidence and addresses hesitancy for COVID-19 vaccine rollout.

(JMIR Preprints 01/02/2021:27632)

DOI: <https://doi.org/10.2196/preprints.27632>

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 large, light gray watermark is oriented diagonally across the center of the page. It consists of the word 'Preprint' in a large, sans-serif font, followed by a circular logo containing a network diagram of three nodes connected by lines. To the right of the logo, the words 'JMIR Publications' are written in a smaller, sans-serif font.](http</p></div><div data-bbox=)

Original Manuscript

Original paper

Zhiyuan Hou, PhD ^{#*1}, Yixin Tong, undergraduate ^{#1}, Fanxing Du, MPH ¹, Linyao Lu, undergraduate ¹, Sihong Zhao, undergraduate ¹, Kexin Yu, undergraduate ¹, Simon J. Piatek, PhD ², Heidi Larson, PhD ², Leesa Lin, PhD ²

Affiliations:

1. School of Public Health, Global Health Institute, NHC Key Laboratory of Health Technology Assessment, Fudan University, Shanghai, China
2. Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, UK

#co-first author

***Corresponding author:** Zhiyuan Hou. School of Public Health, Global Health Institute, NHC Key Laboratory of Health Technology Assessment, Fudan University, Shanghai, China.

Address: 130 Dong'an Road, Shanghai 200032, China.

Email: zyhou@fudan.edu.cn

Assessing COVID-19 vaccine hesitancy, confidence and public engagement: a global social listening study**Abstract**

Background: Monitoring public confidence and hesitancy is crucial for COVID-19 vaccine rollout. Social media listening can not only monitor public attitudes on COVID-19 vaccine timely, but also assess the dissemination of and public engagement with these opinions.

Objectives: To assess global hesitancy, confidence and public engagement towards COVID-19 vaccination.

Methods: We collected posts mentioning the COVID-19 vaccine between June and July, 2020 on Twitter from New York (United States), London (United Kingdom), Mumbai (India), and Sao Paulo (Brazil), and Sina Weibo posts from Beijing (China). In total, we manually coded 12886 posts from the five global metropolises with high COVID-19 burdens, and after assessment, 7032 posts were included in analysis. We manually double-coded these posts using a coding framework developed according to the WHO's Confidence, Complacency, and Convenience model of vaccine hesitancy, and conducted engagement analysis to investigate public communication about COVID-19 vaccine on social media.

Results: Among social media users, 36.4% (571/1568) in New York, 51.3% (738/1440) in London, 67.3% (144/214) in Sao Paulo, 69.8% (726/1040) in Mumbai, and 76.8% (2128/2770) in Beijing indicated that they intended to accept COVID-19 vaccination. With high perceived risk of getting COVID-19, more tweeters in New York and London expressed lack of confidence in vaccine safety, distrust in governments and experts, and widespread misinformation and rumors. Tweeters from Mumbai, Sao Paulo and Beijing worried more about vaccine production and supply, whereas tweeters from New York and London had more concerns on vaccine distribution and inequity. Negative tweets expressing lack of vaccine confidence and misinformation or rumors, had more followers and attracted more public engagement online.

Conclusions: COVID-19 vaccine hesitancy is prevalent across the world, and negative tweets attract higher engagement on social media. It is urgent to develop an effective vaccine campaign that boosts public confidence and addresses hesitancy for COVID-19 vaccine rollout.

Key words: COVID-19 vaccine, hesitancy, confidence, acceptance, engagement, social media.

Introduction

As of January 2021, the coronavirus disease 2019 (COVID-19) pandemic has led to more than 100 million cases and 2 million deaths worldwide [1]. Although personal prevention measures such as mask wearing and social distancing have been shown to be effective in curbing the spread of COVID-19 [2], vaccination is expected to be the key to the long-term prevention and control of the pandemic [3,4]. The COVID-19 pandemic has triggered intense global research and development of vaccines against the disease. Several candidate vaccines advanced to Phase III clinical trials in mid-2020, including Oxford/AstraZeneca, Sinopharm, Sinovac, BioNTech/Pfizer, and Moderna vaccines, and had finished clinical trials at the end of 2020. These vaccines have been approved for use by December 2020 in some countries such as the United Kingdom, United States, China [5]. To ensure universal vaccination coverage, governments must enhance public confidence, address the issue of vaccine hesitancy, and design community engagement strategies for COVID-19 vaccine rollout.

Although immunization has proved successful in reducing the global burden of illness and death, a range of concerns have converged to affect public confidence in vaccines. When vaccine confidence breaks down, hesitancy can cause serious consequences such as delays, refusals, and disruptions to research and delivery programmes, and sometimes leads to the resurgence of disease outbreaks [6]. Vaccine hesitancy has proliferated over the decades, and was cited by the World Health Organization (WHO) as one of the top ten global health threats in 2019 [7,8]. Vaccine hesitancy is complex, multi-factorial, and influenced by a combination of emotional, cultural, social, spiritual and political factors. It can vary across countries, vaccines, and time. Reports have indicated that hesitancy towards general vaccines is prevalent among caregivers of children across the world, at rates such as 45.8% in France (2016) [7], 31.8% in the US (2014) [9], 24.6% in Italy (2017) [10], and 23.0% in Brazil (2016) [11]. The prevalence of vaccine hesitancy among healthcare providers ranges from 2% to 16% across different countries [12]. The accelerated research and development process of the COVID-19 vaccine may further exacerbate public concern on its safety and effectiveness [13]. Similarly, the novelty of the disease, politicization of the vaccine and distrust in experts and governments have increased uncertainty about COVID-19 vaccination [3]. Therefore, it is necessary to assess public confidence and acceptance towards the COVID-19 vaccine to prepare for vaccine introduction.

Social media has become a source of data for detecting outbreaks and understanding public attitudes and behaviors during public health emergencies [14-16]. Large amounts of real time data posted on social media platforms can be used to quickly identify public attitudes on COVID-19 vaccines as a way to support health communication and health promotion messaging. A growing body of literature has used social media platforms such as Twitter and Facebook for public health research [15]. Compared with traditional surveys, social media listening can not only monitor public attitudes in a timely manner, but also assess the dissemination of and public engagement with these opinions [17]. Individuals are increasingly using social media to communicate with each other, and public engagement can assess how various messages around COVID-19 vaccine spread on social media. Although several previous studies investigated public acceptance towards COVID-19 vaccines using questionnaire surveys, it is unknown how these opinions spread among the public.

By using social media listening data from the largest metropolises around the world, this study aimed to assess global vaccine hesitancy and confidence towards COVID-19 vaccination, and public engagement in communications about COVID-19 vaccines. Previous studies showed that Twitter users sharing misinformation tend to

Methods

Data collection

Twitter is one of the most popular social media platforms in the world; Sina Weibo, Twitter in China, is the most influential social media platform in China, with over 500 million users. Users can share information or opinions by tweets or posts on these platforms. Using the Meltwater platform [18], we collected posts mentioning COVID-19 vaccines on Twitter from New York (United States), London (United Kingdom), Mumbai (India), Sao Paulo (Brazil), and Sina Weibo posts from Beijing (China). The five metropolises were selected due to their high disease burden of COVID-19. The data covered the period from June 13 to July 31, 2020 when five COVID-19 vaccines started their Phase III clinical trials around the world, including the Oxford/AstraZeneca, Sinopharm, Sinovac, BioNTech/Pfizer, and Moderna vaccines [19]. The keywords included "COVID vaccin*" OR "COVID-19 vaccin*" OR "COVID19 vaccin*" OR "coronavirus vaccin*" OR "vaccin* for coronavirus" OR "vaccin* for COVID". Each post record comprised account name, contents, post time, the number of followers, and engagement data. Since our study aimed to assess public attitudes towards COVID-19 vaccination, we only included tweets or Weibo posts from individual accounts, and excluded those from news and organizational accounts. Duplicate tweets, tweets with identical text but different tweet identifications, retweets, quotes without comments, and irrelevant tweets were removed [20]. This study was exempt from ethical review because it examined retrospective, publicly-available data.

Content analysis

We identified and classified posts describing personal opinions or discussion of COVID-19 vaccines from Twitter and Sina Weibo. A coding framework (Multimedia Appendix) was developed for content analysis according to the World Health Organization's Confidence, Complacency, and Convenience ("3 Cs") model of vaccine hesitancy [21], and validated through manual annotation of the subset with 500 posts. All posts were double-coded independently, and a third-coder resolved disagreements. Posts were initially classified as relevant or irrelevant to personal opinions towards COVID-19 vaccines, and relevant posts were further classified to the predefined categories in Appendix Table 1. Predefined categories include (a) attitudes towards COVID-19 vaccination (accept, neutral, doubt, or refuse); (b) expectations of COVID-19 vaccine R&D and introduction (positive, neutral, or negative); (c) confidence in COVID-19 vaccine importance (important or not); (d) confidence in COVID-19 vaccine effectiveness (effective or not); (e) confidence in COVID-19 vaccine safety (safe or not); (f) trust in governments (trust or not); (g) trust in experts (trust or not); (h) misinformation & rumors about all vaccines; (i) complacency (perceived risk of getting COVID-19: high or low); (j) COVID-19 vaccine convenience (accessibility, distribution, or affordability); (k) COVID-19 vaccine types (AstraZeneca, Moderna, Pfizer, or Chinese vaccines); and (l) others. Each post could be classified as one category, multiple categories, or no category. We described simple counts and percentages of posts for each topic on COVID-19 vaccination.

Social media engagement analysis

We also conducted social media engagement analysis to investigate public communication and interaction about various topics relating to the COVID-19 vaccine online. In Twitter, engagement metric measures all actions viewers have taken as a result of seeing tweets and engaging with tweets [22]. At the level of public engagement and interaction, Twitter engagement refers to retweets, follows, replies, favorites, and clicks on the tweets. It covers the three metrics of social media engagement: popularity based on the number of likes for tweets, commitment based on the number of comments for tweets, and virality based on the number of shares for tweets [22]. The mean number and Standard Deviation (SD) of engagement with and followers of tweeters were presented for each topic of COVID-19 vaccine.

Results

Figure 1 shows the process of data selection and analysis. In total, we collected 12886 social media posts mentioning a COVID-19 vaccine. There were 3028 tweets on Twitter from New York, 2672 tweets from London, 2166 tweets from Mumbai, 396 tweets from Sao Paulo, and 4624 Sina Weibo posts from Beijing. After assessment, 7032 posts met the inclusion criteria and were included in our content analysis, including 1568 tweets from New York, 1440 tweets from London, 1040 tweets from Mumbai, 214 tweets from Sao Paulo, and 2770 Sina Weibo posts from Beijing.

COVID-19 vaccine hesitancy and confidence

Figure 2 presents social media users' attitudes towards COVID-19 vaccination. Among social media users, 36.4% (571/1568) in New York and 51.3% (738/1440) in London reported willingness to accept COVID-19 vaccines, much lower than observed in metropolises in lower-and-middle income countries (LMIC), such as 69.8% (726/1040) in Mumbai, 67.3% (144/214) in Sao Paulo, and 76.8% (2128/2770) in Beijing. While about 10-20% of users doubted the safety or effectiveness of COVID-19 vaccines in each metropolis, 20.0% (313/1568) in New York and 15.1% (218/1440) in London expressed refusal, much higher than other metropolises (<5%), leading to prevalent vaccine hesitancy. About 50-80% of users discussed the R&D and introduction of the COVID-19 vaccine globally, with most showing positive expectations, and the remaining users did not mention the R&D and introduction of COVID-19 vaccines (Figure 3). The discussion levels on the R&D and introduction of the COVID-19 vaccine in New York and London were lower than those in Mumbai, Beijing and Sao Paulo.

Table 1 conveys confidence, complacency, and convenience related to COVID-19 vaccination on social media. Although most tweets in each metropolis regarded COVID-19 vaccination as important and effective, proportions of tweets perceiving COVID-19 vaccine is unsafe in high income countries (HIC, 8.5% [133/1568] in New York and 10.9% [157/1440] in London) were much higher than those in other metropolises (<4%). These HIC users expressed distrust in governments (>10%) and experts (around 5%). Furthermore, 11.8% (185/1568) and 11.0% (158/1440) of users in New York and London mentioned misinformation, rumors, and anti-vaccine campaigns, but few users in LMIC metropolises (<7%) shared these concerns.

Regarding complacency, most social media users perceived the risk of getting COVID-19 to be high despite only a small number of them (around 5% except Beijing) directly mentioning the risk of getting COVID-19. For vaccination convenience, approximately 5% of HIC users mentioned vaccine accessibility, whereas more users in LMIC metropolises (10%-20%) worried about vaccine accessibility, including production and supply capacity for COVID-19 vaccines. Nearly 20% of HIC users worried about vaccine distribution, including priority vaccination groups, whereas less users in LMIC metropolises (around 10%) mentioned this subject. Few users (about 5%) considered vaccine affordability in mid-

2020. There were differential concerns on vaccine accessibility across the world.

The proportion of tweets mentioning specific vaccine types was 16.5% in New York, 57.0% in Sao Paulo, and around 30% in the remaining three metropolises. Tweets in New York mainly discussed COVID-19 vaccines produced in the United States, and tweets in other metropolises discussed vaccines produced around the world. In terms of specific COVID-19 vaccines which had started Phase III clinical trials, social media users in metropolises except for Beijing mostly mentioned the Oxford/AstraZeneca vaccine, followed by the BioNTech/Pfizer and Moderna vaccines; the Chinese vaccine was the least frequently mentioned. This indicated that public attention was consistent to the development progress of specific vaccines.

Engagement levels of COVID-19 vaccine related tweets

Table 2 presents the engagement metrics of COVID-19 vaccine-related tweets by topic. On average, tweeters who posted COVID-19 vaccine-related tweets had 3634 followers, and their followers ranged from 1637 (expressing COVID-19 vaccination is important) to 4778 followers (expressing COVID-19 vaccination is unsafe) and 7430 followers (expressing COVID-19 vaccination is ineffective). Overall, tweeters expressing negative attitudes towards COVID-19 vaccines, such as doubt or refusal to vaccinate, lack of confidence in vaccines (importance, effectiveness, and safety), and misinformation and rumors, had more followers than those expressing positive attitudes.

COVID-19 vaccine-related tweets attracted 6.1 engagement values on average. When comparing engagement levels on different topics regarding COVID-19 vaccines, we found that the most active topics were misinformation (engagement value 14.0), followed by confidence in vaccine safety (13.1), vaccine effectiveness (10.9), and governments (9.7), whereas the least active topics were vaccine convenience (2.6-5.8), confidence in vaccine importance (3.6) and experts (3.8), and complacency (5.4). For topics on vaccine hesitancy and expectations in vaccine R&D and introduction, tweets with clear attitudes attracted more engagement than neutral tweets, but there was little difference in public engagement between tweets with positive and negative attitudes. For each topic regarding vaccine confidence except distrust in governments, lack of confidence propagated more than did having confidence, especially for vaccine safety (16.7 engagement for tweets expressing vaccine is unsafe vs 10.3 engagement for tweets expressing vaccine is safe); trust in governments attracted much more engagement than distrust (52.8 vs 5.4), representing the public's common expectations of governments taking an active role in controlling the epidemic.

Discussion

Principal Findings

This social listening study in large metropolitan areas in five countries examined hesitancy towards COVID-19 vaccination, perceptions of vaccine confidence, complacency, and convenience, and level of online public engagement. We found that COVID-19 vaccine hesitancy was prevalent across the world, with the highest prevalence in New York and London, followed by Sao Paulo and Mumbai, and the lowest in Beijing. With high perceived risk of getting COVID-19, social media users in HIC, including the United States and United Kingdom, expressed low acceptance of COVID-19 vaccines, serious concerns regarding vaccine safety, and distrust in governments and experts. There were different concerns about vaccination convenience and accessibility between HIC and LMIC. Negative tweets expressing lack of vaccine confidence and misinformation or rumors had more followers and attracted more online public engagement.

Overall, social media users expressed relatively high hesitancy towards COVID-19 vaccination across metropolises with high COVID-19 burdens. There were no metropolises where willingness to accept a COVID-19 vaccine exceeded 80%, and only between 36-50% of tweeters accepted the COVID-19 vaccine especially in New York and London. These social listening results were similar to global survey studies [13,23], which showed that COVID-19 vaccine acceptance was higher in China, India, and Brazil, but lower in the United States and United Kingdom. Over the past decade, vaccine refusal has accelerated across the world, and anti-vaccine activities have been amplified through political activities and social media [24]. The prevalent hesitancy towards COVID-19 vaccines could potentially lead to low vaccination coverage, which will further delay global control of the pandemic and societal and economic recovery. Recent studies estimated a COVID-19 R_0 of around 3.87 for Europe and around 3.45 for the US [25,26], implying herd immunity thresholds of 74% and 71%, respectively. The current acceptance levels among tweeters in New York, London, Mumbai and Sao Paulo is insufficient to reach this 71-74% threshold; one exception is Beijing with an acceptance level at 76.8%. As governments are preparing to introduce COVID-19 vaccines and initiate post-pandemic recovery, the need to develop an effective vaccine campaign for the rollout that boosts public confidence and addresses hesitancy is urgent.

A previous study showed that confidence in vaccines and governments was strongly associated with vaccine acceptance and uptake [13]. However, our findings showed lack of confidence in vaccine safety, distrust in governments and experts, and widespread misinformation and rumors, especially in HIC. Public confidence in the safety and effectiveness of COVID-19 vaccines was far lower than the confidence level in general vaccines. In the 2018 Global Monitor Survey,

79.0% and 84.0% of the public agreed that vaccines were safe and effective globally [27]. During COVID-19, the development of a vaccine for a new pathogen has been pushed much faster than ever before, and new bioscience technologies (e.g. mRNA vaccine) are being used in humans for the first time [5,28]. In light of the accelerated research and development process, any negative news related to vaccine failure may weaken public confidence in its safety and effectiveness [13,29]. Therefore, effective communication campaigns should be designed to explain the safety of COVID-19 vaccines to the public and clarify misinformation and rumors, especially in HIC. Communication campaigns should be supported by the scientific community to address public concerns in COVID-19 vaccines. Through creating a space for a collaborative dialogue between the scientific community and the public, these campaigns would aim to not only provide the public with the latest information, but also build public confidence in vaccine programs.

Social media users in HIC and LMIC expressed different concerns about vaccination convenience and accessibility. HIC users had less concerns about production and supply capacity of COVID-19 vaccines, but more concerns on vaccine distribution and inequity. In contrast, LMIC users worried more about vaccine production and supply instead of vaccine distribution. While HIC can ensure the supply of COVID-19 vaccines, systemic racial and economic disparities have existed for a long time in many fields, including health care and vaccination. The COVID-19 pandemic has disproportionately affected low-income groups and communities of colour [24]. Disparities in access to COVID-19 vaccines possibly still occur and may be a point of concern that needs to be addressed in HIC. There is also much concern regarding access to COVID-19 vaccines and roll-out gaps in LMIC countries compared with HIC countries. Many HIC countries have sought to gain priority access to COVID-19 vaccines by striking advance purchase agreements with vaccine manufacturers, instead of through WHO's global allocation mechanism, leaving few vaccine doses for LMIC [23,30]. Therefore, vaccine campaigns should be tailored to each context to address local concerns: LMIC should take efforts to address vaccine supply issues, and HIC should focus more on equitable distribution within countries. This study calls for strengthened international partnerships and coordination to address the equitable access to COVID-19 vaccines across the world, and WHO should be empowered to take a leading role in guiding more preparedness actions to control the epidemic [31].

According to social media engagement analysis, we found that negative tweets had more followers and higher engagement than positive tweets. This finding might demonstrate that users on Twitter are more interested in communicating and disseminating negative messages, such as those expressing misinformation or rumors and lack of vaccine confidence. Previous studies showed that Twitter users sharing misinformation tend to be more connected and clustered [32], and false information travels faster than true information does in social networks [33]. Although scientific experts received a considerable attention early on in the COVID-19 pandemic, online attention shifted towards political communities as the pandemic developed [34]. During the evolution of the pandemic, scientific experts lost some of their influence, and it became harder for scientific information to reach a broad audience. In addition, the spread of misinformation and rumors on social media has been found to be significantly associated with vaccine hesitancy [35]. Social media is therefore a double-edged sword that can help disseminate public health knowledge directly to the public, but which can also, through inappropriate use, be destructive to public health efforts, especially during a public health emergency [36-38]. Therefore, more efforts are needed to build a more proactive public health presence on social media, and health systems should "listen" to tweets from the public to help inform policies related to public health response.

Limitations

This study captured routine populations who may not be represented in traditional research designs, and social media data can eliminate reporting bias that occurs from speaking with a researcher [20]. However, this study has some limitations. First, there is an inherent bias shared among all social media studies, where users might present themselves differently online and/or represent a skewed-younger population [39]. Second, we did not extract demographic data because of difficulty to refer to users' profiles. Third, we used manually coding methods rather than automated annotation, which increased the length of time taken to annotate. Therefore, we only assessed vaccine hesitancy in large metropolises in a short period, and it reduced the generalizability of our findings. In the future, based on our coding dataset from this study, we plan to develop a machine learning programme to continuously track public attitudes towards COVID-19 vaccination.

Author Contributions: Concept and design: Hou, Lin. Acquisition, analysis, or interpretation of data: All authors. Statistical analysis: Tong. Administrative, technical, or material support: Hou, Piatek, Larson. Drafting of the manuscript: Hou, Tong. Critical revision of the manuscript for important intellectual content: Hou, Lin, Larson. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Acknowledgments: The authors thank Xiaoyu Zhou, Qian Lv, Simeng Hu from School of Public Health, Fudan University, and Annika Wilder-Smith from Edinburgh Medical School for their help on data analysis. The authors also thank Fiona Yueqian Sun from Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine for her help on Twitter data collection, and Xinyu Zhou from School of Public Health, Fudan University for his help on Sina Weibo data collection.

Funding: Zhiyuan Hou acknowledges financial support from the National Natural Science Foundation of China (No. 71874034) and the National Key R&D Program of China (No. 2018YFC1312600 and 2018YFC1312604). Leesa Lin was supported by the UK Foreign, Commonwealth and Development Office and Wellcome (Grant no. 215373/Z/19/Z). The funders had no role in the study design, data collection, data analysis, data interpretation, or writing of the report.

Declaration of interests: Heidi Larson is on the Merck Vaccine Confidence Advisory Board. Her research group, the Vaccine Confidence Project, received research grants from GSK and Merck on vaccine confidence issues. None of those research grants are related to this paper.

Ethics: All data are publicly available.

Reference

1. Coronavirus disease (COVID-19) pandemic: World Health Organization; 2020 [updated 2020-12-23]. Available from: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>.
2. Coronavirus disease (COVID-19) advice for the public: World Health Organization; 2020 [updated 2020-12-23]. Available from: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public#:~:text=If%20COVID%2D19%20is%20spreading,a%20bent%20elbow%20or%20tissue>.
3. Chou WS, Budenz A. Considering emotion in COVID-19 vaccine communication: addressing vaccine hesitancy and fostering vaccine confidence. *Health Commun.* 2020;35(14):1718-22. PMID: 33124475
4. Gao Q, Bao L, Mao H, Wang L, Xu K, Yang M, et al. Development of an inactivated vaccine candidate for SARS-CoV-2. *Science.* 2020;369(6499):77-81. PMID: 32376603
5. Haque A, Pant AB. Efforts at COVID-19 vaccine development: challenges and successes. *Vaccines.* 2020;8(4):739. PMID: 33291245
6. Larson HJ, Smith DM, Paterson P, Cumming M, Eckersberger E, Freifeld CC, et al. Measuring vaccine confidence: analysis of data obtained by a media surveillance system used to analyse public concerns about vaccines. *The Lancet Infectious Diseases.* 2013;13(7):606-13. PMID: 23676442
7. Bocquier A, Fressard L, Cortaredona S, Zaytseva A, Ward J, Gautier A, et al. Social differentiation of vaccine hesitancy among French parents and the mediating role of trust and commitment to health: A nationwide cross-sectional study. *Vaccine.* 2018;36(50):7666-73. PMID: 30391054
8. Larson HJ, Jarrett C, Schulz WS, Chaudhuri M, Zhou Y, Dube E, et al. Measuring vaccine hesitancy: The development of a survey tool. *Vaccine.* 2015;33(34):4165-75. PMID: 25896384
9. Chung Y, Schamel J, Fisher A, Frew PM. Influences on Immunization Decision-Making among US Parents of Young Children. *Maternal and Child Health Journal.* 2017;21(12):2178-87. PMID: 28755045
10. Bianco A, Mascaro V, Zucco R, Pavia M. Parent perspectives on childhood vaccination: How to deal with vaccine hesitancy and refusal? *Vaccine.* 2019;37(7):984-90. PMID: 30655175
11. Brown AL, Sperandio M, Turssi CP, Leite RMA, Berton VF, Succi RM, et al. Vaccine confidence and hesitancy in Brazil. *Cadernos De Saude Publica.* 2018;34(9):e00011618. PMID: 30281705
12. Paterson P, Meurice F, Stanberry LR, Glismann S, Rosenthal SL, Larson HJ. Vaccine hesitancy and healthcare providers. *Vaccine.* 2016;34(52):6700-6706. PMID: 27810314.
13. Lazarus JV, Ratzan SC, Palayew A, Gostin LO, Larson HJ, Rabin K, et al. A global survey of potential acceptance of a COVID-19 vaccine. *Nature Medicine.* 2021;27(2):225-228. PMID: 33082575
14. Shah Z, Surian D, Dyda A, Coiera E, Mandl KD, Dunn AG. Automatically appraising the credibility of vaccine-related web pages shared on social media: a Twitter surveillance study. *J Med Internet Res.* 2019;21(11):e14007. PMID: 31682571
15. Sinnenberg L, Buttenheim AM, Padrez K, Mancheno C, Ungar L, Merchant RM. Twitter as a tool for health research: a systematic review. *Am J Public Health.* 2017;107(1):e1-e8. PMID: 27854532
16. Steffens MS, Dunn AG, Wiley KE, Leask J. How organisations promoting vaccination respond to misinformation on social media: a qualitative investigation. *BMC Public Health.* 2019;19(1):1348. PMID: 31640660
17. Milinovich GJ, Williams GM, Clements AC, Hu W. Internet-based surveillance systems for monitoring emerging infectious diseases. *The Lancet Infectious diseases.* 2014;14(2):160-8. PMID: 24290841
18. Meltwater. [updated 2020-12-23]. Available from: <https://www.meltwater.com/>.
19. The COVID-19 candidate vaccine landscape: World Health Organization; 2020 [updated 2020-12-23]. Available from: <https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines>.
20. Golder S, O'Connor K, Hennessy S, Gross R, Gonzalez-Hernandez G. Assessment of Beliefs and Attitudes About Statins Posted on Twitter: A Qualitative Study. *JAMA network open.* 2020;3(6):e208953. PMID: 32584408
21. MacDonald NE. Vaccine hesitancy: Definition, scope and determinants. *Vaccine.* 2015;33(34):4161-4. PMID: 25896383
22. Stone JA, Can SH. Linguistic analysis of municipal twitter feeds: Factors influencing frequency and engagement. *Government Information Quarterly.* 2020; 37(4): 101468. <https://www.sciencedirect.com/science/article/pii/S0740624X19304873>
23. Wouters OJ, Shadlen KC, Salcher-Konrad M, et al. Challenges in ensuring global access to COVID-19 vaccines: production, affordability, allocation, and deployment. *The Lancet.* 2021;397(10278):1023-1034. PMID: 33587887
24. Hotez PJ, Cooney RE, Benjamin RM, et al. Announcing the Lancet Commission on vaccine refusal, acceptance, and demand in the USA. *The Lancet.* 2021; 397(10280): 1165-7. PMID: 33639088
25. Flaxman S, Mishra S, Gandy A, Unwin H, Coupland H, Mellan T, et al. Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in European countries: technical description update. 2020. <https://arxiv.org/abs/2004.11342>
26. Pitzer VE, Chitwood M, Havumaki J, Menzies NA, Perniciaro S, Warren JL, et al. The impact of changes in diagnostic testing practices on estimates of COVID-19 transmission in the United States. *Am J Epidemiol.* 2021; kwab089. PMID: 33831148.

27. Gallup. Wellcome Global Monitor: How does the world feel about science and health? First wave findings. 2019. <https://static4.libertatea.ro/wp-content/uploads/2019/06/wellcome-global-monitor-2018.pdf>
28. Kamidani S, Rostad CA, Anderson EJ. COVID-19 vaccine development: a pediatric perspective. *Curr Opin Pediatr*. 2021;33(1):144-151. PMID: 33278108
29. Mahase E. Covid-19: Oxford researchers halt vaccine trial while adverse reaction is investigated. *BMJ*. 2020;370:m3525. PMID: 32907856
30. Lancet Commission on COVID-19 Vaccines and Therapeutics Task Force Members. Urgent needs of low-income and middle-income countries for COVID-19 vaccines and therapeutics. *The Lancet*. 2021; 397(10274): 562-4. PMID: 33516284
31. Hou Z, Du F, Zhou X, Jiang H, Martin S, Larson H, et al. Cross-Country Comparison of Public Awareness, Rumors, and Behavioral Responses to the COVID-19 Epidemic: Infodemiology Study. *J Med Internet Res*. 2020;22(8):e21143. PMID: 32701460
32. Pierri F, Piccardi C, Ceri S. Topology comparison of Twitter diffusion networks effectively reveals misleading information. *Sci Rep*. 2020;10(1):1372. PMID: 31992754
33. Vosoughi S, Roy D, Aral S. The spread of true and false news online. *Science*. 2018;359(6380):1146-1151. PMID: 29590045
34. Durazzi F, Müller M, Salathé M, Remondini D. International expert communities on Twitter become more isolated during the COVID-19 pandemic. 2020. <https://arxiv.org/abs/2011.06845>
35. Aquino F, Donzelli G, De Franco E, Privitera G, Lopalco PL, Carducci A. The web and public confidence in MMR vaccination in Italy. *Vaccine*. 2017;35:4494-98. PMID: 28736200
36. Pagoto S, Waring ME, Xu R. A call for a public health agenda for social media research. *J Med Internet Res*. 2019;21(12):e16661. PMID: 31855185
37. Abd-Alrazaq A, Alhuwail D, Househ M, Hamdi M, Shah Z. Top concerns of Tweeters during the COVID-19 pandemic: Infoveillance study. *J Med Internet Res*. 2020;22(4):e19016. PMID: 32287039
38. Wilson SL, Wiysonge C. Social media and vaccine hesitancy. *BMJ Global Health*. 2020;5:e004206. PMID: 33097547
39. Sloan L, Morgan J, Burnap P, Williams M. Who tweets? Deriving the demographic characteristics of age, occupation and social class from twitter user meta-data. *PloS ONE*. 2015;10(3):e0115545. PMID: 25729900

Figure legend

Figure 1. Flowchart of data process and analysis

Figure 2. Social media users' attitudes towards COVID-19 vaccination

Figure 3. Expectations of COVID-19 vaccine R&D and introduction

Table 1. Confidence, complacency and convenience related to COVID-19 vaccine on social media, 2020, n(%)

Number and topics	New York, US	London, UK	Beijing, China	Mumbai, India	Sao Paulo, Brazil
Number of total posts^a	3028	2672	4624	2166	396
Irrelevant posts	1460	1232	1854	1126	182
Relevant posts	1568	1440	2770	1040	214
Topics among relevant posts^b					
<i>Vaccine confidence</i>					
Perceived importance of vaccine	115 (7.3)	130 (9.0)	712 (25.7)	140 (13.5)	22 (10.3)
Important	96 (6.1)	115 (8.0)	651 (23.5)	132 (12.7)	22 (10.3)
Unimportant	19 (1.2)	16 (1.1)	61 (2.2)	8 (0.8)	0 (0.0)
Perceived effectiveness of vaccine	217 (13.8)	361 (25.1)	928 (33.5)	294 (28.3)	71 (33.2)
Effective	142 (9.1)	305 (21.2)	674 (24.3)	261 (25.1)	68 (31.8)
Ineffective	75 (4.8)	56 (3.9)	254 (9.2)	32 (3.1)	3 (1.4)
Perceived safety of vaccine	168 (10.7)	354 (24.6)	201 (7.3)	190 (18.3)	47 (22.0)
Safe	35 (2.2)	197 (13.7)	143 (5.2)	149 (14.3)	43 (20.1)
Unsafe	133 (8.5)	157 (10.9)	58 (2.1)	41 (3.9)	4 (1.9)
Trust in governments	294 (18.8)	177 (12.3)	148 (5.3)	72 (6.9)	8 (3.7)
Trust	14 (0.9)	19 (1.3)	138 (5.0)	16 (1.5)	1 (0.5)
Distrust	280 (17.9)	158 (11.0)	10 (0.4)	56 (5.4)	7 (3.3)
Trust in experts	90 (5.7)	111 (7.7)	-	78 (7.5)	2 (0.9)
Trust	14 (0.9)	44 (3.1)	-	24 (2.3)	2 (0.9)
Distrust	76 (4.8)	67 (4.7)	-	54 (5.2)	0 (0.0)
<i>Information around vaccine</i>					
Misinformation & rumors	185 (11.8)	158 (11.0)	188 (6.8)	30 (2.9)	12 (5.6)
<i>Complacency</i>					
Perceived risk of getting COVID-19	89 (5.7)	52 (3.6)	742 (26.8)	45 (4.3)	4 (1.9)
High	80 (5.1)	45 (3.1)	688 (24.8)	41 (3.9)	4 (1.9)
Low	9 (0.6)	7 (0.5)	54 (1.9)	4 (0.4)	0 (0.0)
<i>Vaccine convenience</i>					
Vaccine accessibility	94 (6.0)	76 (5.3)	283 (10.2)	99 (9.5)	49 (22.9)
Vaccine distribution	309 (19.7)	261 (18.1)	325 (11.7)	63 (6.1)	26 (12.1)
Vaccine affordability	107 (6.8)	40 (2.8)	197 (7.1)	33 (3.2)	20 (9.3)
<i>Vaccine types</i>	259 (16.5)	470 (32.6)	941 (34.0)	270 (26.0)	122 (57.0)
AstraZeneca	102 (6.5)	388 (26.9)	82 (3.0)	218 (21.0)	63 (29.4)
Moderna	51 (3.3)	30 (2.1)	37 (1.3)	31 (3.0)	21 (9.8)
Pfizer	88 (5.6)	26 (1.8)	34 (1.2)	18 (1.7)	27 (12.6)
Chinese vaccines	25 (1.6)	26 (1.8)	831(30.0)	19 (1.8)	13 (6.1)
<i>Others</i>	83 (5.3)	51 (3.5)	990 (35.7)	34 (3.3)	5 (2.3)

Note: ^a We assessed 50% random samples from New York and London due to large sample size, and full samples in Beijing, Mumbai and Sao Paulo.

^b Topics are calculated among relevant posts.

Table 2. Followers and engagements of COVID-19 vaccine related tweets, 2020, mean(SD)

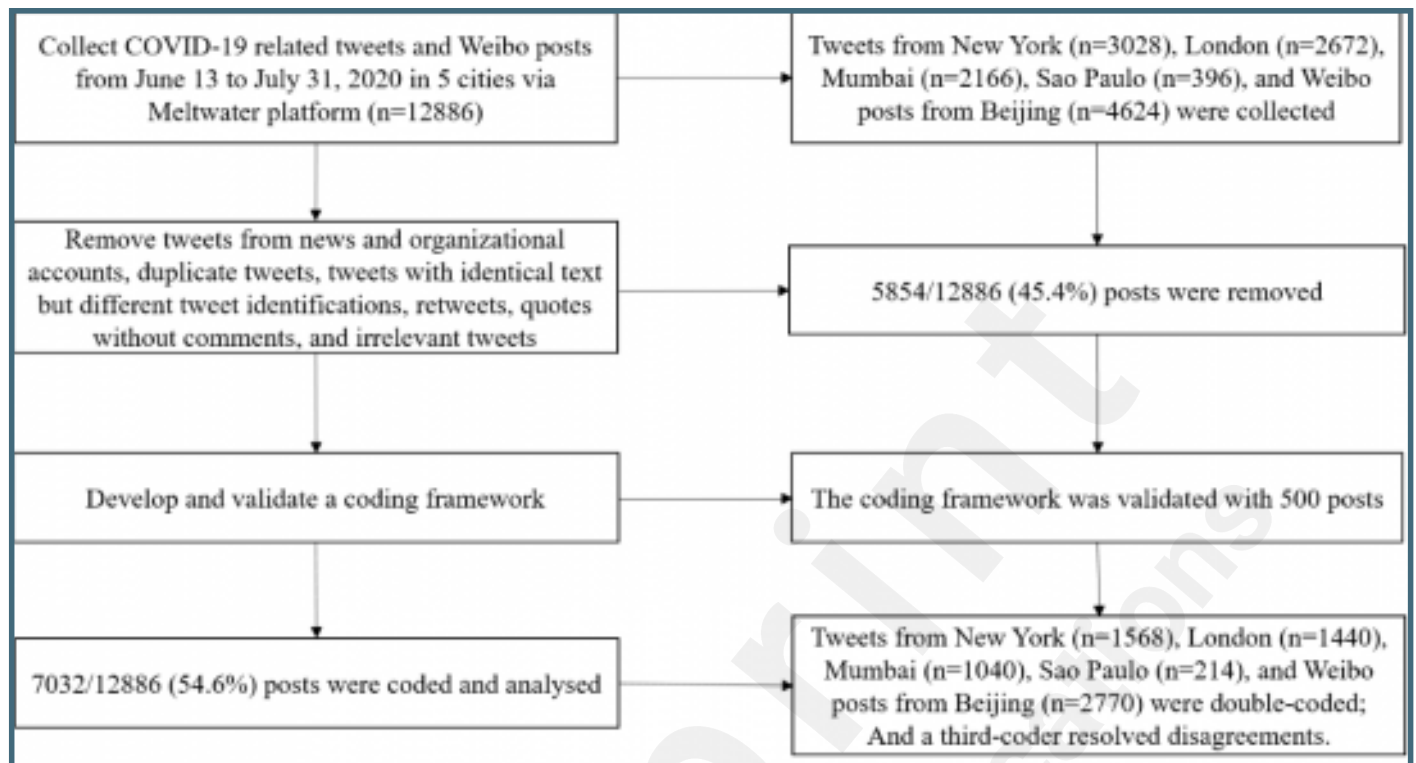
Topics	Followers	Engagements
<i>Vaccine hesitancy</i>		
Attitudes towards COVID-19 vaccination	3633.6 (12614.4)	6.1 (61.4)
Accept	2448.8 (8614.6)	7.9 (93.3)
Neutral	4333.0 (16250.7)	5.8 (28.5)
Doubt	4593.3 (18245.7)	10.0 (93.4)
Refuse	2708.4 (12504.3)	6.4 (35.4)
Expectations of COVID-19 vaccine R&D and introduction	3011.0 (11759.4)	8.4 (94.4)
Positive	2527.7 (9083.2)	8.8 (104.3)
Neutral	4449.8 (16399.9)	5.9 (31.9)
Negative	3629.4 (14833.2)	8.2 (84.0)
<i>Vaccine confidence</i>		
Perceived importance of vaccine	1763.6 (3727.9)	3.6 (10.7)
Important	1636.5 (3167.2)	3.4 (9.9)
Unimportant	2843.1 (6811.7)	5.4 (16.5)
Perceived effectiveness of vaccine	3373.4 (13960.7)	10.9 (115.5)
Effective	2505.6 (8366.8)	10.6 (119.5)
Ineffective	7430.1 (27615.6)	12.2 (94.7)
Perceived safety of vaccine	3274.3 (14058.0)	13.1 (137.2)
Safe	2086.2 (6191.7)	10.3 (137.4)
Unsafe	4778.0 (19897.0)	16.7 (137.1)
Trust in governments	3865.2 (16363.8)	9.7 (111.8)
Trust	3055.6 (6573.5)	52.8 (360.4)
Distrust	3946.0 (17036.6)	5.4 (28.7)
Trust in experts	1887.9 (4812.1)	3.8 (14.7)
Trust	2029.1 (4076.9)	3.0 (6.5)
Distrust	1827.7 (5102.0)	4.1 (17.1)
<i>Information around vaccine</i>		
Misinformation & rumors	3127.7 (11610.5)	14.0 (121.2)
<i>Vaccine complacency</i>		
Perceived risk of getting COVID-19	3017.9 (8282.9)	5.4 (15.6)
High	2909.7 (7848.9)	5.1 (15.5)
Low	3937.6 (11554.1)	8.1 (16.5)
<i>Vaccine convenience</i>		
Vaccine accessibility	2351.5 (8915.4)	2.6 (7.8)
Vaccine distribution	3531.5 (12541.6)	5.8 (26.2)
Vaccine affordability	2525.3 (4895.6)	3.1 (8.3)

Note: SD: Standard Deviation.

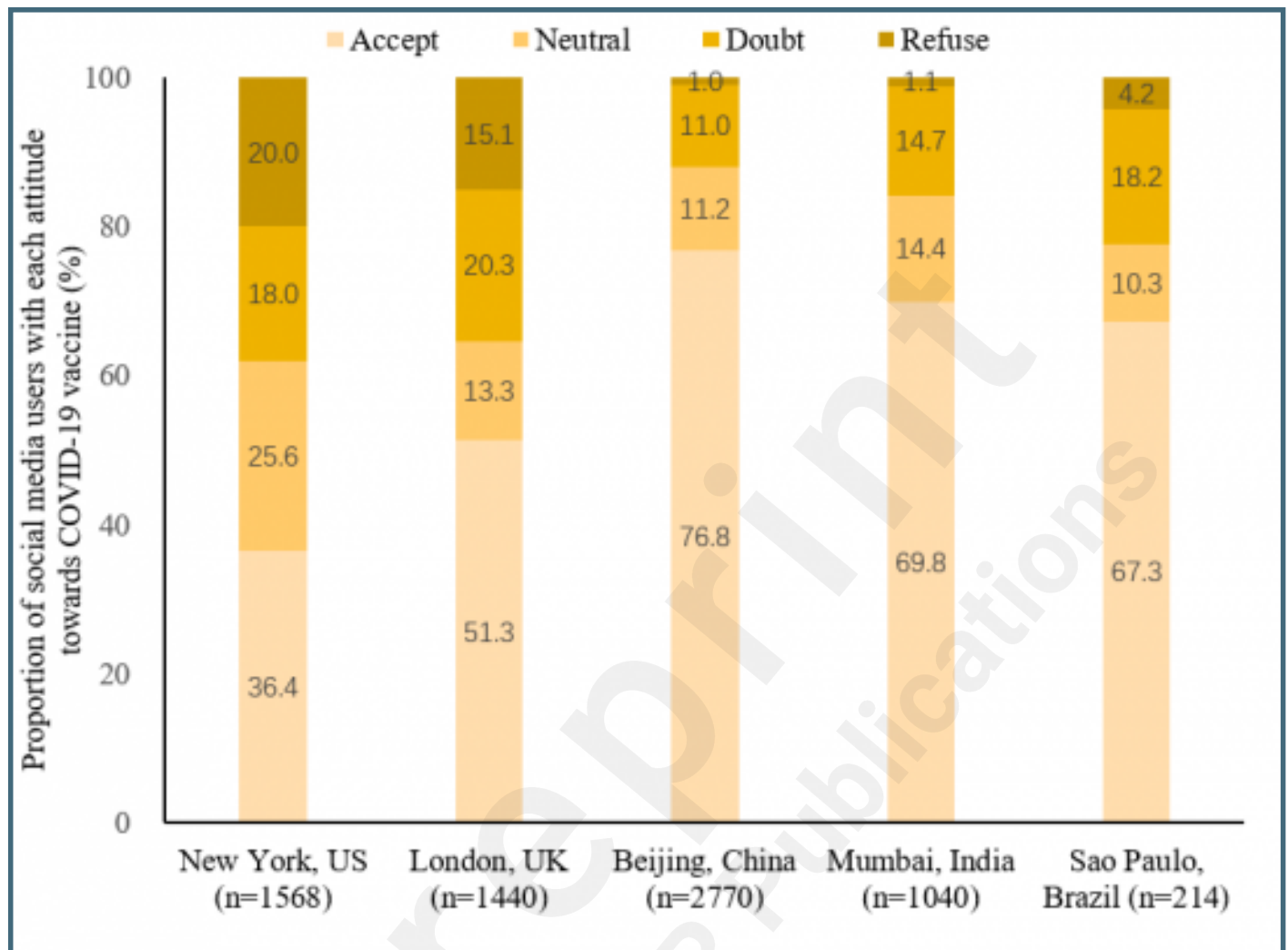
Supplementary Files

Figures

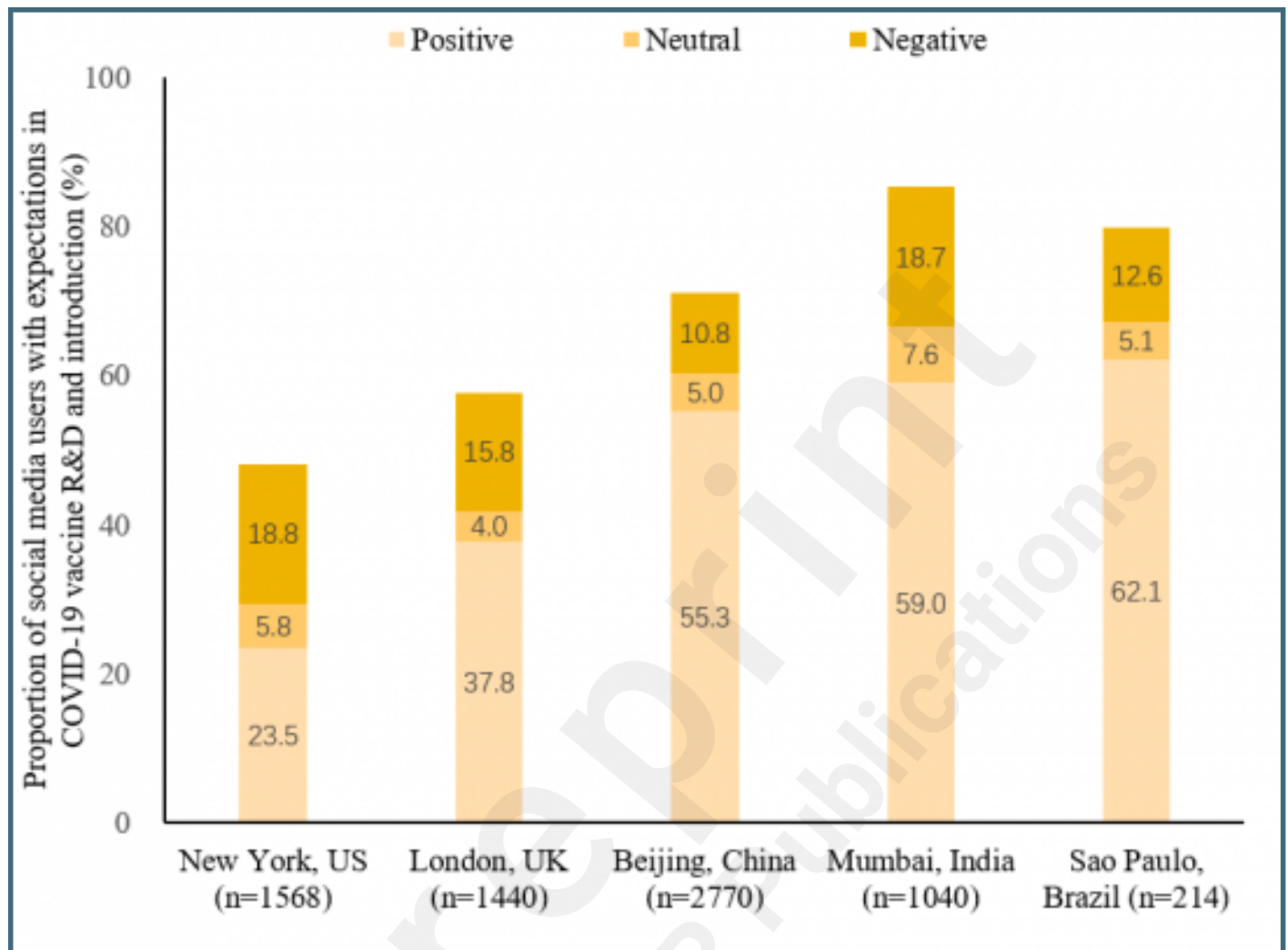
Flowchart of data process and analysis.



Social media users' attitudes towards COVID-19 vaccination.



Expectations of COVID-19 vaccine R&D and introduction.



Multimedia Appendixes

Coding framework for COVID-19 vaccine posts on social media.

URL: <http://asset.jmir.pub/assets/061c9ef1a710ce67706da921a8760a52.docx>

