Tracking the impact of COVID-19 and lockdown policy on public mental health using social media

Minghui Li
Zhejiang University https://orcid.org/0000-0003-3673-5925
Yining Hua
Harvard Medical School
Yanhui Liao
Zhejiang University
Li Zhou
Brigham and Women's Hospital, Harvard Medical School
Xue Li
Zhejiang University
Ling Wang
King's College London
Jie Yang (jieynlp@gmail.com)
Zhejiang University https://orcid.org/0000-0001-5696-363X

Article

Keywords:

Posted Date: March 29th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1498436/v1

Tracking the impact of COVID-19 and lockdown policy on public mental health using social media

Minghui Li[1], Yining Hua[2,3], Yanhui Liao[4], Li Zhou[2,3], Xue Li[1], Ling Wang[5], Jie Yang[1]

[1] Department of Big Data in Health Science School of Public Health, and Center of Clinical Big Data and Analytics of The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

[2] Harvard Medical School, Boston, Massachusetts, USA

[3] Division of General Internal Medicine and Primary Care, Department of Medicine, Brigham and Women's Hospital, Boston, MA

[4] Department of Psychiatry, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, China

[5] Florence Nightingale Faculty of Nursing, Midwifery & Palliative Care, King's College London, London, UK

Corresponding author:

Jie Yang, Ph.D.

Department of Big Data in Health Science School of Public Health, and Center of Clinical Big Data and Analytics of The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

Address: 866 Yuhangtang Rd, Hangzhou, P.R. China 310058

Email: y@zju.edu.cn

Telephone: +86-19157731185

Manuscript word count: 2807 words

Abstract

The COVID-19 pandemic and its corresponding preventive and control measures have increased the mental burden on the public. Social media serve as important platforms to timely track public mental status. In this study, we conducted social-media-based analyses on temporal, geographical and occupational distributions of public mental health status during the pandemic, and how the public reacted to the lock-down policy from the perspective of mental health. We extracted 2,973,319 mental health-related tweets of 1,778,140 users from February 1, 2020 to September 30, 2021. We found that, compared to the general public, healthcare workers had higher concerns on three types of mental health problems (depression, insomnia, addiction) (P<0.001) and focused more on clinical topics while the public worried more about daily life issues. The lockdown policy in New York was correlated with a proportional decrease of mental health-related tweets, while Florida had an opposite correlation (both P<0.05). Our findings indicated that the mental burden brought by the pandemic varied across occupations and locations and changed over time.

Introduction

The persistent coronavirus disease 2019 (COVID-19) pandemic has drastically changed people's daily lives since the first confirmed case in December 2019[1]. As of February 25, 2022, there had been over 428 million confirmed COVID-19 cases and about 6 million deaths worldwide[2]. The pandemic not only led to high hospitalization and fatality but also brought negative impacts on public mental health[3,4]. The reasons that led to mental health problems during the pandemic include but are not limited to the infection and death of relatives and friends, the fear of infection, quarantine and isolation[5,6], and the stress from unemployment[7]. Using a national cross-sectional survey, Gualano et al[8] reported that the prevalence of depression (24.7%), anxiety (23.2%), and sleep disturbances (42.2%) were high in Italy. Moreover, certain subpopulations such as children and adolescents[9,10], students[11,12], COVID-19 patients[13], and healthcare workers[14,15] are particularly vulnerable to psychological disorders.

Healthcare workers suffered from more mental stresses due to the higher exposure risk to the virus and overwhelming workload during the pandemic[16,17]. To reduce burnout and mental pressure, it is important to track their mental status, identify their general concerns, and provide timely support to healthcare workers[18,19]. Pappa et al. conducted a meta-analysis containing 13 cross-sectional studies (12 studies in China) and reported the pooled prevalence of three psychological symptoms among healthcare workers: anxiety (23.2%), depression (22.8%), and insomnia (34.32%)[20]. Similarly, Firew et al[21] studied the psychological distress of healthcare workers in the United States of America (USA) and reported high levels of anxiety, depression, and burnout symptoms. However, existing researches on healthcare workers' mental health are mostly cross-sectional[15,17,20-23], while the dynamic characteristics remain less studied.

Due to their large scale, immediacy, and wide coverage, social media platforms (such as Twitter, Facebook, Instagram, and Weibo) have been vital data sources for analyzing public perception, discussion, and sentiment. With the features of large scale and timely availability, they provide opportunities to timely monitor the public health status[24]. There have been a few studies applying social media analysis for content analysis[25], surveillance[26], and early warning[27] during the outbreak of infectious diseases. During the 2009 H1N1 outbreak, Chew and Eysenbach filtered the HIN1-related tweets on Twitter and found that there were about 4.5% of tweets were identified as misinformation[25]. During the Zika outbreak in 2016, Masri et al. tracked the Zika related tweets for disease surveillance and found that the Zika new cases can be predicted one week ahead based on the related tweets[26]. After the outbreak of COVID-19, a series of studies used social media data to monitor public perceptions on topics such as enforced remote work[28] and vaccines[29,30], and analyze public sentiments[31-36]. Topic model and sentiment analysis were widely used in such studies. For example, Zhang et al. found topic themes related to remote work tweets involved cybersecurity, work-life balance, teamwork, and leadership[28]. Xie et al. analyzed the discussions about COVID-19 vaccines on Twitter and found that positive public responses were related to the positive progress of vaccine development, while negative emotions mainly came from concerns of vaccine availability[29].

The main objective of this study is to analyze public mental health status, common symptoms, and their temporal and geographic distributions during the COVID-19 pandemic. In particular, we aim to investigate the following research questions: 1) What mental health-related topics did the public concern the most and how did their discussions change over time? 2) Are there differences in mental health concerns between the general population and healthcare workers? 3) Do lockdown policies have significant impacts on public mental health?

We filtered COVID-19 related tweets with keywords (**Supplementary Table 1**) that can be categorized into four common mental health problems: anxiety, depression, insomnia, and addiction. We visualized the distributions of mental health-related tweets by time and by the states of the USA. We used a topic model to summarize these tweets into 17 topics, and compared the topic distributions among healthcare workers and the general population. Finally, we investigated the influence of lockdown policies on public mental health in different states of the USA.

Results

Dataset structure

A total number of 2,973,319 mental health-related tweets (from 1,778,140 users) were selected after the data preprocessing steps (**Fig. 1**). Among them, 46.85% tweets (1,393,025) were in the "depression" group, 43.41% tweets (1,290,793) were in the "anxiety" group, 8.91% tweets (265,002) were in the "insomnia" group, and 3.32% tweets (98,791) were in the "addiction" group.[1] Additionally, 562,277 (18.91%) tweets were extracted with their geographical information and 40,798 (1.37%) tweets (from 21,520 users) were posted by healthcare workers.

Temporal distribution of mental health-related tweets

[1] The sum of the four proportions was larger than 100% because some tweets included multiple keywords which belong to different mental health subgroups.

The trends of the weekly numbers of mental health-related tweets in four subgroups are shown in **Fig. 2a**. The tweet numbers of mental health problems reached a peak from February 29 to April 4, 2020. To eliminate the effect of tweets number fluctuation, we calculated and visualized the proportions of mental health-related tweets among all COVID-19 related tweets in **Fig. 2b**. The proportion curve of anxiety-related tweets had a dominant peak in Mar 2020. And both the number and the proportion of anxiety-related tweets had a remarkable increase in September 2021.

Geographical distribution of mental health-related tweets in the USA

Fig. 3a shows the proportion of mental health-related tweets among all COVID-19 related tweets in each state of the USA from February 1, 2020 to September 30, 2021. Massachusetts, Oregon, and Vermont were three states that had the highest proportions of mental health-related tweets, while West Virginia, Kansas, and Mississippi were the ones with the lowest proportions. **Fig. 3b** visualizes the monthly tweet proportion for all the 50 USA states. The first two months had a larger proportion of mental health-related tweets compared with the following months across most of the states. Compared with other states, the proportion of mental health-related tweets in Vermont was higher from February to December 2020, and West Virginia kept a low proportion during the whole period of the pandemic.

Topics of mental health-related tweets

The most frequent words and word bigrams of mental health-related tweets are "people", "worry", "shame", "panic", "lockdown", "panic_buying", "wear_mask", etc., which were shown in **Supplementary Fig. 1**. We chose 17 to be the number of topics based on the perplexity and coherence (**Supplementary Methods**). Topics and the corresponding top 20 most probable unigrams and bigrams are displayed in **Supplementary Table 2**. We concluded each topic with a topic name based on these words. For example, the topic included keywords "stress", "school", "student", "kid", "exam" indicates that tweets in this topic probably focused on "education". Except for the topics related to COVID-19 itself, such as "social distancing", "panic attack", "COVID case", the public also showed special interests in topics such as "economic collapse”, “presidential election” and “vaccine”. The 17 topics were then categorized into 6 topic groups: “COVID-19 pandemic”, “preventive measures”, “concerns”, “economics”, “politics”, “education”. **Fig. 4** shows the dynamic distributions of the investigated topics in relative tweet proportions. The topic “family concerns” occupied a dominant position during most time of the pandemic. “Social distancing” was frequently mentioned at the beginning of the pandemic, but it went back to a normal level after June 2020. The topic proportion of “hospital situations” notably fluctuated in the whole research period and was relatively large from February 2020 to March 2020, and from August 2021 to September 2021.

The mental health of healthcare workers

We assessed the differences in the proportions of four mental health symptoms-related tweets between healthcare workers and the general population in **Table 1**. Statistical results showed that proportions of “depression”, “insomnia” and “addiction” related tweets were significantly higher in healthcare workers than that in the general public (all $P<0.001$), while the proportion of anxiety-related tweets did not show a significant difference between the two groups. **Fig. 5a** shows the distribution of average number of tweets per user on different topics. “family concerns” is the top topic which discussed in both healthcare workers and the general population. To visualize the difference of the topic distribution between healthcare workers and the general population, we showed the ratios of the average number of tweets by topics of the two groups in **Fig.5b**. It demonstrated that healthcare workers discussed more on 12 topics, especially in clinical-related topics such as “COVID death”, “COVID symptoms” and “test result”, while the general population concentrated on topics such as “hospital situations”, “mental concerns” and “education”.

Impacts of the lockdown policy

We selected tweets from California, New York, Texas, and Florida to explore the effect of the lockdown policy on public mental status because these four states had the largest number of mental health-related tweets. As **Fig. 6** shows, the proportions of the four mental health-related tweets changed after the lockdown policy was effected in

California, New York, and Florida, but not in Texas. **Table 2** lists the results of interrupted time series analyses[37] of the lockdown policy on public mental health. The coefficient of “intervention”, meaning the change of intercept, was significant in the model of New York and Florida. That is to say, the proportion of mental health-related tweets decreased in New York after the lockdown policy on March 22, 2020, whereas in Florida it increased on April 3, 2020, with both $P<0.05$. For California and Texas, there was no statistically significant change in proportions after the lockdown policy.

Discussions

The pandemic itself, strict lockdown policies, economic collapse, and the pressure of employment, had severely affected public mental health. Our study tracked the public mental health status in a long period of the COVID pandemic, explored the discussion themes of related tweets, examined the differences between healthcare workers and the general population, and analyzed the influence of lockdown policy on public mental health.

We observed the initial evidence of the psychological numbing during the COVID-19 pandemic on social media data with large numbers of users in this study. The proportion of anxiety-related tweets climbed to a significant peak in Mar 2020 and backed to normal quickly in the initial stage of the COVID-19 pandemic, then kept in a stable tendency after the first peak between February 29 to April 4, 2020. A possible explanation is that: in the initial stage, the outbreak of COVID-19 caused various social problems, such as the shortage of necessities and unemployment, these problems raised a temporal but intense fear to the public. As the pandemic continues, the public concerns fall back to the normal level. This observation is consistent with the study finding of Joel Dyer and Blas Kolic that public emotional response diminishes as the pandemic intensifies[34].

The proportion of “depression” related tweets slightly increased in March 2020 and was then kept at a stable level. The history of mental health problems[38-40] and the financial worries[38,39] played a vital role in the depression emotions of populations. The proportion of “insomnia” related tweets climbed continuously until February 2020 and fluctuated then (**Supplementary Fig. 2**), while the proportion of "addiction" kept stable during the whole research period. Some affecting factors of insomnia referred to the COVID symptoms, the history of mental health problems, loneliness, social support, financial worries, etc[41-43].

Our topic model showed 6 groups of topics indicating the main concerns of populations: "COVID-19 pandemic", "preventive measures", "concerns", "economics", "politics", "education". This shows that the public was not only concerned about the pandemic and its prevention, but also the economic and educational problems caused by COVID-19. Compared with the previous work[32,33,35,36,44,45], topics such as "discriminatory names", "developing bioweapon", "racial hostility", "racism", "increased racism", "travel bans and warnings", "travel restrictions" were not observed in our study. Those works focused on the tweets on the early stage of pandemic (mostly from January to August 2020), while our study covered a longer period. Topics in our study, such as "social distancing", "COVID case", "test results", "world pandemic", "negative emotions", "COVID news", "economic collapse", "vaccine" were also observed in previous work [32,33,35,36,44,45], indicating that these topics are a common focus of the public. In addition to the above common discussed topics, our study found two new topics: "presidential election" and "education". Our data collection included tweets during the presidential election(December 3, 2020), and COVID-19 was one of the main topics in the election debates. The education system was severely disturbed by the lockdown policy, and students (especially children and adolescents) were more vulnerable to psychological disorders[46]. Leeb et al. found the proportion of emergency department visits among children increased until October 2020 while the number decreased in case of the lockdown policy[10]. The monthly proportions of tweets in several topics changed over time obviously. People paid attention to the topic "social distancing" at the initial 5 months and reduced the concerns gradually, and a higher proportion of "hospital situations" often related to the new epidemic.

Compared with the general population, healthcare workers had more concerns about the mental health problems "depression", "insomnia" and "addiction" during the pandemic, while the proportion of “anxiety” tweets was similar in the two groups. Healthcare professionals with medical training and clinical practice experience are more likely to discuss the pandemic from the scientific point of view. Healthcare workers were more likely to suffer from mental health problems as they have a higher likelihood of being infected[16] and suffering severe environmental pressure. They may also get mental burnout due to issues like the lack of personal protective equipment (particularly during the early pandemics) and increased workload. One potential reason for higher proportion of “addiction” among healthcare workers is that they tend to share information about their patients or call on the public to reduce the substance and behavior addiction on social media. In previous work[47-50], the prevalence of “anxiety” and “depression” was compatible between healthcare workers and the general population. The inconsistency in the prevalence of “depression” may be caused by the regional disparity, given that the pandemic of COVID-19 in the USA was serious and healthcare workers there suffered more workload. However, many studies reported a higher prevalence of “insomnia” among healthcare workers than the general population[48,50-53], in accord with our result. In addition, compared with the general population, healthcare workers focused more on discussing the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus and preferred to share COVID-related news or experiences.

The lockdown policy had various effects on mental health discussions among different states of the USA. The lockdown policy of California and New York had reduced public mental health discussions, while this phenomenon was not observed in Texas and Florida. A possible explanation is that California and New York adopted prevention measures and implemented the stay-at-home policy without delay, so people witnessed a positive effect brought by the lockdown policy. Texas and Florida applied a lockdown policy about two weeks later after the pandemic had become severe, as a result, the policy might have exacerbated public panic. Mittal et al. found that most Twitter users shared positive opinions toward lockdown policies in related tweets from March 22 to April 6, 2020[54]. While another study focusing on Twitter users in Massachusetts found increased anxiety expression after the enforcement of the Massachusetts State of Emergency and US State of Emergency[55]. Wang et al. found the sentiment toward the lockdown policy appears to be positive in most of the states (including California and Florida) of the USA, and negative among several states such as New York and Texas[56], which also demonstrated the geographic variation of public reactions to the lockdown policy.

This study has the following limitations. First, the evaluation of public mental health on social media may be biased because older people and people with lower income have less access to social media. What's more, the confirmed mental health problems should be diagnosed by professional psychologists. Data from social media platforms can only reflect the general situation among populations. In addition, tweets that contain keywords do not always reflect the users' mental health status, they can be comments on news or other people. To reduce this noise, we removed Tweets containing URLs in our preprocessing step, as these tweets were usually summarizations or quotes of other information sources. Third, keyword search identification suffers from the negation problem. For instance, the tweet containing "not panic" may be identified as an anxiety-related tweet. We manually check the extracted tweets and found the negation problem is rare (<10%), so this problem only affected our research to a limited extent.

The impact of the COVID-19 pandemic and the corresponding control measures on the public's mental status is dynamic and shows variability among different cohorts. In this study, we developed a full pipeline of tracking public mental status through social media during a pandemic from various aspects. This study demonstrated that social media such as Twitter can be a useful and timely platform for public health monitoring, and our study can be extended to track the mental health status of other important cohorts, such as sex minorities, adolescents, low-socioeconomic status groups.

Methods

Data collection

We collected and downloaded COVID-19 related Tweets from February 1, 2020, to September 30, 2021, using Twitter's Application Programming Interface based on the unique Tweet Identifier (Tweet ID) from an open-source database[57]. The downloaded data contained full tweet texts and the corresponding metadata including created times, user information, tweet status, etc. Only English and self-posted tweets (non-retweets) were kept, resulting in 368,816,761 tweets. The data collection process strictly followed Twitter's privacy and data use management. This study was conducted with approval by the Institutional Review Board of Zhejiang University and Mass General Brigham.

Data preprocessing and filtering

We removed tweets that contained URLs because such tweets often only included summaries or quotations of the original tweets. There were 133,002,625 tweets left after this step.

A mental health lexicon with 231 keywords was curated by a psychiatrist and a psychologist, it was categorized into four subgroups including anxiety, depression, insomnia, and addiction (**Supplementary Table 1**). This lexicon was used to extract the mental health and COVID-19 related tweets through keyword matching, 2,973,319 mental health-related tweets were kept after this step. **Fig. 1** shows an overview of the data preprocessing process.

Geographic information extraction

The user geographical information was collected from two fields of the tweet database: 1) the "place" field in tweet metadata, and 2) the "location" field nested in the "user" field, which is contained in each tweet's metadata. The "place" information was chosen as the primary evidence of the user's geographical information and the "location" information as the secondary evidence, as the former (from the Global Position System) is more accurate than the latter (user self-report). We built a list of state names of the USA and used it to extract the geographic information of users

(**Supplementary Methods**). Tweets whose user was identified in more than one state were removed.

Healthcare workers identification

To identify healthcare workers, we built a healthcare worker identification lexicon that can be divided into three groups: occupation, degree, and title of the association (**Supplementary Methods**). The lexicon contains 47 keywords, such as “doctor”, “MD”, “Doctor of Medicine” and “FACP”, etc. (**Supplementary Table 3**). We used this lexicon to filter the user’s description and extracted 40798 tweets from healthcare workers.

Topic model Analysis

The Latent Dirichlet Allocation (LDA) model[58] was used to conclude the main topics of mental health-related tweets. To create the corpora of topic modeling, we concatenated all tweets of interest and removed all stop words[59] as well as numbers and symbols. The topic model was implemented using the *LdaModel* function of the *Genism* package[59]. The number of topics was a hyperparameter, we selected it based on perplexity and topic coherence (detail information in **Supplementary Methods** and **Supplementary Fig. 3**).

Statistical analysis

We applied standard descriptive statistics to summarize four types of mental health-related tweets proportion, including median and interquartile ranges. Wilcoxon matched-pairs signed-ranks test was used to compare differences between healthcare workers and the general population because the differences between the two groups didn’t fit the normal distribution.

Interrupted time series analysis[37] was applied to analyze the lockdown policy effects on public mental health. We chose the proportion of mental health-related tweets as the outcome variable and selected three independent variables in the linear regression model: 1) time, a continuous variable encoding the day number of the research period (15 days before and after lockdown); 2) intervention, a binary variable, encoded as 0 before the lockdown policy, and 1 after the policy; and 3) their interaction term[60]. In regression models, the intercept and coefficient of "time" represent the baseline of the level and slope, respectively; the coefficient of "intervention" and interaction term represents the change of level and slope after the lockdown policy. We used the Durbin-Watson test to detect the autocorrelation of residuals. We applied Orthogonal Least Square (OLS) regression model when no autocorrelation existed, otherwise Generalized Least Square (GLS) model was employed. We used Python 3.6 to conduct the statistical analyses and chose a p-value of 0.05 to be the statistically significant threshold.

Data availability.

The data and code that support the findings of the study are available at https://github.com/zjumh/mental-health-during-COVID.

Acknowledgments

This study was partially supported by the Key Laboratory of Intelligent Preventive Medicine of Zhejiang Province (2020E10004).

Author Contributions

M.L. and J.Y. designed the study and drafted the manuscript. Y.H. collected the data, helped draft and revise the manuscript. M.L. performed data and statistical analysis. Y.L. and L.W. built the mental health keywords. Y.L., L.Z. and X.L. provided critical reviews. All authors reviewed the manuscript. M.L. takes responsibility for the integrity of the work.

Competing interests

The authors declare no competing interests.

References

1. Zhu, N. *et al.* A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* **382**, 727-733, doi:10.1056/NEJMoa2001017 (2020).
2. WHO. *WHO Coronavirus (COVID-19) Dashboard*, <https://covid19.who.int/> (2021).
3. Vigo, D. *et al.* Mental Health of Communities during the COVID-19 Pandemic. *Canadian Journal of Psychiatry-Revue Canadienne De Psychiatrie* **65**, 681-687, doi:10.1177/0706743720926676 (2020).
4. Moreno, C. *et al.* How mental health care should change as a consequence of the COVID-19 pandemic. *Lancet Psychiatry* **7**, 813-824, doi:10.1016/S2215-0366(20)30307-2 (2020).
5. Xin, M. *et al.* Negative cognitive and psychological correlates of mandatory quarantine during the initial COVID-19 outbreak in China. *Am Psychol* **75**, 607-617, doi:10.1037/amp0000692 (2020).
6. Wang, Y. *et al.* The impact of quarantine on mental health status among general population in China during the COVID-19 pandemic. *Mol Psychiatry* **26**, 4813-4822, doi:10.1038/s41380-021-01019-y (2021).
7. Witteveen, D. & Velthorst, E. Economic hardship and mental health complaints during COVID-19. *Proc Natl Acad Sci U S A* **117**, 27277-27284, doi:10.1073/pnas.2009609117 (2020).
8. Gualano, M. R., Lo Moro, G., Voglino, G., Bert, F. & Siliquini, R. Effects of Covid-19 Lockdown on Mental Health and Sleep Disturbances in Italy. *International journal of environmental research and public health* **17**, doi:10.3390/ijerph17134779 (2020).
9. Golberstein, E., Wen, H. & Miller, B. F. Coronavirus Disease 2019 (COVID-19) and Mental Health for Children and Adolescents. *JAMA Pediatr* **174**, 819-820, doi:10.1001/jamapediatrics.2020.1456 (2020).
10. Leeb, R. T. *et al.* Mental Health-Related Emergency Department Visits Among Children Aged <18 Years During the COVID-19 Pandemic - United States, January 1-October 17, 2020. *MMWR Morb Mortal Wkly Rep* **69**, 1675-1680, doi:10.15585/mmwr.mm6945a3 (2020).
11. Copeland, W. E. *et al.* Impact of COVID-19 Pandemic on College Student Mental Health and Wellness. *J Am Acad Child Adolesc Psychiatry* **60**, doi:10.1016/j.jaac.2020.08.466 (2021).
12. Wang, X. *et al.* Investigating Mental Health of US College Students During the COVID-19 Pandemic: Cross-Sectional Survey Study. *Journal of medical Internet research* **22**, e22817, doi:10.2196/22817 (2020).
13. Naidu, S. B. *et al.* The high mental health burden of "Long COVID" and its association with on-going physical and respiratory symptoms in all adults discharged from hospital. *Eur Respir J* **57**, doi:10.1183/13993003.04364-2020 (2021).
14. Lin, K. *et al.* The Mental Health Effects of COVID-19 on Health Care Providers in China. *Am J Psychiatry* **177**, 635-636, doi:10.1176/appi.ajp.2020.20040374 (2020).
15. Bryant-Genevier, J. *et al.* Symptoms of Depression, Anxiety, Post-Traumatic Stress Disorder, and Suicidal Ideation Among State, Tribal, Local, and Territorial Public Health Workers During the COVID-19 Pandemic - United States, March-April 2021. *MMWR Morb Mortal Wkly Rep* **70**, 1680-1685, doi:10.15585/mmwr.mm7048a6 (2021).
16. Rudberg, A.-S. *et al.* SARS-CoV-2 exposure, symptoms and seroprevalence in healthcare workers in Sweden. *Nat Commun* **11**, 5064, doi:10.1038/s41467-020-18848-0 (2020).

17 Manzano García, G. & Ayala Calvo, J. C. The threat of COVID-19 and its influence on nursing staff burnout. *J Adv Nurs* **77**, 832-844, doi:10.1111/jan.14642 (2021).

18 Rimmer, A. Covid-19: Drop the hero narrative and support doctors' mental health, says charity. *BMJ* **372**, n337, doi:10.1136/bmj.n337 (2021).

19 Rimmer, A. Covid-19: Two fifths of doctors say pandemic has worsened their mental health. *BMJ* **371**, m4148, doi:10.1136/bmj.m4148 (2020).

20 Pappa, S. *et al.* Prevalence of depression, anxiety, and insomnia among healthcare workers during the COVID-19 pandemic: A systematic review and meta-analysis. *Brain Behav Immun* **88**, 901-907, doi:10.1016/j.bbi.2020.05.026 (2020).

21 Firew, T. *et al.* Protecting the front line: a cross-sectional survey analysis of the occupational factors contributing to healthcare workers' infection and psychological distress during the COVID-19 pandemic in the USA. *BMJ Open* **10**, e042752, doi:10.1136/bmjopen-2020-042752 (2020).

22 Amin, F., Sharif, S., Saeed, R., Durrani, N. & Jilani, D. COVID-19 pandemic- knowledge, perception, anxiety and depression among frontline doctors of Pakistan. *BMC Psychiatry* **20**, 459, doi:10.1186/s12888-020-02864-x (2020).

23 Chew, N. W. S. *et al.* A multinational, multicentre study on the psychological outcomes and associated physical symptoms amongst healthcare workers during COVID-19 outbreak. *Brain Behav Immun* **88**, 559-565, doi:10.1016/j.bbi.2020.04.049 (2020).

24 Hay, S. I., George, D. B., Moyes, C. L. & Brownstein, J. S. Big data opportunities for global infectious disease surveillance. *PLoS Med* **10**, e1001413, doi:10.1371/journal.pmed.1001413 (2013).

25 Chew, C. & Eysenbach, G. Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. *PloS one* **5**, e14118, doi:10.1371/journal.pone.0014118 (2010).

26 Masri, S. *et al.* Use of Twitter data to improve Zika virus surveillance in the United States during the 2016 epidemic. *BMC Public Health* **19**, 761, doi:10.1186/s12889-019-7103-8 (2019).

27 van de Belt, T. H. *et al.* Social media posts and online search behaviour as early-warning system for MRSA outbreaks. *Antimicrob Resist Infect Control* **7**, 69, doi:10.1186/s13756-018-0359-4 (2018).

28 Zhang, C., Yu, M. C. & Marin, S. Exploring Public Sentiment on Enforced Remote Work During COVID-19. *Journal of Applied Psychology* **106**, 797-810, doi:10.1037/apl0000933 (2021).

29 Xie, Z. *et al.* Public Perception of COVID-19 Vaccines on Twitter in the United States. *medRxiv : the preprint server for health sciences*, doi:10.1101/2021.10.16.21265097 (2021).

30 Melton, C. A., Olusanya, O. A., Ammar, N. & Shaban-Nejad, A. Public sentiment analysis and topic modeling regarding COVID-19 vaccines on the Reddit social media platform: A call to action for strengthening vaccine confidence. *Journal of Infection and Public Health* **14**, 1505-1512, doi:10.1016/j.jiph.2021.08.010 (2021).

31 Ahmed, M. S., Aurpa, T. T. & Anwar, M. M. Detecting sentiment dynamics and clusters of Twitter users for trending topics in COVID-19 pandemic. *Plos One* **16**, doi:10.1371/journal.pone.0253300 (2021).

32 Xue, J. *et al.* Twitter Discussions and Emotions About the COVID-19 Pandemic: Machine Learning Approach. *Journal of Medical Internet Research* **22**, doi:10.2196/20550 (2020).

33 Xue, J. *et al.* Public discourse and sentiment during the COVID 19 pandemic: Using Latent Dirichlet Allocation for topic modeling on Twitter. *PloS one* **15**, e0239441, doi:10.1371/journal.pone.0239441 (2020).
34 Dyer, J. & Kolic, B. Public risk perception and emotion on Twitter during the Covid-19 pandemic. *Applied Network Science* **5**, doi:10.1007/s41109-020-00334-7 (2020).
35 Chandrasekaran, R., Mehta, V., Valkunde, T. & Moustakas, E. Topics, Trends, and Sentiments of Tweets About the COVID-19 Pandemic: Temporal Infoveillance Study. *Journal of medical Internet research* **22**, e22624, doi:10.2196/22624 (2020).
36 Abd-Alrazaq, A., Alhuwail, D., Househ, M., Hamdi, M. & Shah, Z. Top Concerns of Tweeters During the COVID-19 Pandemic: Infoveillance Study. *Journal of medical Internet research* **22**, e19016, doi:10.2196/19016 (2020).
37 Biglan, A., Ary, D. & Wagenaar, A. C. The Value of Interrupted Time-Series Experiments for Community Intervention Research. *Prevention science* **1**, 31-49, doi:10.1023/A:1010024016308 (2000).
38 McCracken, L. M., Badinlou, F., Buhrman, M. & Brocki, K. C. Psychological impact of COVID-19 in the Swedish population: Depression, anxiety, and insomnia and their associations to risk and vulnerability factors. *Eur Psychiatry* **63**, e81, doi:10.1192/j.eurpsy.2020.81 (2020).
39 Solomou, I. & Constantinidou, F. Prevalence and Predictors of Anxiety and Depression Symptoms during the COVID-19 Pandemic and Compliance with Precautionary Measures: Age and Sex Matter. *International journal of environmental research and public health* **17**, doi:10.3390/ijerph17144924 (2020).
40 Şahin, M. K., Aker, S., Şahin, G. & Karabekiroğlu, A. Prevalence of Depression, Anxiety, Distress and Insomnia and Related Factors in Healthcare Workers During COVID-19 Pandemic in Turkey. *J Community Health* **45**, 1168-1177, doi:10.1007/s10900-020-00921-w (2020).
41 Voitsidis, P. *et al.* Insomnia during the COVID-19 pandemic in a Greek population. *Psychiatry Res* **289**, 113076, doi:10.1016/j.psychres.2020.113076 (2020).
42 Wang, J. *et al.* Maternal insomnia during the COVID-19 pandemic: associations with depression and anxiety. *Soc Psychiatry Psychiatr Epidemiol* **56**, 1477-1485, doi:10.1007/s00127-021-02072-2 (2021).
43 Pizzonia, K. L. *et al.* Insomnia during the COVID-19 pandemic: the role of depression and COVID-19-related risk factors. *Cogn Behav Ther* **50**, 246-260, doi:10.1080/16506073.2021.1879241 (2021).
44 Kwon, S. & Park, A. Understanding user responses to the COVID-19 pandemic on Twitter from a terror management theory perspective: Cultural differences among the US, UK and India. *Computers in human behavior*, 107087, doi:10.1016/j.chb.2021.107087 (2021).
45 Li, S., Wang, Y., Xue, J., Zhao, N. & Zhu, T. The Impact of COVID-19 Epidemic Declaration on Psychological Consequences: A Study on Active Weibo Users. *International journal of environmental research and public health* **17**, doi:10.3390/ijerph17062032 (2020).
46 Singh, S. *et al.* Impact of COVID-19 and lockdown on mental health of children and adolescents: A narrative review with recommendations. *Psychiatry Res* **293**, 113429, doi:10.1016/j.psychres.2020.113429 (2020).

47 Shi, L. *et al.* Prevalence of and Risk Factors Associated With Mental Health Symptoms Among the General Population in China During the Coronavirus Disease 2019 Pandemic. *JAMA Netw Open* **3**, e2014053, doi:10.1001/jamanetworkopen.2020.14053 (2020).

48 Cénat, J. M. *et al.* Prevalence of symptoms of depression, anxiety, insomnia, posttraumatic stress disorder, and psychological distress among populations affected by the COVID-19 pandemic: A systematic review and meta-analysis. *Psychiatry Res* **295**, 113599, doi:10.1016/j.psychres.2020.113599 (2021).

49 Luo, M., Guo, L., Yu, M., Jiang, W. & Wang, H. The psychological and mental impact of coronavirus disease 2019 (COVID-19) on medical staff and general public - A systematic review and meta-analysis. *Psychiatry Res* **291**, 113190, doi:10.1016/j.psychres.2020.113190 (2020).

50 Huang, Y. & Zhao, N. Generalized anxiety disorder, depressive symptoms and sleep quality during COVID-19 outbreak in China: a web-based cross-sectional survey. *Psychiatry Res* **288**, 112954, doi:10.1016/j.psychres.2020.112954 (2020).

51 Li, Y. *et al.* Insomnia symptoms during the early and late stages of the COVID-19 pandemic in China: a systematic review and meta-analysis. *Sleep Med*, doi:10.1016/j.sleep.2021.09.014 (2021).

52 Jahrami, H. *et al.* Sleep problems during the COVID-19 pandemic by population: a systematic review and meta-analysis. *J Clin Sleep Med* **17**, 299-313, doi:10.5664/jcsm.8930 (2021).

53 Wu, T. *et al.* Prevalence of mental health problems during the COVID-19 pandemic: A systematic review and meta-analysis. *J Affect Disord* **281**, 91-98, doi:10.1016/j.jad.2020.11.117 (2021).

54 Mittal, R., Ahmed, W., Mittal, A. & Aggarwal, I. Twitter users exhibited coping behaviours during the COVID-19 lockdown: an analysis of tweets using mixed methods. *Information Discovery and Delivery*, doi:10.1108/idd-08-2020-0102.

55 Thorpe Huerta, D., Hawkins, J. B., Brownstein, J. S. & Hswen, Y. Exploring discussions of health and risk and public sentiment in Massachusetts during COVID-19 pandemic mandate implementation: A Twitter analysis. *SSM - population health* **15**, 100851, doi:10.1016/j.ssmph.2021.100851 (2021).

56 Wang, J. *et al.* Global evidence of expressed sentiment alterations during the COVID-19 pandemic. *Nat Hum Behav* **6**, 349-358, doi:10.1038/s41562-022-01312-y (2022).

57 Chen, E., Lerman, K. & Ferrara, E. Tracking Social Media Discourse About the COVID-19 Pandemic: Development of a Public Coronavirus Twitter Data Set. *JMIR public health and surveillance* **6**, e19273, doi:10.2196/19273 (2020).

58 Blei, D. M., Ng, A. Y. & Jordan, M. I. Latent Dirichlet allocation. *Journal of Machine Learning Research* **3**, 993-1022, doi:10.1162/jmlr.2003.3.4-5.993 (2003).

59 Rehurek, R. & Sojka, P. in *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks.* 45-50 (ELRA).

60 Bernal, J. L., Cummins, S. & Gasparrini, A. Interrupted time series regression for the evaluation of public health interventions: A tutorial. *International journal of epidemiology* **46**, 348-355, doi:10.1093/ije/dyw098 (2017).

Figure Legend

- **Fig. 1 Data collection and preprocessing**
- **Fig. 2 The trend of four types of mental health symptoms-related tweets by a.** number; **b.** proportion.
- **Fig. 3 Proportion distribution of mental health-related tweets in the USA. a.** Geographical distributions for the tweets from February 1st, 2020 to September 30th, 2021. **b.** The monthly proportion of mental health-related tweets in the 50 states of the USA. (The proportion was calculated based on the number of mental health and COVID-19 related tweets divided by the number of COVID-19 related tweets.)
- **Fig. 4 Dynamic characteristics of topic proportions** (Topic proportion equals to the tweets number in each topic divided by total tweets number of this month)
- **Fig. 1 The distribution of tweets in topics for healthcare workers and the general population. a.** Average number of tweets per user in each topic. **b.** Logarithmic ratio of the average number of tweets between healthcare workers and general population in each topic. (The ratio equals to the average number of tweets per user among healthcare workers divided by the average number among the general population.)
- **Fig. 6 Daily proportion of mental health-related tweets before and after the lockdown policy**. (0 in the X-axis indicates the effective date of the lockdown policy. The model included tweets from 15-days before and after the effective date.)

Table Legend

- **Table 1. Comparison of proportions of mental health-related tweets between healthcare workers and the general population**
- **Table 2. The impact of lockdown policy on public mental health**

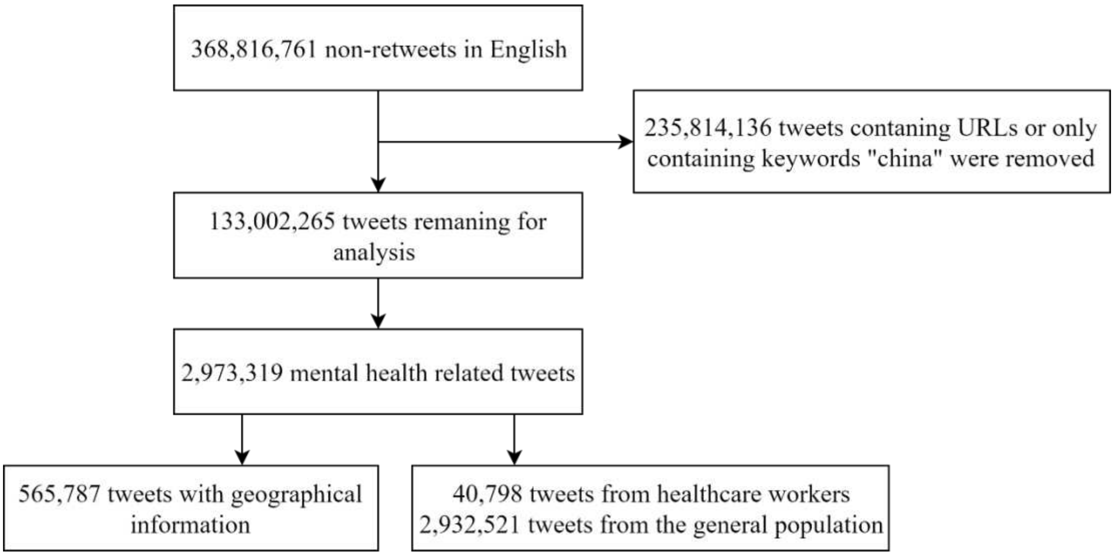


Fig. 2 Data collection and preprocessing

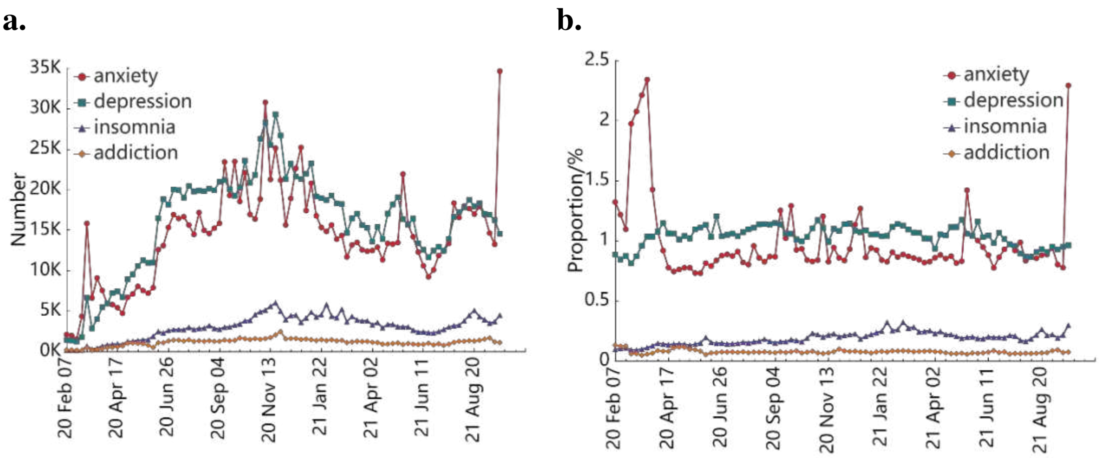
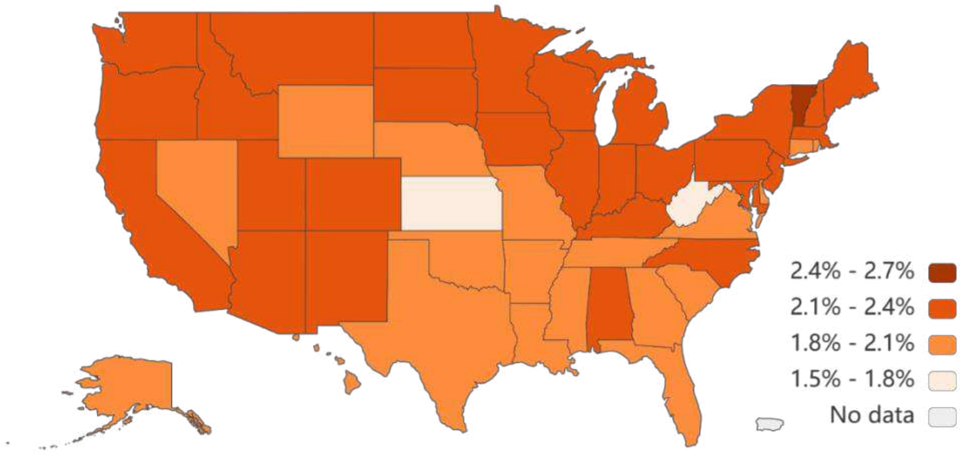


Fig. 3 The trend of four types of mental health symptoms-related tweets by a. number; **b**. proportion.

a.



b.

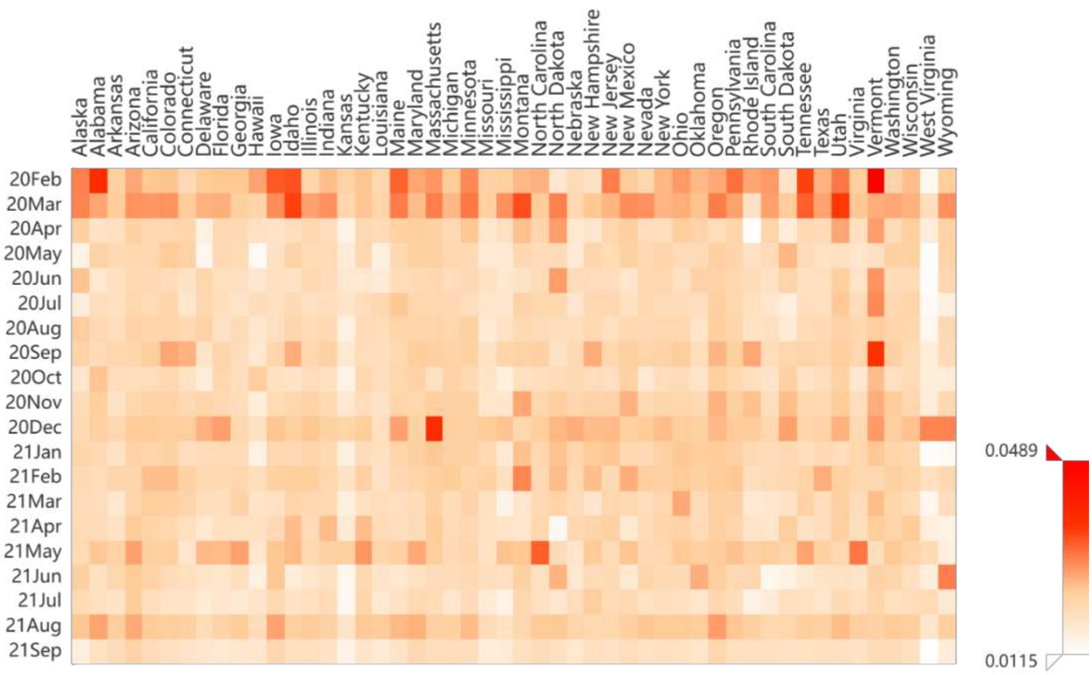


Fig. 4 Proportion distribution of mental health-related tweets in the USA. a. Geographical distributions for the tweets from February 1st, 2020 to September 30th, 2021. **b.** The monthly proportion of mental health-related tweets in the 50 states of the USA. (The proportion was calculated based on the number of mental health and COVID-19 related tweets divided by the number of COVID-19 related tweets.)

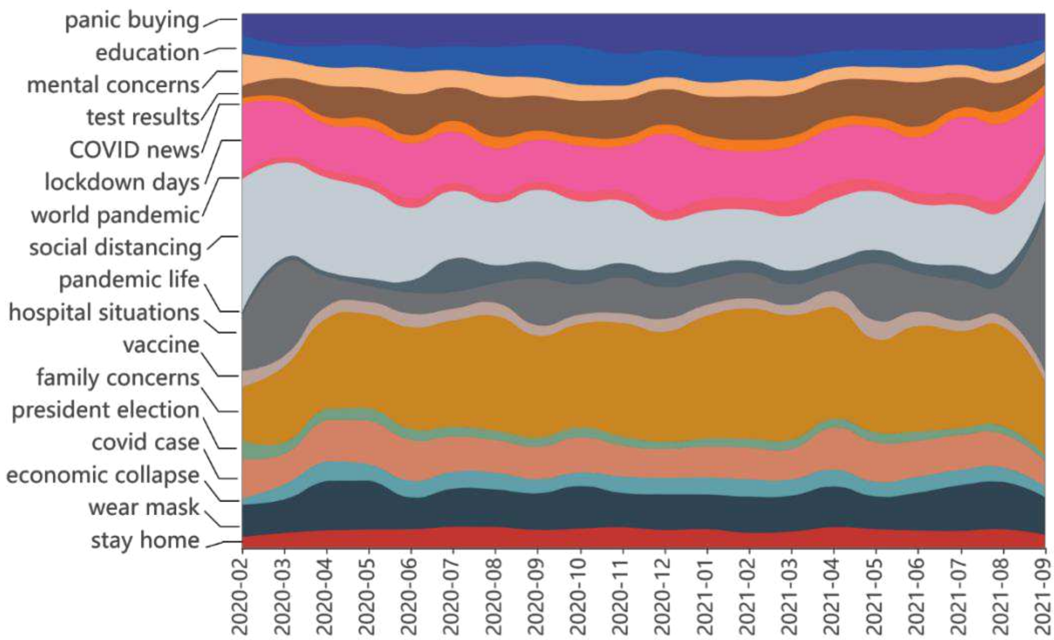
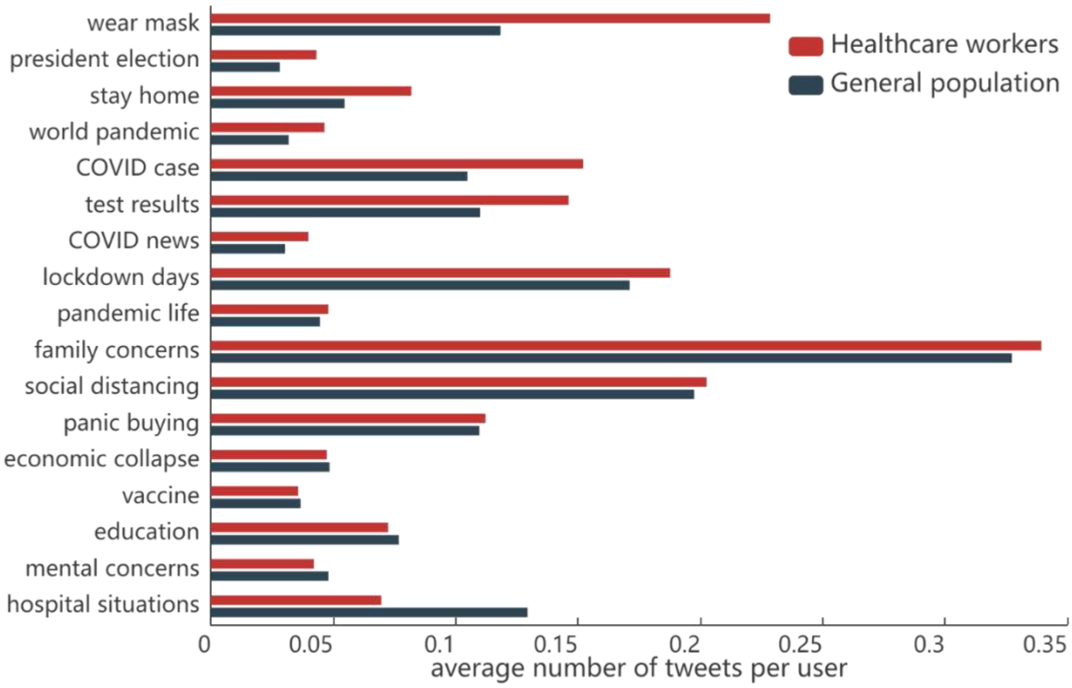


Fig. 5 Dynamic characteristics of topic proportions (Topic proportion equals to the tweets number in each topic divided by total tweets number of this month)

a.



b.

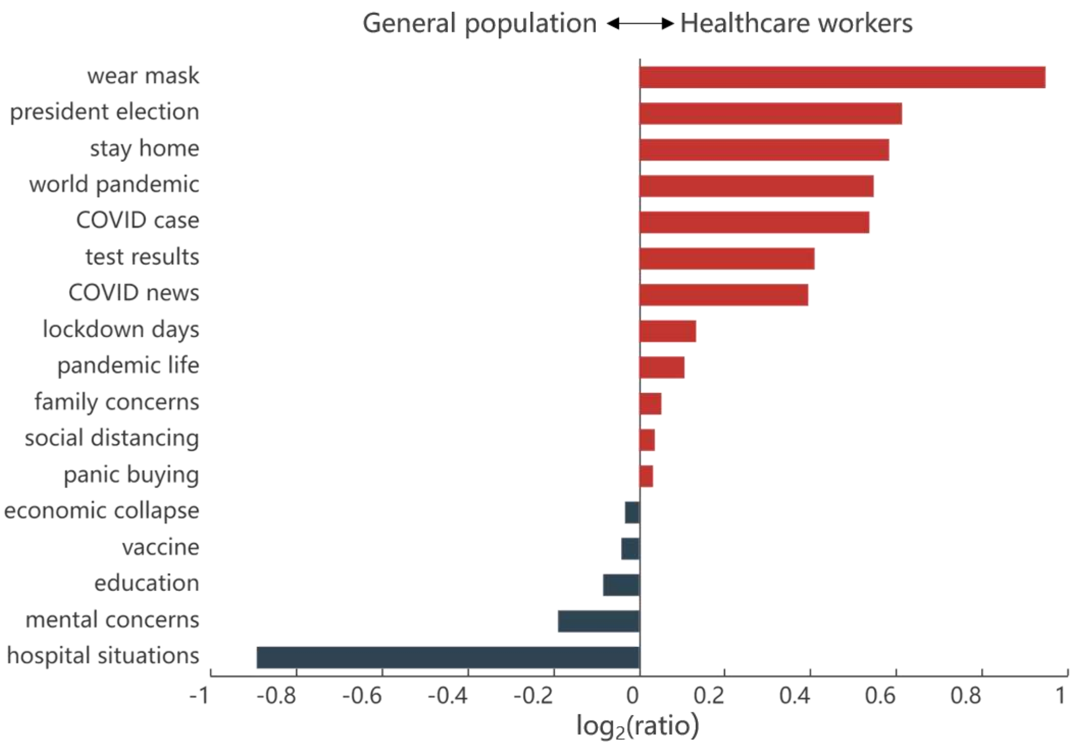


Fig. 6 The distribution of tweets in topics for healthcare workers and the general population. a. Average number of tweets per user in each topic. **b.** Logarithmic ratio of the average number of tweets between healthcare workers and general population in each topic. (The ratio equals to the average number of tweets per user among healthcare workers divided by the average number among the general population.)

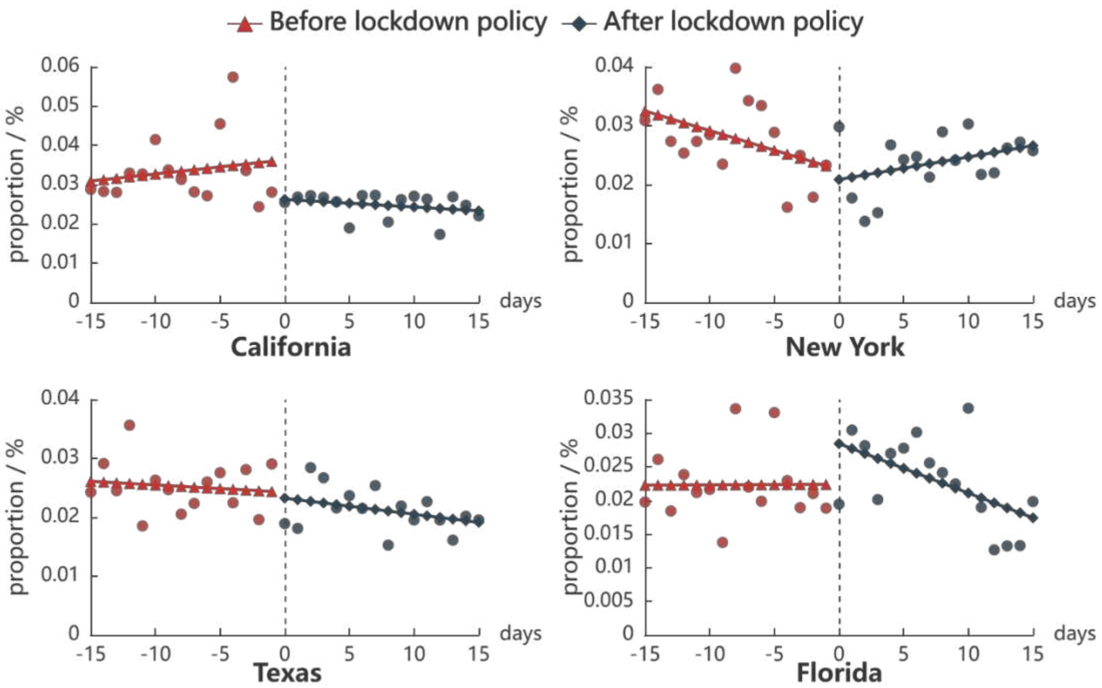


Fig. 7 Daily proportion of mental health-related tweets before and after the lockdown policy. (0 in the X-axis indicates the effective date of the lockdown policy. The model included tweets from 15-days before and after the effective date.)

Table 1. Comparison of proportions of mental health-related tweets between healthcare workers and the general population

	Healthcare workers Median (IQR)	General population Median (IQR)	W	P
Anxiety	0.892(0.183)	0.872(0.133)	1865	0.836
Depression	1.242(0.170)	1.049(0.112)	37	<0.001
Insomnia	0.230(0.126)	0.192(0.0685)	282	<0.001
Addiction	0.111(0.040)	0.074(0.014)	61	<0.001

(IQR: Interquartile Range, and Wilcoxon matched-pairs signed-ranks test was applied to compare the differences between two groups.)

Table 2. The impact of lockdown policy on public mental health

	Date	**Intercept**	**Time**	**Intervention**	**Time* Intervention**	**F statistic**
California	Mar 19, 2020	0.0306*	0.0004	-0.0012	-0.0006	4.911*
New York	Mar 22, 2020	0.0332*	-0.0007*	-0.0184*	0.0010	3.548*
Texas	Apr 2, 2020	0.0262*	-0.0001	0.0014	-0.0002	3.251*
Florida	Apr 3, 2020	0.0223*	8.425e-6	0.0179*	-0.0007	1.989

(The meaning of the independent variables: 1) time: a continuous variable encoding the number of days in the research period (15 days before and after lockdown), 2) intervention: a binary variable, encoded as 0 before the lockdown policy and 1 after the policy, 3) time*intervention, the interaction term of time and intervention. The coefficients of time in the model of California and Florida were not reported because their absolute values were less than 0.0001. Statistical significance: P<0.05*.)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInformationforXXXTrackingtheimpactofCOVID19andlockdownpolicyonpublicmentalhealthusingsocialmediaXXX.pdf