



Published in final edited form as:

J Affect Disord. 2023 December 01; 342: 63–68. doi:10.1016/j.jad.2023.08.141.

Estimating national and state-level suicide deaths using a novel online symptom search data source

Steven A. Sumner^{a,*}, Alen Alic^a, Royal K. Law^a, Nimi Idaikkadar^a, Nimesh Patel^b

^aNational Center for Injury Prevention and Control, U.S. Centers for Disease Control and Prevention, Atlanta, GA, USA

^bCenter for Forecasting and Outbreak Analytics, U.S. Centers for Disease Control and Prevention, Atlanta, GA, USA

Abstract

Background: Suicide mortality data are a critical source of information for understanding suicide-related trends in the United States. However, official suicide mortality data experience significant delays. The Google Symptom Search Dataset (SSD), a novel population-level data source derived from online search behavior, has not been evaluated for its utility in predicting suicide mortality trends.

Methods: We identified five mental health related variables (suicidal ideation, self-harm, depression, major depressive disorder, and pain) from the SSD. Daily search trends for these symptoms were utilized to estimate national and state suicide counts in 2020, the most recent year for which data was available, via a linear regression model. We compared the performance of this model to a baseline autoregressive integrated moving average (ARIMA) model and a model including all 422 symptoms (All Symptoms) in the SSD.

Results: Our Mental Health Model estimated the national number of suicide deaths with an error of –3.86 %, compared to an error of 7.17 % and 28.49 % for the ARIMA baseline and All Symptoms models. At the state level, 70 % (N = 35) of states had a prediction error of <10 % with the Mental Health Model, with accuracy generally favoring larger population states with higher number of suicide deaths.

Conclusion: The Google SSD is a new real-time data source that can be used to make accurate predictions of suicide mortality monthly trends at the national level. Additional research is needed to optimize state level predictions for states with low suicide counts.

Keywords

Suicide; Google; Online; Forecasting

*Corresponding author at: Centers for Disease Control and Prevention, National Center for Injury Prevention and Control, 4770 Buford Highway, NE, Mailstop F-63, Atlanta, GA 30341, USA. hvo5@cdc.gov (S.A. Sumner).

CRediT authorship contribution statement

SAS conceptualized the study, supervised the research, developed data visualizations, and led the writing. AA led the analysis and data visualization. NI and NP assisted in data acquisition and analysis. RL supervised the research and aided with methodology. All authors interpreted findings, provided methodological input, and reviewed and edited the manuscript.

Declaration of competing interest

We declare no competing interest.

1. Introduction

1.1. Suicide mortality data

Suicide continues to be a leading public health problem in the United States, with age-adjusted rates near their highest levels in more than two decades (Centers for Disease Control and Prevention, 2022b). Despite the urgency of suicide prevention and concerns about population-level deterioration of mental health following the COVID-19 pandemic (Ettman et al., 2022; Vahratian et al., 2021), national data on suicide rates from official mortality statistics are delayed, hampering a timely understanding and response to the problem (Centers for Disease Control and Prevention, 2023). Mortality statistics on suicide can take considerable time to process and finalize as a result of multiple factors such as the complexity of intent determination, the considerable time needed for post-mortem toxicology testing, and limitations in electronic records infrastructure at the local level (Spencer and Ahmad, 2017). The Centers for Disease Control and Prevention's (CDC) National Vital Statistics System (NVSS) has worked to improve data delays by releasing provisional data, which reduces the lag on national suicide data down from over one year for final data to approximately 7 months for provisional data (Centers for Disease Control and Prevention, 2023).

1.2. Modeling suicide deaths

To address these challenges, in part, researchers and public health experts have sought ways to model or estimate mortality statistics in the absence of timely data. Recent efforts in both suicide and opioid overdose have demonstrated that proxy data sources can accurately “nowcast” or estimate death rates in near real-time (Choi et al., 2020; Sumner et al., 2022). These approaches use statistical or machine learning based models to impute mortality trends from morbidity-related data.

While there is a need for ongoing prospective validation of these approaches, these methodologies have gained widespread usage in both infectious and non-infectious disease modeling (Lu et al., 2018; Reich et al., 2019; Rosenfeld and Tibshirani, 2021). As a consequence, researchers continue to seek and evaluate a wide variety of novel data sources and their applicability to such modeling efforts. One leading category of novel data sources includes information derived from online sources, such as online search or social media data.

1.3. Novel symptom search data

One leading online data source for suicide related research has historically been Google search trends data. The popularity of Google search data for health-related concerns emerged and grew after publication of initial studies pertaining to seasonal influenza (Dugas et al., 2012; Ginsberg et al., 2009). More than a decade of research has subsequently sought to refine and optimize the utility of this data source for other health concerns including suicide (Barros et al., 2019; Kristoufek et al., 2016; McCarthy, 2010; Sinyor et al., 2020). Google search data has traditionally been available via the Google Trends product (<https://trends.google.com/>). Google Trends uses information from search terms entered into Google's internet search page to produce publicly available metrics revealing the popularity

of a given keyword over time. However, results from cross-sectional studies examining the association of various suicide related keywords and suicide rates have yielded mixed results (Tran et al., 2017).

During the COVID-19 pandemic, as part of efforts to enhance realtime public health surveillance efforts, Google released a new dataset for health research using an improved methodology for health symptom identification. The dataset, known as the COVID-19 Search Trends Symptoms Dataset (commonly referred to as the Symptom Search Dataset or SSD), included information on 422 signs, symptoms, and health conditions (Gabrilovich, 2020; Google LLC, n.d.). Whereas the original Google Trends service provided data based on the relative popularity of search terms, the SSD employed an updated methodology based on Google's Knowledge Graph that enhanced the accuracy of identification of health-related searches by utilizing both the words used in queries and the health-related entities present in web pages subsequently viewed (Vaidyanathan et al., 2022).

While many of the symptoms in the SSD are relevant for infectious disease monitoring, other health conditions are represented as well, including measures of poor mental health. Given the need for timely data at both the national and state level to assess suicide mortality trends, we sought to evaluate the new Google Symptom Search Dataset for its utility in modeling population level trends in suicide deaths.

2. Methods

2.1. Data sources

Our primary predictor data was symptom trends as measured in the new Google SSD. From the SSD, we downloaded daily symptom trends information for the U.S. at the county-level. Data from the SSD are reported as normalized values ranging from 0 to 100, with higher values indicating increased popularity of searches related to this symptom relative to other periods of time. As noted above, the SSD contains information on >400 symptoms and we use information from all of these symptoms as well as a mental-health focused subset, as discussed below. As the temporal focus of analysis for this study was monthly, we summed daily level values to obtain total monthly values.

Our primary outcome of interest was monthly counts of suicide fatalities in the United States. Suicide deaths were identified from death certificate data from the National Vital Statistics System using the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, underlying cause of death codes U03, X60-X84, and Y87.0.

Monthly suicide deaths by county of residence were linked to county-level Google symptom data from January 1st, 2018, to December 31st, 2020. We focused on this time period as 2018 was the earliest full year that Google symptom data was available and 2020 was the most recent year of mortality data available. As detailed below, suicide count predictions were made at the county level then aggregated into the final state level predictions.

2.2. Machine learning approach

Three model variations were built for this study: a model using mental health related symptoms from the Google SSD as identified by CDC suicide subject matter experts (Mental Health model), a model incorporating all 422 symptoms in the SSD (All Symptoms), and a baseline autoregressive integrated moving average (ARIMA baseline) model. Consistent with best practices in machine learning, when building and testing the Google symptom-based models we used 2018 data in model training, 2019 data in model validation (e.g., selection of models and optimal hyperparameters), and 2020 data for model evaluation (out-of-sample testing). As Google SSD is available in near realtime, we used the current month's SSD values to predict that month's suicide count.

The Mental Health Model was our main model of interest and consisted of five mental health related variables (suicidal ideation, self-harm, depression, major depressive disorder, and pain) in the Google SSD, as identified by the authors of this study. These variables were selected given their strong association with suicide risk as demonstrated through extensive prior literature (Angst et al., 1999; Hooley et al., 2014). Three statistical machine learning algorithms were tested for making predictions from these five variables: Linear Multivariate Regression, Support Vector Regression, and Random Forest Regression. These three algorithms were selected for testing as they are among the model widely-used statistical approaches for prediction and each employ a distinct approach to model optimization. The optimal algorithm among these three for making monthly state-level suicide count predictions was chosen from results on the validation sample of the data (2019) and then predictions were made for 2020 and evaluated against actual suicide counts observed for 2020.

Our second model, the All Symptoms model, provided an important comparison as it attempted to leverage information from all 422 variables present in the Google SSD. This test was important as conceivably there is latent information present in other symptom trends that could improve prediction accuracy. Given the large number of predictor variables in this model relative to the number of observations in the dataset, we employed a LASSO (least absolute shrinkage and selection operator) model (Tibshirani, 1996). LASSO models are a form of linear regression which implement a penalization in the model fitting process that shrinks the beta-coefficients associated with unimportant features. Thus, the model aims to prevent overfitting in the setting of a large number of predictor variables.

Lastly, we compared our models to a baseline ARIMA model, one of the most widely used approaches to predicting injury mortality (Faust et al., 2021). The ARIMA model was trained on historical suicide count data from 2010 to 2018.

3. Results

Table 1 presents a national-level summary of model performance on the 2020 testing year. The Mental Health model had the best prediction performance, predicting a total of 43,722 suicides when aggregating information from all states. This corresponds to an error of only 3.86 %, compared to errors of 7.17 % and 28.49 % error for the ARIMA baseline and All Symptoms models, respectively. However, there was notable variability at the state-level,

even for the best performing model, and errors ranged from a low of 0.36 % to a high of 57.48 % for the Mental Health Model.

Fig. 1 further explores this state-level variability in prediction accuracy by plotting a histogram of the distribution of the percentage errors for the Mental Health Model. Although there were notable outliers, the majority of percentage errors clustered around zero with (35 or 70 %) of states having a prediction error <10 %.

As an effort to better explore reasons for the observed state-level variability, Fig. 2 displays the percentage error for each state by the actual number of suicides in a given state, with larger circles representing states with a larger number of suicide fatalities. In general, the model predicted suicide fatalities with a smaller percentage error for states with more suicide fatalities (which are also generally states with a higher total population), such as Texas and California. States with a smaller number of suicides, such as Hawaii and Rhode Island, were more vulnerable to having high prediction error.

Fig. 3 plots month-to-month trends for 2020 monthly suicide fatalities overlayed with the corresponding predictions during the same time period by state for the Mental Health model. Accuracy on a monthly level was variable. Many states experienced a predicted decrease in suicide fatalities in mid-2020, which was not ultimately seen in mortality data that later became available.

4. Discussion

This study is the first, to our knowledge, to evaluate the performance of a new real-time data source, the Google Symptom Search Dataset (SSD), on estimating suicide death counts. This work represents an important effort to assess the utility of this novel dataset for informing suicide-related trends in the United States. While the use of near realtime, symptom-based data sources has potential promise in adding to our knowledge of suicide-related trends, they are not intended to supplant use of gold standard data, such as official mortality records.

We found excellent performance in aggregate—national level suicide counts were predicted with a <4 % error. This is a robust result considering that only a single proxy data source is used in the modeling. Interestingly, this compares favorably to a prior study by our group which used the original Google Trends product in a similar modeling effort; in that study, Google Trends data alone had a 8.29 % error in estimating the national suicide rate (Choi et al., 2020). While these studies were performed across different years and are thus not a head-to-head comparison, this finding lends support to the notion that the Google Symptom Search Dataset may represent a superior signal for health-related research. This is consistent with the design of the SSD in that it more explicitly takes into account user intention when producing search trend based estimates as opposed to simply reporting on keyword trends (Vaidyanathan et al., 2022).

The original Google Trends data has been used extensively for mental health related research owing to its accessibility and that it is free of cost (Jimenez et al., 2020; McCarthy, 2010). However, there has been considerable debate in the literature about the reliability of this tool as a singular data source when used to estimate epidemiologic trends, with

systematic reviews and meta-analyses reporting variable performance based on keywords used and topic being modeled (Tran et al., 2017). It is possible that the SSD may improve this heterogeneity in mental health research using search trends data, however, it is important to note that at present the SSD has a limited and predefined number of mental health constructs based on previous health concepts built as part of Google's Knowledge Graph work (Vaidyanathan et al., 2022). We used these variables in our model (suicidal ideation, self-harm, depression, major depressive disorder, and pain); however, it is not possible for external researchers to inspect or modify the definitions for these symptoms/variables.

When assessing performance at the state-level, we also noted favorable results with most states exhibiting a <10 % error rate in annual suicide death estimation. However, performance was notably poorer for states with small populations and small suicide counts. The difficulty in fitting models for rare outcomes presents a significant challenge for efforts to develop accurate forecasts for small geographic areas, such as counties or cities, where focused prevention efforts may be most cost-effective.

It is important to note that our test year for model evaluation was 2020, the year in which COVID-19 emerged in the U.S. and vastly disrupted expected patterns of health-related trends and thus may have been a particularly challenging year to predict. Available evidence from household surveys suggests that symptoms of depressive disorder worsened nationally throughout 2020 (Vahratian et al., 2021) and the rate of emergency department visits for mental health conditions and suicide attempts similarly increased nationwide following the emergence of the pandemic (Holland et al., 2021). However, mortality data from 2020 reveals a decrease in the crude rate of suicide deaths (13.95 per 100,000 persons) compared to 2019 (14.47 per 100,000 persons) (Centers for Disease Control and Prevention, 2022b). The precise reasons for the divergence of morbidity and mortality trends for 2020 are not fully understood but may be the result of many factors including changes in behavioral health treatment seeking (Busch et al., 2021). Many researchers have also pointed out that suicide deaths are particularly hard to certify for coroners and medical examiners and may be increasingly misclassified, particularly with rapidly rising overdose deaths nationally which may be unintentional or intentional (Stone et al., 2017).

Nonetheless, this speaks to the complicated environment in which internet search trends must be evaluated. For example, the SSD shows that depression related searches were at their lowest point in March of 2020, which is the time point at which the COVID-19 pandemic first emerged as a major public health crisis in the U.S. It is likely that patients with depression did not experience sudden relief of their depression but rather internet search patterns for these individuals changed as other health related searches became of greater importance. These and other complicated behavioral patterns in 2020 likely challenged efforts to make accurate predictions from this sole data source.

Because we only had a single year (2020) for a fully held-out test set given the historical limited availability of the Google SSD as a new product, we did also examine 2019 prediction results as a sensitivity analysis. These findings should be cautiously interpreted as prediction results on a validation sample can be overly optimistic; however, findings revealed a predicted 45,728 suicides nationally in 2019 for a -2.69 % error. This is similar

to the error on national level prediction results for 2020 (−3.86 %). Furthermore, a similar number of states (N = 36) had a prediction error under 10 % for state level estimates in 2019 as compared to 2020 (N = 35).

Lastly, the poor performance of the model incorporating all 422 symptoms is an important finding. Although it is conceivable that a sophisticated model could harness information from a broader selection of symptoms, we found that this was not easily achieved with a widely used statistical machine learning model that incorporates variable selection and regularization and generally performs well for forecasting tasks (Sumner et al., 2022). Although there exists more sophisticated neural network based models for forecasting (Kapoor et al., 2020) that can and should be tested in future research, it is unclear how these models would perform given that a fundamental challenge in this dataset is the small number of observations on which to perform training. Because the SSD is a new data source, we only had a single year of data for model training; model performance and the ability to apply more sophisticated models to higher dimensional data will likely improve as more longitudinal data becomes available.

Important limitations of this work should be noted. This research primarily served as an evaluation of the potential performance of a novel dataset for suicide mortality estimation and revealed both promising findings as well as some sub-optimal results, such as in the accuracy of predictions for small states. This suggests that further evaluation work is needed to ensure that use of new data sources and prediction models are reliable for all localities. This should include prospective validation using additional years of data. Nonetheless, this study represents the first comparison of this data source to suicide mortality data and helps contribute to a better understanding of its potential strengths and limitations. Second, it is important to interpret all results cautiously as the study time period incorporated 2020, during which the COVID-19 pandemic had profound impacts on population level mental health and behavioral patterns which are still being elucidated. Nonetheless, model performance on this atypical year was respectable and is encouraging for continued use and evaluation of this data source. Additionally, there are fundamental limitations to the underlying data; for example, while the data provides population level trends in disease burden, it does not provide information on individual risk or protective factors. Third, there are practical implications for use of such a modeling approach. Our model evaluation uses training data from 2018 (2 years prior to the test year) and validation data from 2019 (one year prior to the test data). Data lags that affect access to recent mortality data could prevent retraining of deployed models in a timely fashion. Lastly, there remain many important directions for future work which we were unable to assess in this manuscript. These include testing additional types of machine learning models, such as those which may incorporate additional spatial information such as graph neural networks (Kapoor et al., 2020) and testing the utility of the SSD when combined with signals from other data sources (Choi et al., 2020).

Nevertheless, this research presents new and useful information in the ongoing search for novel data sources that can inform a more realtime understanding of suicide-related trends in the U.S. At minimum, more timely data can help inform funding for suicide prevention activities, ensuring that the level of resources matches the public health burden.

Population-level interventions exist for suicide prevention (Centers for Disease Control and Prevention, 2022a) but there is often lack of timely data to study the implementation of these interventions and measure their effectiveness in a rapid feedback loop. Some interventions—such as safe media messaging about suicide—can be implemented on a national scale and benefit from more timely national-level data (Niederkrötenhaler et al., 2021). With suicide rates near the highest levels in over two decades, improving the timeliness of suicide-related data will continue to be an important goal in assisting prevention efforts for this critical public health problem.

Role of funding source

No funding to declare. All authors were employees of the U.S. government.

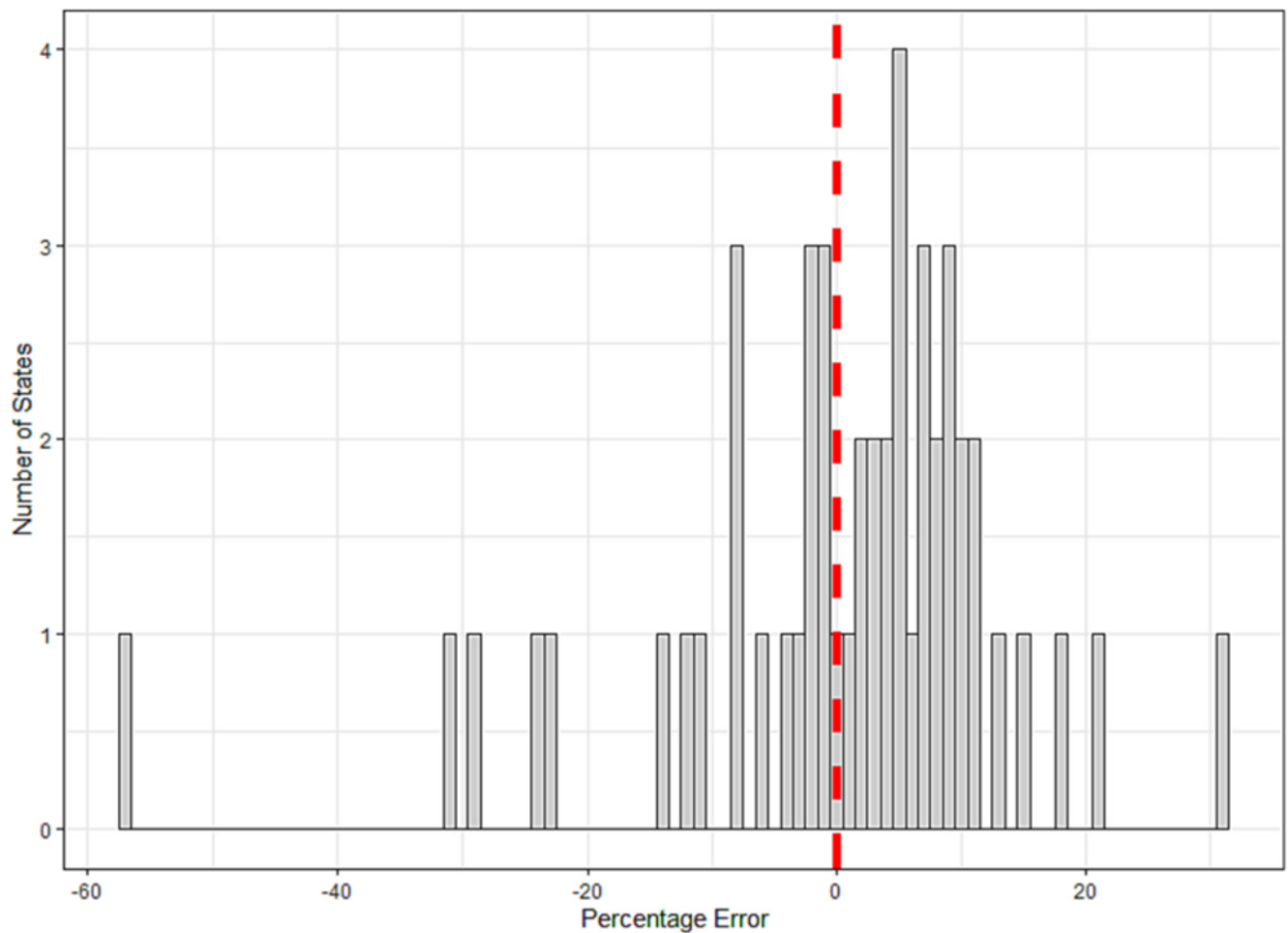
CDC disclaimer

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

References

- Angst J, Angst F, Stassen HH, 1999. Suicide risk in patients with major depressive disorder. *J. Clin. Psychiatry* 60 (2), 57–62.
- Barros JM, Melia R, Francis K, Bogue J, O’Sullivan M, Young K, Bernert RA, Rebholz-Schuhmann D, Duggan J, 2019. The validity of Google Trends search volumes for behavioral forecasting of national suicide rates in Ireland. *Int. J. Environ. Res. Public Health* 16 (17), 3201. [PubMed: 31480718]
- Busch AB, Sugarman DE, Horvitz LE, Greenfield SF, 2021. Telemedicine for treating mental health and substance use disorders: reflections since the pandemic. *Neuropsychopharmacology* 46 (6), 1068–1070. [PubMed: 33479513]
- Centers for Disease Control and Prevention, 2022a. Suicide Prevention Resource for Action: A Compilation of the Best Available Evidence. Atlanta, GA. <https://www.cdc.gov/suicide/pdf/preventionresource.pdf>.
- Centers for Disease Control and Prevention, 2022b. WISQARS - Fatal Injury Reports. <https://wisqars.cdc.gov/fatal-reports>.
- Centers for Disease Control and Prevention, 2023. CDC WONDER. Multiple Cause of Death Data <https://wonder.cdc.gov/mcd.html>.
- Choi D, Sumner SA, Holland KM, Draper J, Murphy S, Bowen DA, Zwald M, Wang J, Law R, Taylor J, 2020. Development of a machine learning model using multiple, heterogeneous data sources to estimate weekly US suicide fatalities. *JAMA Netw. Open* 3 (12), e2030932. [PubMed: 33355678]
- Dugas AF, Hsieh Y-H, Levin SR, Pines JM, Mareiniss DP, Mohareb A, Gaydos CA, Perl TM, Rothman RE, 2012. Google Flu Trends: correlation with emergency department influenza rates and crowding metrics. *Clin. Infect. Dis* 54 (4), 463–469. [PubMed: 22230244]
- Ettman CK, Cohen GH, Abdalla SM, Sampson L, Trinquart L, Castrucci BC, Bork RH, Clark MA, Wilson I, Vivier PM, 2022. Persistent depressive symptoms during COVID-19: a national, population-representative, longitudinal study of US adults. *The Lancet Regional Health-Americas* 5, 100091. [PubMed: 34635882]
- Faust JS, Du C, Mayes KD, Li S-X, Lin Z, Barnett ML, Krumholz HM, 2021. Mortality from drug overdoses, homicides, unintentional injuries, motor vehicle crashes, and suicides during the pandemic, March-August 2020. *JAMA* 326 (1), 84–86. [PubMed: 34019096]
- Gabrilovich E, 2020. Using Symptoms Search Trends to Inform COVID-19 Research. <https://blog.google/technology/health/using-symptoms-search-trends-inform-covid-19-research/>.
- Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L, 2009. Detecting influenza epidemics using search engine query data. *Nature* 457 (7232), 1012–1014. [PubMed: 19020500]

- Google LLC. Google COVID-19 Search Trends Symptoms Dataset. <http://goo.gle/covid19symptomdataset>. Accessed: May 19, 2021.
- Holland KM, Jones C, Vivolo-Kantor AM, Idaikkadar N, Zwald M, Hoots B, Yard E, D’Inverno A, Swedo E, Chen MS, 2021. Trends in US emergency department visits for mental health, overdose, and violence outcomes before and during the COVID-19 pandemic. *JAMA Psychiatry* 78 (4), 372–379. [PubMed: 33533876]
- Hooley JM, Franklin JC, Nock MK, 2014. Chronic pain and suicide: understanding the association. *Curr. Pain Headache Rep* 18 (8), 1–6.
- Jimenez A, Santed-Germán M-A, Ramos V, 2020. Google Searches and suicide rates in Spain, 2004-2013: correlation study. *JMIR Public Health Surveill.* 6 (2), e10919. [PubMed: 32281540]
- Kapoor A, Ben X, Liu L, Perozzi B, Barnes M, Blais M, & O’Banion S (2020). Examining covid-19 forecasting using spatio-temporal graph neural networks. *arXiv preprint arXiv:2007.03113*.
- Kristoufek L, Moat HS, Preis T, 2016. Estimating suicide occurrence statistics using Google trends. *EPJ Data Sci.* 5, 1–12.
- Lu FS, Hou S, Baltrusaitis K, Shah M, Leskovec J, Hawkins J, Brownstein J, Conidi G, Gunn J, Gray JJJPH, surveillance, 2018. Accurate influenza monitoring and forecasting using novel Internet data streams: a case study in the Boston Metropolis, 4 (1), e4.
- McCarthy MJ, 2010. Internet monitoring of suicide risk in the population. *J. Affect. Disord* 122 (3), 277–279. 10.1016/j.jad.2009.08.015. [PubMed: 19748681]
- Niederkröthaler T, Tran US, Gould M, Sinyor M, Sumner S, Strauss MJ, Voracek M, Till B, Murphy S, Gonzalez F, 2021. Association of Logic’s hip hop song “1-800-273-8255” with lifeline calls and suicides in the United States: interrupted time series analysis. *BMJ* 375.
- Reich NG, Brooks LC, Fox SJ, Kandula S, McGowan CJ, Moore E, Osthus D, Ray EL, Tushar A, Yamana TK, 2019. A collaborative multiyear, multimodel assessment of seasonal influenza forecasting in the United States. *Proc. Natl. Acad. Sci* 116 (8), 3146–3154. [PubMed: 30647115]
- Rosenfeld R, Tibshirani RJ, 2021. Epidemic tracking and forecasting: lessons learned from a tumultuous year. *Proc. Natl. Acad. Sci* 118 (51), e2111456118. [PubMed: 34903658]
- Sinyor M, Spittal MJ, Niederkröthaler T, 2020. Changes in suicide and resilience-related Google searches during the early stages of the COVID-19 pandemic. *Can. J. Psychiatr* 65 (10), 741–743.
- Spencer M, Ahmad F, 2017. Timeliness of Death Certificate Data for Mortality Surveillance and Provisional Estimates. National Center for Health Statistics.
- Stone DM, Holland KM, Bartholow B, Logan J, LiKamWa McIntosh W, Trudeau A, Rockett IR, 2017. Deciphering suicide and other manners of death associated with drug intoxication: a Centers for Disease Control and Prevention consultation meeting summary. *Am. J. Public Health* 107 (8), 1233–1239. [PubMed: 28640689]
- Sumner SA, Bowen D, Holland K, Zwald ML, Vivolo-Kantor A, Guy GP, Heuett WJ, Pressley DP, Jones CM, 2022. Estimating weekly national opioid overdose deaths in near real time using multiple proxy data sources. *JAMA Netw. Open* 5 (7), e2223033. [PubMed: 35862045]
- Tibshirani R, 1996. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. B. Methodol* 58 (1), 267–288.
- Tran US, Anel R, Niederkröthaler T, Till B, Ajdacic-Gross V, Voracek M, 2017. Low validity of Google Trends for behavioral forecasting of national suicide rates. *PLoS One* 12 (8), e0183149. [PubMed: 28813490]
- Vahratian A, Blumberg SJ, Terlizzi EP, Schiller JS, 2021. Symptoms of anxiety or depressive disorder and use of mental health care among adults during the COVID-19 pandemic—United States, August 2020–February 2021. *Morb. Mortal. Wkly Rep* 70 (13), 490.
- Vaidyanathan U, Sun Y, Shekel T, Chou K, Galea S, Gabrilovich E, Wellenius GA, 2022. An evaluation of Internet searches as a marker of trends in population mental health in the US. *Sci. Rep* 12 (1), 1–9. [PubMed: 34992227]

**Fig. 1.**

Distribution of state-level percentage errors for suicide fatality predictions.

Note: Figure shows percentage error in suicide predictions for all states for 2020 test year; most prediction errors are within $\pm 10\%$ of zero. Predictions generated using Google Symptom Search Dataset and evaluated using death certificate data from the National Vital Statistics System using the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, underlying cause of death codes U03, X60-X84, and Y87.0.

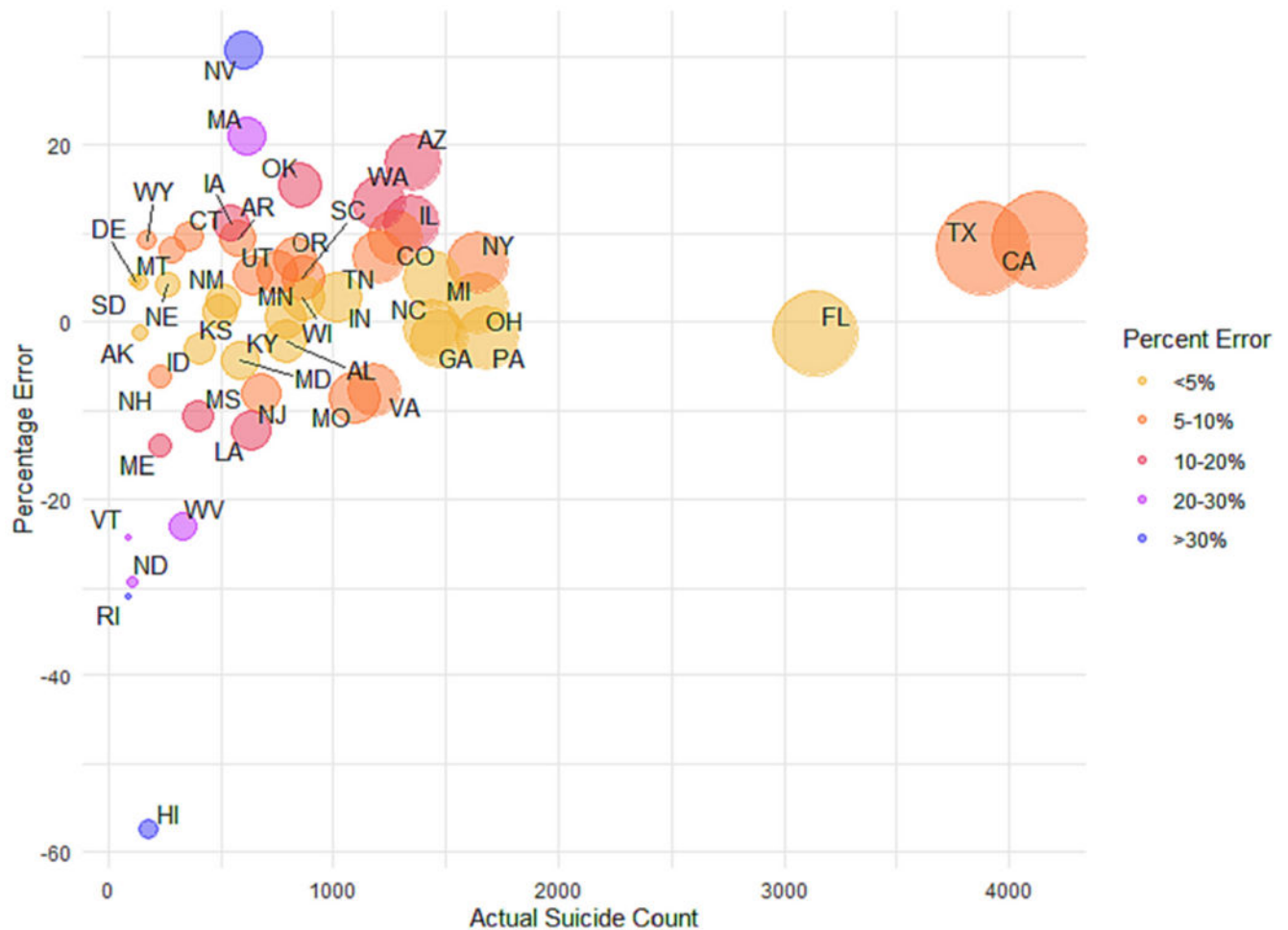
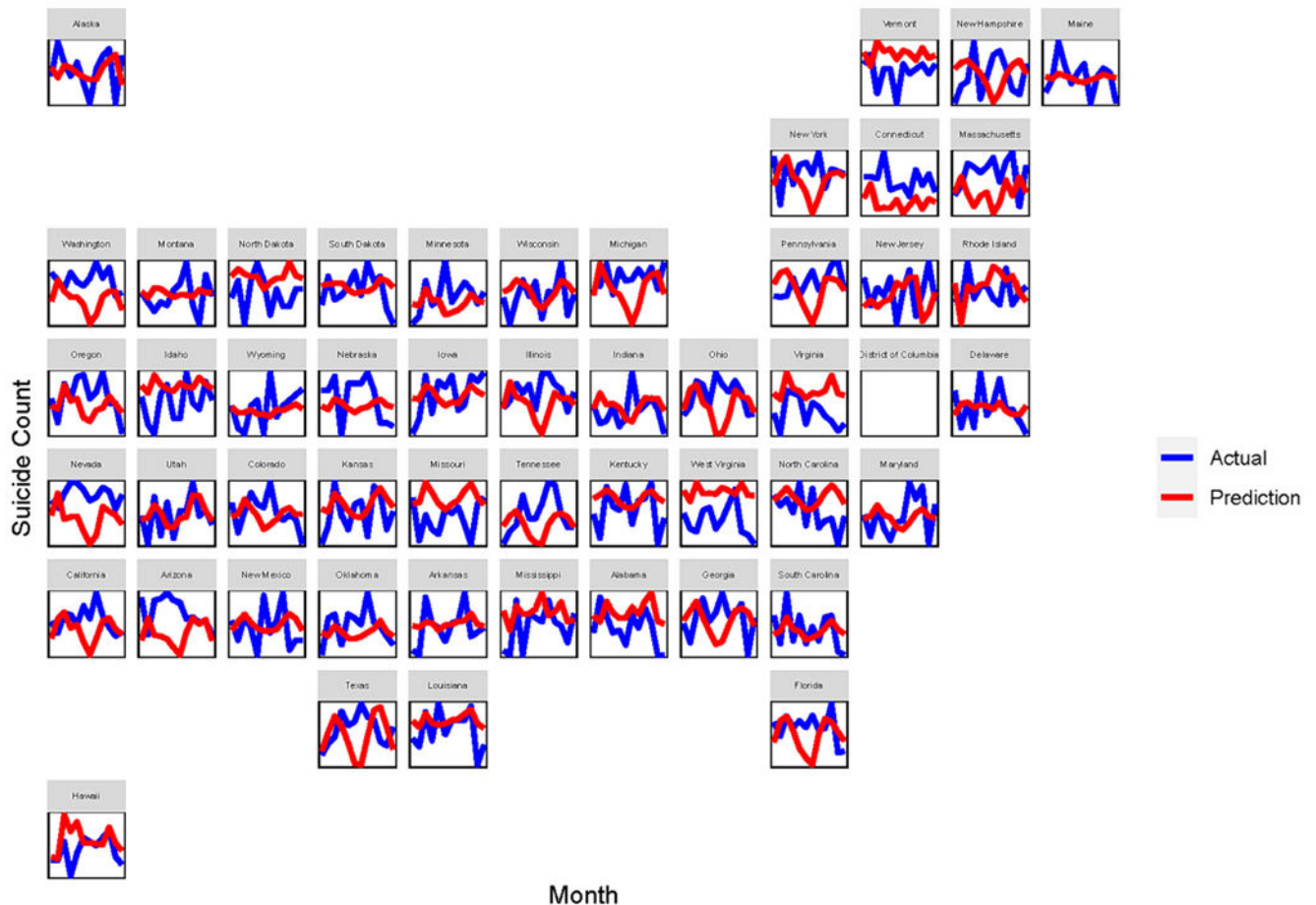


Fig. 2.

Percentage error of state suicide predictions by state suicide count.

Note: Figure displays percentage error in suicide predictions for all states for 2020 by magnitude of suicide deaths in a given state, revealing that states with a larger number of suicide generally have lower and more stable prediction errors. Predictions generated using Google Symptom Search Dataset and evaluated using death certificate data from the National Vital Statistics System using the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, underlying cause of death codes U03, X60-X84, and Y87.0.

**Fig. 3.**

Actual and predicted monthly suicide count trends by state, 2020.

Note: Results displayed are for Mental Health Model. X-axis for each state panel represents time in months for the test year of 2020; Y-axis displays suicide counts in each state and are scaled to the magnitude of suicide deaths in each state to ensure visibility of trend. Magnitude of differences may appear larger than actual in certain states as counts <10 not displayed due to mortality data privacy provisions. Predictions generated using Google Symptom Search Dataset and evaluated using death certificate data from the National Vital Statistics System using the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, underlying cause of death codes U03, X60-X84, and Y87.0.

Table 1

Performance of baseline and symptom search models for predicting national suicide fatalities, 2020.

	Predicted suicides	Error (%)	Standard deviation of error (%)	Range (absolute %)
ARIMA Baseline	48,739	7.17 %	7.78 %	0.17 % - 37.66 %
Mental Health Model	43,722	-3.86 %	10.27 %	0.36 % - 57.48 %
All Symptoms Model	58,433	28.49 %	31.61 %	0.17 % - 179.97 %

Note: Actual number of suicides in 2020 with geographic information at the U.S. county-level was 45,477.