

Depression and Anxiety on Twitter During the COVID-19 Stay-At-Home Period in 7 Major U.S. Cities



Danielle Levanti¹, Rebecca N. Monastero, MD,² Mohammadzaman Zamani, PhD,³ Johannes C. Eichstaedt, PhD,⁴ Salvatore Giorgi, MA,⁵ H. Andrew Schwartz, PhD,³ Jaymie R. Meliker, PhD⁶

Introduction: Although surveys are a well-established instrument to capture the population prevalence of mental health at a moment in time, public Twitter is a continuously available data source that can provide a broader window into population mental health. We characterized the relationship between COVID-19 case counts, stay-at-home orders because of COVID-19, and anxiety and depression in 7 major U.S. cities utilizing Twitter data.

Methods: We collected 18 million Tweets from January to September 2019 (baseline) and 2020 from 7 U.S. cities with large populations and varied COVID-19 response protocols: Atlanta, Chicago, Houston, Los Angeles, Miami, New York, and Phoenix. We applied machine learning-based language prediction models for depression and anxiety validated in previous work with Twitter data. As an alternative public big data source, we explored Google Trends data using search query frequencies. A qualitative evaluation of trends is presented.

Results: Twitter depression and anxiety scores were consistently elevated above their 2019 baselines across all the 7 locations. Twitter depression scores increased during the early phase of the pandemic, with a peak in early summer and a subsequent decline in late summer. The pattern of depression trends was aligned with national COVID-19 case trends rather than with trends in individual states. Anxiety was consistently and steadily elevated throughout the pandemic. Google search trends data showed noisy and inconsistent results.

Conclusions: Our study shows the feasibility of using Twitter to capture trends of depression and anxiety during the COVID-19 public health crisis and suggests that social media data can supplement survey data to monitor long-term mental health trends.

AJPM Focus 2023;2(1):100062. © 2022 The Author(s). Published by Elsevier Inc. on behalf of The American Journal of Preventive Medicine Board of Governors. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

From the ¹Undergraduate Studies, Cornell University, Ithaca, New York; ²Renaissance School of Medicine, Stony Brook University, Stony Brook, New York; ³Department of Computer Science, College of Engineering and Applied Sciences, Stony Brook University, Stony Brook, New York; ⁴Department of Psychology, School of Humanities and Sciences, Stanford University, Palo Alto, California; ⁵Department of Computer and Information Science, University of Pennsylvania, Philadelphia, Pennsylvania; and ⁶Program in Public Health, Department of Family, Population, & Preventive Medicine, Renaissance School of Medicine, Stony Brook

University, Stony Brook, New York

Address correspondence to: Jaymie R. Meliker, PhD, Program in Public Health, Department of Family, Population, & Preventive Medicine, Renaissance School of Medicine, Stony Brook University, HSC L3, Room 071, Stony Brook NY 11790-8338. E-mail: jaymie.meliker@stonybrook.edu.

2773-0654/\$36.00

<https://doi.org/10.1016/j.focus.2022.100062>

INTRODUCTION

After the advent of coronavirus disease (COVID-19) in the U.S. in early 2020, the U.S. implemented multiple measures in an attempt to slow down the spread of this highly transmissible virus, including social distancing, mask wearing, and closure of nonessential businesses and schools. Between March 1, 2020 and May 31, 2020, 42 states and territories issued stay-at-home orders.¹ Although the duration and enforcement of these orders varied, they shared common quarantine-like mandates of limiting population movement for essential tasks to lower COVID-19 transmission, resulting in an unprecedented and widespread shift in the daily lives of most U.S. residents to be highly isolated.²

Social isolation during the stay-at-home period presented a particular threat to well-being. A wide body of work has established that social relationships are among the leading determinants of subjective well-being.³ We are a fundamentally social species; the physical presence of others regulates our emotions, including those occurring in response to threat.⁴ By contrast, loneliness, the result of insufficient quality and quantity of social relationships, has detrimental impacts on mental health⁵ as well as physical health and lifespan in general.⁶ For example, among older adults, loneliness and isolation may be predictors for symptoms of depression and anxiety.^{7,8}

In addition to social isolation, the experience of the pandemic has been marked by emotions associated with sympathetic arousals, such as fear, worry, and agitation, with their behavioral manifestations such as rumination and hypervigilance.⁹ Sources of fear include contracting the virus, losing loved ones, lack of resources, and economic struggles as well as uncertainty about the future. Both negative emotions and social isolation during the pandemic may contribute to poor mental health outcomes.

Previous work has identified negative mental health outcomes in earlier epidemics requiring quarantine, including severe acute respiratory syndrome, Middle East respiratory syndrome, and Ebola.¹⁰ In a study of healthcare workers with potential severe acute respiratory syndrome exposure, quarantined individuals showed increased symptoms of acute stress disorder, anxiety under specified conditions, irritability, insomnia, poor concentration, and decreased quality of their work compared with those who were not quarantined.¹¹ Another study examined post-traumatic stress symptoms in adults and children in quarantine and found that those in quarantine faced significantly higher post-traumatic stress scores than those who had not been quarantined.¹²

Several scholars predicted negative mental health outcomes during the COVID-19 pandemic and stay-at-

home periods because of loneliness and isolation and concluded that increased mental health support would be particularly necessary to prevent increased suicides, depression, and anxiety.^{13,14} Indeed, recent studies have noted the increased prevalence of mental health struggles in the U.S. during COVID-19 through survey data. For example, 1 survey of U.S. adults in late June 2020 showed that 40% of respondents were struggling with mental health or substance abuse issues during the pandemic; 13% of respondents reported anxiety or depression symptoms.¹⁵ An international study during the pandemic that collected survey data from 78 countries showed that poor mental health was prevalent among 10% of survey respondents and that moderate mental health was prevalent among 50%.¹⁶ A recent review of 19 mental health-related studies during the pandemic found a prevalence of depressive symptoms of 15%–48% compared with a prepandemic prevalence of 4%–7%.¹⁷ In the same review, the prevalence of anxiety during the pandemic was identified as ranging between 6% and 51%.

There are many strengths to these surveys; however, they are limited in (1) time (moments or weeks measured), (2) sample (number of residents represented), and (3) scope (narrowly limited to anxiety and depression outside of the rapidly changing everyday concerns during COVID-19 and stay-at-home onset). Publicly available Twitter data have been shown to provide an alternative window into population mental health insights and how those trends change over time.^{18,19} Although imperfect samples, previous work has shown high convergence between Twitter-based psychological assessments and representative mental and physical health outcomes.^{20–22} In addition, they provide a more readily available source of data should we need to identify social and psychological trends in a future pandemic.

In this study, we seek to characterize the relationship between stay-at-home orders because of COVID-19 and anxiety and depression in 7 major U.S. cities utilizing Twitter data. We selected 7 U.S. cities located around the country to examine the role of local case trends and stay-at-home orders. We compare monthly data on anxiety and depression from January to September 2020 with data from the same period in 2019.

METHODS

COVID-19 Data

COVID-19 case counts in the states of Arizona, California, Florida, Georgia, Illinois, New York, and Texas were captured from the *New York Times* COVID-19 website: <https://www.nytimes.com/interactive/2021/us/covid-cases.html>. Dates of statewide stay-at-home orders in 2020 were recorded from sources listed in

[Appendix Table](#) (available online) and cross-referenced with articles in local newspapers in each state and state-issued executive orders.

Twitter Data

We compiled monthly Twitter-based data between January and September of 2019 and 2020 in 7 U.S. cities: Atlanta, GA; Chicago, IL; Houston, TX; Los Angeles, CA; Miami, FL; New York City, NY; and Phoenix, AZ. These cities were selected for analysis owing to their large populations and varied COVID-19 response protocols and stay-at-home orders. To build the Twitter data set, we began with a random sample of 1% of all publicly available tweets from January to March 2020. The tweets in this random sample were mapped to U.S. cities on the basis of free-response location fields and geographic coordinates. This approach was found to agree with human judgments of the intended city 94% of the time.²⁰ Next, we took all Twitter users mapped to the 7 cities mentioned earlier and collected each user's tweet history dating back to January 1, 2019. We then restricted the tweet sample to between January 1 and September 30 of both 2019 and 2020. Following thresholds set by the County Tweet Lexical Bank,²³ a large open-source database of U.S. location-mapped Twitter data, each user posted at least 30 tweets, and we randomly sampled 10,000 users per city. This resulted in a total of 56,411,200 tweets from 70,000 Twitter users. Full statistics are reported in [Appendix Table 2](#) (available online), with example tweets in [Appendix Table 3](#) (available online).

Artificial intelligence-based language assessment. The text of the tweet was input into 2 artificial intelligence (AI)-based assessments of depression and anxiety developed by Schwartz et al.^{24,25} These models required 3 types of linguistic features: (1) relative frequencies of words and phrases, (2) binary indicators of words and phrases, and (3) topic prevalence scores. Words and phrases are sequences of 1 to 3 words in a row. Their relative frequency was recorded by counting each word or phrase mentioned and dividing by the total number of words or phrases mentioned by the Twitter user. The binary indicator for words and phrases simply indicated whether each word or phrase is present (1) or not (0).

Analytic strategy. To apply the approach of Schwartz et al.,²⁴ all word, phrase, and topic features were extracted for each Twitter user for each month in the data set. Once extracted, all features were averaged across all Twitter users within each of the 7 cities, resulting in city-level word, phrase, and topic scores for each month. We then applied the depression and anxiety AI assessments to each of the city-month feature sets and then calculated the difference between the 2019 and 2020 scores. We then standardized the monthly scores for both depression and anxiety for each of the 7 cities by subtracting the mean and dividing by the SD across all city-month data points across 2019 (baseline) and 2020 (also known as a z-score). All analyses, described earlier, were performed using the open-source python package DLATK (Differential Language Analysis ToolKit).²⁵

Google Trends Data

As a secondary analysis, we also compiled Google Trends deidentified data at the city level. This is a more simplistic method than the AI-based assessment we used for Twitter. This method relies on weekly internet search term volumes, which represent the

entire number of searches standardized on a 0–100 scale: 0 indicates no searches in a given time frame for a given term, and 100 represents the maximum number of searches in a given time frame for a given term. Search data were collected for the same 7 U.S. cities. Location at the time of search was determined by Google through user location data, including Internet Protocol address and GPS data, among others.

Search terms were generated on the basis of terms used in previous work evaluating mental health through Internet search trends and were tailored to specifically identify patients with mental health issues.^{26,27} *Keywords with insufficient data*, defined as a search volume of 0 for any month for 2 or more cities for a given term, were excluded from our analysis. Search terms included the following mental health-related terms or phrases: *Anxiety*, *Signs of anxiety*, *Symptoms of anxiety*, *Depression*, *Signs of depression*, *Symptoms of depression*, and *Panic attack*. Trends for anxiety and depression are presented in the [Appendix Figures](#) (available online); the other search terms yielded similarly ambiguous results and are not shown. We present relative search volume graphs at the monthly level per city. We calculated the monthly difference in relative search volume scores for each of the 7 cities, comparing the average from 2020 with the city-month average from 2018 and 2019. Owing to minor changes in data when accessed on different days with the same specifications, for consistency, all data were gathered on February 1, 2021.²⁸

This work was approved as exempt, nonhuman subjects by an academic IRB. This approval was granted because it utilized publicly available and aggregate data from Google Trends and Twitter.

RESULTS

A timeline of U.S. COVID-19 waves of cases in the 7 states and statewide stay-at-home orders is shown in [Figure 1](#). In [Appendix Figure 1](#) (available online), we see the locations of the 7 cities on a U.S. map, and in [Appendix Figure 2](#) (available online), we see that cases began to sharply increase in New York in March and in Illinois in April 2020 but did not sharply increase until June and July 2020 in the other locations: Georgia, Texas, California, Arizona, and Florida. Stay-at-home orders were adopted in the 7 states in the spring in response to the national spread of COVID-19 but were removed by early May in Arizona, Texas, Florida, and Georgia before the increased number of cases in some of the states in June.

Depression

Twitter-based depression scores resulted in trends of increased depression during the early phase of the pandemic, with a peak in early summer and a subsequent decline in late summer in each of the 7 cities, although Phoenix showed increasing depression rates in September 2020 ([Figure 2](#)). Depression trends appeared to be more in sync with the national prevalence of COVID-19 case trends ([Appendix Figure 2](#), available online) than with local trends of cases or stay-at-home orders in individual states.

COVID-19 Timeline

January 2020	First COVID-19 case reported in US
March 2020	Initial wave of cases in New York
March 19 – April 3, 2020	Stay at Home order initiated in Arizona, California, Florida, Georgia, Illinois, New York, Texas
April 2020	Initial wave of cases in Illinois
Early May 2020	Stay at Home order begins to be lifted in Arizona, Florida, Georgia, Texas
June 2020	Initial wave of cases in California, Texas, Florida, Georgia, Arizona Stay at Home order begins to be lifted in California, Illinois, New York

Figure 1. Timeline of major COVID-19–related events in 2020, as relevant to the 7 cities under study.

Anxiety

Anxiety levels based on Twitter data were overall elevated when compared with those in the previous year, most consistently elevated beginning in April and continuing into the Autumn, with Phoenix again showing an increase over all previous values in September (Figure 3). Most cities saw an initial peak in anxiety in April, coincident with national COVID-19 case trends, and remained elevated throughout the summer months, regardless of local case trends or stay-at-home orders.

Secondary Analysis: Google Search Terms

In the secondary analysis using relative search volume of Google search terms, no strong patterns were observed for depression- or anxiety-related searchers across cities or over time (Appendix Figures 3 and 4, available online). All cities and search terms that showed considerable increases also showed considerable decreases at other times of the year, although with inconsistent patterns.

DISCUSSION

This study employed a qualitative comparison of mental health trends over the pandemic on the basis of Twitter posts and Google search queries. Our study found overall increases in depression and anxiety expressed on Twitter throughout the early pandemic (until September 2020). The Twitter analysis was consistent with our hypothesis of increased depression and anxiety in the early stages of the pandemic, which is also consistent with survey data during the same period.^{15–17} According to Twitter data, depression was higher earlier and resolved toward the end of the summer, whereas anxiety increased in April and stayed elevated into the autumn. We had hypothesized that these trends would vary by city on the basis of local COVID-19 case trends or stay-at-home orders, but this was not the case; the trends were largely consistent across the cities, reflecting the national milieu and not local factors.

Our secondary analysis relied on a simpler approach to evaluating patterns of search terms in Google Trends data. These data did not show trends similar to those of the Twitter data; when compared with the 2019 baseline period, the difference from 2020 scores appeared noisy with frequent peaks and troughs, and not showing a clear association with COVID-19 trends. The figures appear to approximate fluctuations from normal variation. Recent studies used Google Trends data to characterize mental health–related outcomes during the pandemic similar to what we have done in this study. Results have been mixed, with some showing Google Trends data matching the anticipated increases in anxiety or depression²⁹ and others not reporting increases in depression or anxiety.^{30–32} We also note a recent report indicating that Google Trends data do not appear to be a useful indicator of changing the levels of population mental health during a public health emergency.³³ Previous work also found that some approaches to using Google Trends for forecasting suicide rates were not very accurate.²⁸ By contrast, Twitter data seem to aid in predicting infectious diseases such as influenza³⁴ and more reliably capture mental health trends and be more aligned with survey data during the early phase of the pandemic.^{35,36} We think this is because many on Twitter share their daily thoughts, feelings, and behaviors,¹⁹ whereas Google Trends capture what populations are searching for.³⁷ Conceivably, Google Trends data may be suitable for other use cases, such as viral prevalence surveillance as people search for the particular symptoms they are encountering. For example, some previous work has shown that people mention more influenza symptoms in areas at times with higher rates of influenza.³⁴

We did not observe strong influences of state trends in case rates or stay-at-home orders. By and large, the trends picked up by Twitter appear to follow the nationwide COVID-19 case rates, for example, with increases in depression from April through July in most cities

even before they experienced local increases in cases and after local stay-at-home orders ended. Mortality tended to follow cases on a 2-week delay in 2020 (before vaccinations and better treatments), with mortality peaks generally reflecting the case peaks (Appendix Figure 2, available online); therefore, local mortality trends also were not correlated with local changes in depression or anxiety. This may point to the fact that national (rather than local) reporting of pandemic developments may offer the most salient psychological context that the mental health rates reflect.

Limitations

A limitation of the study is the observational nature of our work as well as the inability to distinguish between the impact of policy and the impact of the pandemic. We do not have the ability to control for individual-level confounders, and we lack knowledge of the demographics of participants. Race, ethnicity, SES, age, and a multitude of other factors may contribute to individuals’ responses to the pandemic. We also acknowledge that the protests and social unrest in June 2020 could influence feelings of depression and/or anxiety, and that is not considered in

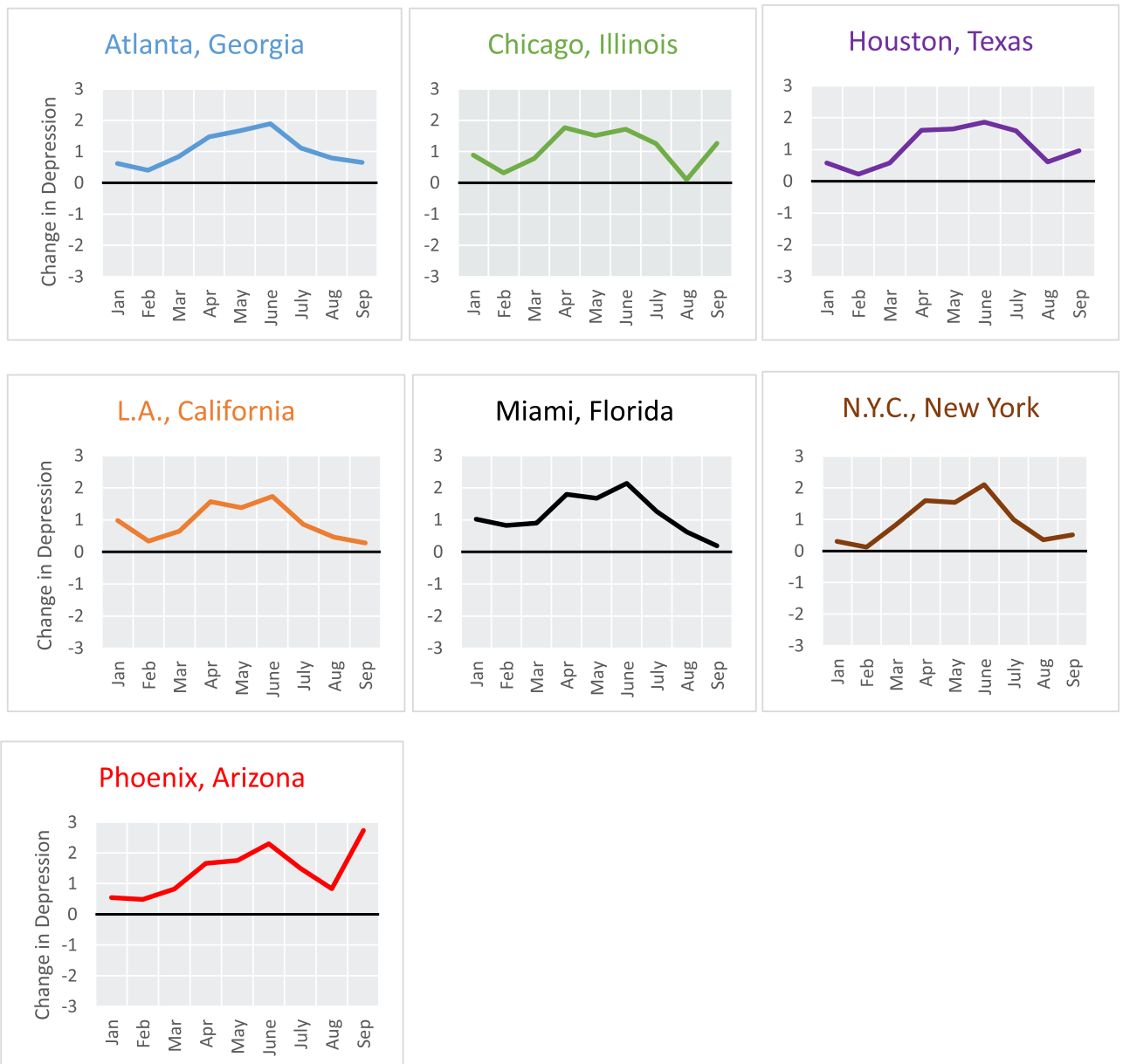


Figure 2. Twitter trends for change in depression, the difference between 2020 and 2019, by city and month, January–September
 Note: Units of Twitter scores reflect an SD change between 2020 and 2019. A change equal to 1 reflects a 1 SD change in depression score.
 Apr, April; Aug, August; Feb, February; Jan, January; L.A., Los Angeles; Mar, March; N.Y.C., New York City; Sep, September.

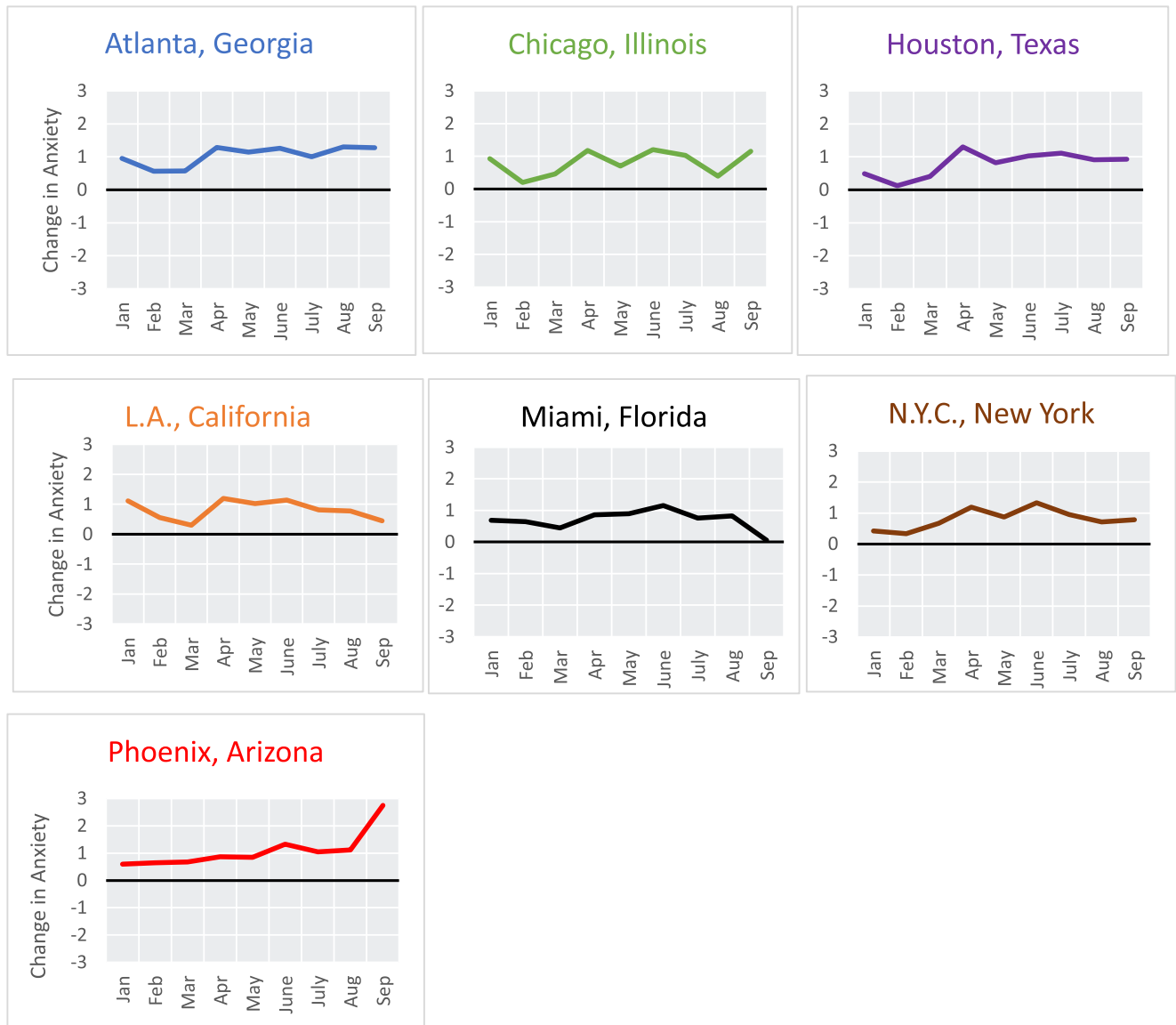


Figure 3. Twitter trends for change in anxiety, the difference between 2020 and 2019, by city and month, January–September
 Note: Units of Twitter scores reflect an SD change between 2020 and 2019. A change equal to 1 reflects a 1 SD change in anxiety score.
 Apr, April; Aug, August; Feb, February; Jan, January; L.A., Los Angeles; Mar, March; N.Y.C., New York City; Sep, September.

this study. We also cannot distinguish between mental health issues because of pandemic stressors (e.g., fear of illness, fear of losing loved ones, and so on) and mental health issues resulting from isolation and social distancing policies. It is likely that both policy and pandemic have contributed to overall mental health, and our work is not equipped to distinguish the two, if separable.

Although Twitter is an important source of big data, it also has limitations in utility. One possible problem with Twitter, Google Trends, or other social media platforms is the issue of ambiguities in language use. One limitation is that words can carry multiple meanings such that an individual writing *great* might not be referring to

how they feel but rather using it as an adjective as in a great sadness. The AI-based approach we use partially solves this by not assuming relationships between words and depressivity but rather learning it from examples.^{24,38} Furthermore, social media are known to skew young and thus introduce a selection bias. On average, most Twitter users in this sample have posted 700–900 tweets (Appendix Table 2, available online), which is roughly 350–450 tweets per year. In general, social media–based language assessments rely on users posting a minimum number of words (e.g., 500–1,000 words) to derive stable estimates. Thus, these findings may not generalize to populations who are not as active on

Twitter. Finally, there may be errors in the geolocation process. For example, we assign a location to each Twitter user's timeline on the basis of a single tweet. In addition, we assign a static location and do not consider that tweets may originate from different locations (e.g., traveling) or the fact that these users may have moved. There may also be a selection bias because Twitter users who do not turn on location data or self-report their location in their profile will not be in the sample. Taken together, these issues (errors in the geolocation process, errors in the mental health estimates, selection biases, and high minimum words and tweet thresholds) should only make the process of estimating community attributes more difficult. Nevertheless, past work has shown on average that community lexical variance does match accepted community measurements: people do tweet more influenza symptoms in areas at times with higher rates of influenza,³⁴ positive emotion words are mentioned more in areas surveying higher life satisfaction,^{20,22} and individuals with medical diagnoses of depression are more likely to talk about depressive symptoms (e.g., low mood, somatic pain, hostility, and loneliness) on social media.³

CONCLUSIONS

Our study observes Twitter trends over a prolonged timeframe during the COVID-19 pandemic to shed insights on possible mental health outcomes of a public health crisis. These data are not meant to replace survey data but can supplement questionnaires to help identify the long-term psychological impacts of social distancing during pandemics. Policy makers and other officials should plan for such events and consider both physical and psychological aspects of human health when deciding on effective public planning. With increased information about the pandemic's effect on mental health, either owing to isolation or owing to other concerns surrounding the pandemic, we will be better equipped to provide resources and mitigate these public health concerns in the future.

ACKNOWLEDGMENTS

Declarations of interest: none.

CREDIT AUTHOR STATEMENT

Danielle Levanti: Conceptualization, Data curation, Formal analysis, Writing – original draft, Writing – review & editing. Rebecca N. Monastero: Conceptualization, Data curation, Project administration, Supervision, Writing – original draft, Writing – review & editing. Mohammadzaman Zamani: Data curation, Formal analysis, Software, Visualization, Writing – review &

editing. Johannes C. Eichstaedt: Data curation, Formal analysis, Software, Visualization, Writing – review & editing. Salvatore Giorgi: Data curation, Formal analysis, Software, Visualization, Writing – review & editing. H. Andrew Schwartz: Conceptualization, Data curation, Formal analysis, Methodology, Project administration, Supervision, Visualization, Writing – original draft, Writing – review & editing. Jaymie R. Meliker: Conceptualization, Methodology, Formal analysis, Project administration, Supervision, Visualization, Writing – original draft, Writing – review & editing.

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found in the online version at [doi:10.1016/j.focus.2022.100062](https://doi.org/10.1016/j.focus.2022.100062).

REFERENCES

- Moreland A, Herlihy C, Tynan MA, et al. Timing of state and territorial COVID-19 stay-at-home orders and changes in population movement — United States, March 1–May 31, 2020. *MMWR Morb Mortal Wkly Rep*. 2020;69(35):1198–1203. <https://doi.org/10.15585/mmwr.mm6935a2>.
- Mervosh S, Lu D, Swales V. See which states and cities have told residents to stay at home. *The New York Times*. April 20, 2020 <https://www.nytimes.com/interactive/2020/us/coronavirus-stay-at-home-order.html>.
- Eichstaedt JC, Yaden DB, Ribeiro F, Adler A, Kern ML. Lifestyle and wellbeing: exploring behavioral and demographic covariates in a large U.S. sample. *Intrl J Wellbeing*. 2020;10(4):87–112. <https://doi.org/10.5502/ijw.v10i4.831>.
- Coan JA, Schaefer HS, Davidson RJ. Lending a hand: social regulation of the neural response to threat. *Psychol Sci*. 2006;17(12):1032–1039. <https://doi.org/10.1111/j.1467-9280.2006.01832.x>.
- Erzen E, Çikrikci Ö. The effect of loneliness on depression: a meta-analysis. *Int J Soc Psychiatry*. 2018;64(5):427–435. <https://doi.org/10.1177/0020764018776349>.
- Cacioppo JT, Hawkley LC, Crawford LE, et al. Loneliness and health: potential mechanisms. *Psychosom Med*. 2002;64(3):407–417. <https://doi.org/10.1097/00006842-200205000-00005>.
- Cacioppo JT, Hughes ME, Waite LJ, Hawkley LC, Thisted RA. Loneliness as a specific risk factor for depressive symptoms: cross-sectional and longitudinal analyses. *Psychol Aging*. 2006;21(1):140–151. <https://doi.org/10.1037/0882-7974.21.1.140>.
- Santini ZI, Jose PE, York Cornwell E, et al. Social disconnectedness, perceived isolation, and symptoms of depression and anxiety among older Americans (NSHAP): a longitudinal mediation analysis. *Lancet Public Health*. 2020;5(1):e62–e70. [https://doi.org/10.1016/S2468-2667\(19\)30230-0](https://doi.org/10.1016/S2468-2667(19)30230-0).
- Richards HJ, Benson V, Donnelly N, Hadwin JA. Exploring the function of selective attention and hypervigilance for threat in anxiety. *Clin Psychol Rev*. 2014;34(1):1–13. <https://doi.org/10.1016/j.cpr.2013.10.006>.
- Brooks SK, Webster RK, Smith LE, et al. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. *Lancet*. 2020;395(10227):912–920. [https://doi.org/10.1016/S0140-6736\(20\)30460-8](https://doi.org/10.1016/S0140-6736(20)30460-8).
- Bai Y, Lin CC, Lin CY, Chen JY, Chue CM, Chou P. Survey of stress reactions among health care workers involved with the SARS outbreak. *Psychiatr Serv*. 2004;55(9):1055–1057. <https://doi.org/10.1176/appi.ps.55.9.1055>.
- Sprang G, Silman M. Posttraumatic stress disorder in parents and youth after health-related disasters. *Disaster Med Public Health Prep*. 2013;7(1):105–110. <https://doi.org/10.1017/dmp.2013.22>.

13. Holmes EA, O'Connor RC, Perry VH, et al. Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. *Lancet Psychiatry*. 2020;7(6):547–560. [https://doi.org/10.1016/S2215-0366\(20\)30168-1](https://doi.org/10.1016/S2215-0366(20)30168-1).
14. Pfefferbaum B, North CS. Mental health and the COVID-19 pandemic. *N Engl J Med*. 2020;383(6):510–512. <https://doi.org/10.1056/NEJMp2008017>.
15. Czeisler MÉ, Lane RI, Petrosky E, et al. Mental health, substance use, and suicidal ideation during the COVID-19 pandemic—United States, June 24–30, 2020. *MMWR Morbid Mortal Wkly Rep*. 2020;69(32):1049–1057. <https://doi.org/10.15585/mmwr.mm6932a1>.
16. Gloster AT, Lamnisos D, Lubenko J, et al. Impact of COVID-19 pandemic on mental health: an international study. *PLoS One*. 2020;15(12):e0244809. <https://doi.org/10.1371/journal.pone.0244809>.
17. Xiong J, Lipsitz O, Nasri F, et al. Impact of COVID-19 pandemic on mental health in the general population: a systematic review. *J Affect Disord*. 2020;277:55–64. <https://doi.org/10.1016/j.jad.2020.08.001>.
18. Paul MJ, Dredze M. You are what you tweet: analyzing Twitter for public health. In: Proceedings of the 5th International AAAI Conference on Weblogs and Social Media; July 17–21, 5; 2011. p. Barcelona, Spain 265–272. <https://doi.org/10.1609/icwsm.v5i1.14137>.
19. Schwartz HA, Ungar LH. Data-driven content analysis of social media: a systematic overview of automated methods. *Ann Am Acad Polit SS*. 2015;659(1):78–94. <https://doi.org/10.1177/0002716215569197>.
20. Schwartz HA, Eichstaedt JCE, Kern ML, et al. Characterizing geographic variation in well-being using tweets. In: Proceedings of the Seventh International AAAIO Conference on Weblogs and Social Media. Cambridge, MA, USA, 7; 2013. p. 583–591. <https://doi.org/10.1609/icwsm.v7i1.14442>.
21. Eichstaedt JC, Schwartz HA, Kern ML, et al. Psychological language on Twitter predicts county-level heart disease mortality. *Psychol Sci*. 2015;26(2):159–169. <https://doi.org/10.1177/0956797614557867>.
22. Jaidka K, Giorgi S, Schwartz HA, Kern ML, Ungar LH, Eichstaedt JC. Estimating geographic subjective well-being from Twitter: a comparison of dictionary and data-driven language methods. *Proc Natl Acad Sci U S A*. 2020;117(19):10165–10171. <https://doi.org/10.1073/pnas.1906364117>.
23. Gorgi A, Preotiuc-Pietro D, Buffone A, et al. Remarkable Benefit of User-Level Aggregation for Lexical-Based Population-Level Predictions. *arXiv*. Preprint. Online August 29, 2018. <https://doi.org/10.48550/arxiv.1808.09600>.
24. Schwartz HA, Eichstaedt J, Kern ML, et al. Towards assessing changes in degree of depression through Facebook. In: Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; June 2014. p. Baltimore, MD 118–125. <https://doi.org/10.3115/v1/W14-3214>.
25. Schwartz HA, Giorgi S, Sap M, Crutchley P, Ungar L, Eichstaedt JDLATK. In: Differential Language Analysis ToolKit. Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, 57–60; September 2017. p. Copenhagen, Denmark 55–60. <https://doi.org/10.18653/v1/D17-2010>.
26. Stijelja S, Mishara BL. COVID-19 and psychological distress—changes in internet searches for mental health issues in New York during the pandemic. *JAMA Intern Med*. 2020;180(12):1703–1706. <https://doi.org/10.1001/jamainternmed.2020.3271>.
27. Yang AC, Tsai SJ, Huang NE, Peng CK. Association of Internet search trends with suicide death in Taipei City, Taiwan, 2004–2009. *J Affect Disord*. 2011;132(1–2):179–184. <https://doi.org/10.1016/j.jad.2011.01.019>.
28. Tran US, Anzel R, Niederkrotenthaler T, Till B, Ajdacic-Gross V, Voracek M. Low validity of Google Trends for behavioral forecasting of national suicide rates. *PLoS One*. 2017;12(8):e0183149. <https://doi.org/10.1371/journal.pone.0183149>.
29. Brodeur A, Clark AE, Fleche S, Powdthavee N. COVID-19, lockdowns and well-being: evidence from Google Trends. *J Public Econ*. 2021;193:104346. <https://doi.org/10.1016/j.jpubeco.2020.104346>.
30. Knipe D, Evans H, Marchant A, Gunnell D, John A. Mapping population mental health concerns related to COVID-19 and the consequences of physical distancing: a Google trends analysis. *Wellcome Open Res*. 2020;5:82. <https://doi.org/10.12688/wellcomeopenres.15870.2>.
31. Lippi G, Henry BM, Sanchis-Gomar F. Putative impact of the COVID-19 pandemic on anxiety, depression, insomnia and stress. *Eur J Psychiatry*. 2021;35(3):200–201. <https://doi.org/10.1016/j.ejpsy.2020.11.006>.
32. Misiak B, Szcześniak D, Koczanowicz L, Rymaszewska J. The COVID-19 outbreak and Google searches: is it really the time to worry about global mental health? *Brain Behav Immun*. 2020;87:126–127. <https://doi.org/10.1016/j.bbi.2020.04.083>.
33. Knipe D, Gunnell D, Evans H, John A, Fancourt D. Is Google Trends a useful tool for tracking mental and social distress during a public health emergency? A time-series analysis. *J Affect Disord*. 2021;294:737–744. <https://doi.org/10.1016/j.jad.2021.06.086>.
34. Paul MJ, Dredze M, Broniatowski D. Twitter improves influenza forecasting. *PLoS Curr*. 2014;6 ecurrents.outbreaks.90b9ed0f59bae4ccaa683a39865d9117. <https://doi.org/10.1371/currents.outbreaks.90b9ed0f59bae4ccaa683a39865d9117>.
35. Guntuku SC, Sherman G, Stokes DC, et al. Tracking mental health and symptom mentions on twitter during COVID-19. *J Gen Intern Med*. 2020;35(9):2798–2800. <https://doi.org/10.1007/s11606-020-05988-8>.
36. Valdez D, Ten Thij M, Bathina K, Rutter LA, Bollen J. Social media insights into U.S. mental health during the COVID-19 pandemic: longitudinal analysis of twitter data. *J Med Internet Res*. 2020;22(12):e21418. <https://doi.org/10.2196/21418>.
37. Jaidka K, Eichstaedt J, Giorgi S, Schwartz HA, Ungar LH. Information-seeking vs. sharing: which explains regional health? An analysis of Google Search and Twitter trends. *Telemat Inform*. 2021;59. <https://doi.org/10.1016/j.tele.2020.101540>.
38. Juhng S, Matero M, Varadarajan V et al. Assessing anxiety from language: discourse-level embeddings improve prediction; 2022.