Published in final edited form as:
Drug Alcohol Depend. 2018 September 01; 190: 166–169. doi:10.1016/j.drugalcdep.2018.05.009.

Internet searches for opioids predict future emergency department heroin admissions

Sean D. Young[a,*], **Kai Zheng**[b], **Larry F. Chu**[c], and **Keith Humphreys**[c,d]

[a]Department of Family Medicine, University of California Institute for Prediction Technology, David Geffen School of Medicine, University of California, Los Angeles, CA, USA

[b]Department of Informatics, Donald Bren School of Information and Computer Sciences, University of California, Irvine, CA, USA

[c]Department of Anesthesia, Stanford University School of Medicine, Stanford, California, USA

[d]Veterans Affairs Health Care System and Stanford University School of Medicine, Palo Alto, California, USA

Abstract

Background: For a number of fiscal and practical reasons, data on heroin use have been of poor quality, which has hampered the ability to halt the growing epidemic. Internet search data, such as those made available by Google Trends, have been used as a low-cost, real-time data source for monitoring and predicting a variety of public health outcomes. We aimed to determine whether data on opioid-related internet searches might predict future heroin-related admissions to emergency departments (ED).

Methods: Across nine metropolitan statistical areas (MSAs) in the United States, we obtained data on Google searches for prescription and non-prescription opioids, as well as Substance Abuse and Mental Health Services Administration (SAMHSA) data on heroin-related ED visits from 2004 to 2011. A linear mixed model assessed the relationship between opioid-related Internet searches and following year heroin-related visits, controlling for MSA GINI index and total number of ED visits.

Results: The best-fitting model explained 72% of the variance in heroin-related ED visits. The final model included the search keywords “Avinza,” “Brown Sugar,” “China White,” “Codeine,”

[*]Corresponding author at: University of California, Institute for Prediction Technology, 10880 Wilshire Blvd. Suite 1800, Los Angeles, CA 90024, USA.

Contributors

Dr. Young had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Young conceptualized and designed the study. Young, Zheng, Chu, and Humphreys contributed to the analysis/interpretation/acquisition of the data. Young and Chu drafted the manuscript. Young, Zheng, Chu, and Humphreys participated in critical revision of the manuscript for important intellectual content. Young and Zheng were responsible for statistical analysis Funding was obtained by Young.: Young, Zheng, Chu, and Humphreys provided administrative, technical, or material support. Study supervision was undertaken by Young and Zheng.

Conflict of interest

None to report.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.drugalcdep.2018.05.009.

"Radian," "Methadone," and "Oxymorphone." We found regional differences in where and how people searched for opioid-related information.

Conclusions: Internet search-based modeling should be explored as a new source of insights for predicting heroin-related admissions. In geographic regions where no current heroin-related data exist, Internet search modeling might be a particularly valuable and inexpensive tool for estimating changing heroin use trends. We discuss the immediate implications for using this approach to assist in managing opioid-related morbidity and mortality in the United States.

Keywords

Opioids; Social media; Internet search data; Heroin; Emergency department

1. Introduction

Opioid-related emergency department (ED) admissions and deaths have rapidly increased in the United States during the past decade (Rudd et al., 2016). In particular, heroin-related deaths have more than quintupled from 2000 to 2014, as some people have shifted from taking prescription opioids to heroin (Compton et al., 2016; Jones et al., 2015; Martins et al., 2017). For a number of fiscal and practical reasons, data on the status and evolution of heroin use have been of poor quality, which has hampered the ability to halt the epidemic (Humphreys, 2017; Kilmer and Caulkins, 2014; Ruhm, 2017). New methods for monitoring prescription and non-prescription opioid use are therefore needed to prevent heroin-related hospital admissions.

Internet search data, such as those provided by Google Trends, have been used as a low-cost, real-time data source for monitoring and predicting a variety of public health outcomes, including influenza, skin cancer, and depression (Bloom et al., 2015; Nuti et al., 2014). Recent models based on internet search and social media data have been improving and overcoming earlier limitations, gaining the attention of health care organizations, public health clinics, and government agencies who are beginning to implement these approaches in disease prevention and monitoring (Young, 2014). Internet search data might similarly be applied as a potential source for insights to address the opioid epidemic; however, no known research has explored this area.

Based on the hypothesis that heroin-related emergency department admissions might be preceded by people searching online for information relevant to opioids, we sought to identify whether Google Trends data could be used to predict the following year's emergency department heroin-related admissions. Because both prescription and non-prescription opioid use has been associated with future heroin use (Compton et al., 2016), we compiled a list of search terms for common descriptors of prescription and non-prescription related opioids, and evaluated whether a model including these search terms might predict the following year's emergency department heroin-related visits.

2. Materials and methods

We obtained annual data on heroin-related ED visits over the most recent six-year period of data availability (2005–2011) from the Substance Abuse and Mental Health Services Administration (SAMHSA). The dataset comprised ten metropolitan statistical areas (MSAs) in the United States: Boston, Chicago, Denver, Detroit, Minneapolis, New York City, Phoenix, San Francisco, Seattle, and Miami. However, Miami was excluded from this study due to missing data from a number of years, leaving nine MSAs. We then identified a list of common descriptors for prescription and non-prescription opioids from the Drug Enforcement Agency (DEA) website listing the clinical and street names for relevant drugs (Drug Enforcement Agency, Houston Division, 2017). Due to the large number of opioid-related terms in the DEA list, we attempted to refine the list by manually entering each opioid-related descriptor into a Google Trends search to determine the descriptors that were most frequently used in Google searches. Based on this initial screening, 12 prescription and non-prescription-related opioids were identified for use in the final model. Then, we obtained aggregated search data on Google searches using the Google Trends Application Programming Interface (API) for the period 2004–2011 (beginning one year prior to our existing heroin data in order to use Internet search data to predict following year heroin-related admissions). Google Trends returned data on the normalized search volume, which is a decimal number between 0 and 1, provided by Google based on the relative frequency of search terms compared to each other, for each of the 12 opioid-related keywords. The searches were conducted on August 4, 2017 for each MSA (see Supplementary materials).

A linear mixed model assessed the relationship between the normalized volume of internet searches using the opioid-related search terms and following year (natural log-transformed) heroin-related visits, controlling for MSA GINI index (a measure of poverty-related inequality) and total number of ED visits for all causes. We employed a two-step process to determine the best-fitting model that incorporated a combined list of the opioid-related search terms. First, we entered all 12 opioid-related search terms into the model and used a step-wise comparison, ranking all models of search keyword combinations by BIC statistic. Second, we assessed the quality of fit of the model by analyzing the marginal R2 value (Nakagawa and Schielzeth, 2013). Analyses were conducted using R software in August 2017. The IRB waived this study, as data are anonymous and aggregated.

3. Results

Across all MSA’s, heroin-related admissions increased from 1194.7 per 100,000 population in 2005, to 1358.3 in 2011. Minneapolis had the largest annual percentage increase in heroin-related ED visits (22.7%) over the study period, primarily due to a 53.5% increase in admissions between 2010 and 2011. San Francisco had the largest annual decrease over the study period (−20.6%). The largest percentage increase across all cities occurred in 2011 (11.3%).

The analysis revealed regional differences based on where people searched for opioid-related information. We found relatively fewer searches for “Avinza” across all MSA’s, except in Detroit and New York. We also found relatively fewer searches for “China White” in Denver and Minneapolis; and “Kadian,” in Denver, Phoenix, and Seattle. “Brown Sugar” was the most frequently searched term in all cities except in Phoenix and Seattle, where “Methadone” was the most searched opioid-related keyword.

The best-fitting model explained 72% of the variance in heroin-related ED visits (R^2 = .72). The final model included the search keywords “Avinza,” “Brown Sugar,” “China White,” “Codeine,” “Kadian,” “Methadone,” and “Oxymorphone” (Table 1; Fig. 1). GINI and total number of emergency department visits were not significant predictors and were automatically dropped from the final model. “Brown Sugar” was not a significant predictor but was automatically included in the final model because it was found to improve model fit. “Avinza,” “Methadone,” and “China White” were the strongest model contributors.

4. Discussion

Internet search data may enable novel, low-cost, and early detection of public health events or outcomes (Nuti et al., 2014). These data can be a highly cost-effective source for insights as they are freely and publicly available for research use and provide information on the opioid-related psychology and search behaviors of individuals. In geographic regions or countries where no current heroin-related data exist, internet search might be a valuable tool in providing at least a first estimate for changing heroin use trends.

Internet search data are being increasingly used for public health forecasting and understanding concerns for a variety of public health concerns because they are highly correlated with conventional surveillance data that are often costly to collect and suffer from reporting lag-time (Young et al., 2018). In fact, a primary barrier to addressing the opioid epidemic, as pointed out by public health officials and opioid researchers, is the inability to access recent data on opioid-related outcomes (Kilmer and Caulkins, 2014; Ruhm, 2017). For example, this study used heroin-related admissions from 2004 to 2011 (a seven-year lag time to present) because no more recent publicly-available data across the United States on heroin-related admissions were available. Although efforts are being made to provide access to high quality data on the evolution of the epidemic, most researchers are currently limited in their ability to access these data (Office of the Chief Technology Officer, 2017). However, near real-time or even predictive assessments of opioid-related trends are needed in order for researchers and public health officials to intervene in a timely fashion. Because Internet search data occur in real-time and have already been shown to be associated with current and future health outcomes, they may help to address this issue by allowing researchers and public health officials to gain more timely insights into the evolution of the opioid epidemic in order to intervene and prevent negative opioid-related outcomes, such as heroin overdose.

The ability to use online search behaviors to predict future heroin visits has several immediate implications for managing opioid-related morbidity and mortality in the United States. First, geolocation of search behaviors may offer more precise and targeted community-based opioid overdose education and naloxone distribution interventions to proactively target high-need areas (Kerensky and Walley, 2017). For example, if a model or software program were able to identify an increase in opioid-related search terms associated with future heroin-related admissions (such as in Indiana prior to the 2014 Indiana HIV outbreak), then public health officials could be better prepared for rapid education and medical interventions. Second, agencies such as the Department of Health and Human Services have already requested new methods of using real-time, low-cost data sources to aid as potential solutions in addressing the opioid crisis in the United States. Recent efforts are focused on facilitating collaborations between academia, industry, and government to address the opioid epidemic (e.g., http://predictiontechnology.ucla.edu/hack-a-thon-2018/), and Internet-based opioid prediction models provide a good use case for developing these partnerships. Third, the present methods create a use case for how Internet and even social media data might be used to better understand people's real-time public opioid-related intentions and behaviors. For example, analyzing opioid-related search terms could help to uncover changing trends in the descriptors that people use for prescription and non-prescription opioids, both regionally and longitudinally. These types of social and behavioral data have already been integrated with health care system utilization and outcomes and can be provided to public health officials so that prevention efforts can be quickly assessed and rapidly refined at regional and national levels (Young and Heinzerling, 2017).

This pilot study has several limitations. First, it is subject to the common sample bias associated with Internet search data. For example, not all Americans use Google as their primary search provider, and Google Trends does not provide raw data on the number of searches but rather normalized search volume. Several earlier studies using Google search data have also identified flaws in their predictive ability due to such limitations. However, as newer data science models have been created, more recent studies have been able to address many of these previous limitations. Although the present study incorporates many of these methods that have been shown to address these limitations, we believe that the primary value of the current study is not to provide a definitive model for predicting future heroin-related admissions, but rather to present a case study to illustrate how this new tool can be used (and refined with future work) to address heroin and opioid-related public health problems. Second, our analysis was based on search keywords for opioids in order to explore the feasibility of this new approach. Unlike social media (e.g., Twitter data), which can provide context around people's thoughts and behaviors, search data typically provide only keywords without context for the search. We conducted this study with an a priori list of search terms to explore these methods in a pilot study. Future modeling efforts may build upon this work by incorporating other "red flag" symptoms and behaviors for heroin use, such as searches related to needle sharing, as well as by integrating more contextual social media data (Young, 2014). Third, "Brown Sugar" was the most frequently searched term in all model keywords. This analysis did not filter out the many searches conducted using this keyword related to the baking product rather than to heroin. However, and possibly for this reason, "Brown Sugar" was not a significant predictor of future heroin admission. Finally, we were limited in the recency and breadth of the heroin-related admissions data (complete SAMHSA data were only available for nine MSAs through 2011), as well as reliance on the DEA site for identifying opioid-related search terms. For example, this analysis did not include data on rural towns or sub-urban and micropolitan areas, even though opioid use has been heavily linked to small urban and non-urban areas in the United States (Cicero et al., 2014). Because opioid use and the number of heroin admissions continue to evolve over time, more recent data, including data from small urban and non-urban areas, are urgently needed to be incorporated into models to provide more accurate and timely assessments.

5. Conclusion

Internet search data may be a valuable new source for insights in predicting heroin-related admissions. Due to the lack of available data on heroin and opioid-related behaviors and outcomes, and the urgent need for these data, researchers should further explore whether and how Internet search data may be used as a novel and supplementary source of information for addressing the opioid crisis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Role of the funding source

This study was funded by the National Institute on Drug Abuse, National Institute of Allergy and Infectious Diseases, University of California, Office of the President (UCOP), and National Human Genome Research Institute. The funders played no role in the study planning, analysis, or outcomes. Dr. Humphreys holds a Senior Research Career Scientist award from the Veterans Affairs Health Research and Development Service.

References

Bloom R, Amber KT, Hu S, Kirsner R, 2015 Google search trends and skin cancer: evaluating the US population's interest in skin cancer and its association with melanoma outcomes. JAMA Dermatol. 151, 903–905. 10.1001/jamadermatol.2015.1216. [PubMed: 26061357]

Cicero TJ, Ellis MS, Surratt HL, Kurtz SP, 2014 The changing face of heroin use in the United States: a retrospective analysis of the past 50 years. JAMA Psychiatry 71, 821–826. 10.1001/jamapsychiatry.2014.366. [PubMed: 24871348]

Compton WM, Jones CM, Baldwin GT, 2016 Relationship between nonmedical prescription-opioid use and heroin use. N. Engl. J. Med. 374, 154–163. 10.1056/NEJMra1508490. [PubMed: 26760086]

Drug Enforcement Agency, Houston Division, 2017 DEA Intelligence Report: Drug Slang Code Words. https://ndews.umd.edu/sites/ndews.umd.edu/files/dea-drug-slang-code-words-may2017.pdf.

Humphreys K, 2017 Analysis | The Federal Government Is Systematically Undercounting Heroin Users [WWW Document]. URL. Washington Post (Accessed 05 January 2018). https://www.washingtonpost.com/news/wonk/wp/2017/08/22/the-federal-government-is-systematically-under-counting-heroin-users/?noredirect=on&utm_term=.72712272a4c9.

Jones CM, Logan J, Gladden RM, Bohm MK, 2015 Vital signs: demographic and substance use trends among heroin users — United States, 2002–2013. MMWR Morb. Mortal. Wkly. Rep. 64, 719–725. [PubMed: 26158353]

Kerensky T, Walley AY, 2017 Opioid overdose prevention and naloxone rescue kits: what we know and what we don't know. Addict. Sci. Clin. Pract. 12 (4). 10.1186/s13722-016-0068-3.

Kilmer B, Caulkins JP, 2014 Hard Drugs Demand Solid Understanding [WWW Document]. URL. (Accessed 05 January 2018). https://www.rand.org/blog/2014/03/hard-drugs-demand-solid-understanding.html.

Martins SS, Sarvet A, Santaella-Tenorio J, Saha T, Grant BF, Hasin DS, 2017 Changes in US lifetime heroin use and heroin use disorder: prevalence from the 2001–2002 to 2012–2013 National Epidemiologic Survey on Alcohol and Related Conditions. JAMA Psychiatry 74, 445–455. 10.1001/jamapsychiatry.2017.0113. [PubMed: 28355458]

Nakagawa S, Schielzeth H, 2013 A general and simple method for obtaining R2 from generalized linear mixed-effects models. Methods Ecol. Evol. 4, 133–142. 10.1111/j.2041-210x.2012.00261.x.

Nuti SV, Wayda B, Ranasinghe I, Wang S, Dreyer RP, Chen SI, Murugiah K, 2014 The use of Google Trends in health care research: a systematic review. PLoS One 9, e109583 10.1371/journal.pone.0109583. [PubMed: 25337815]

Office of the Chief Technology Officer, 2017 The HHS Opioid Symposium [WWW Document]. HHS.gov. URL https://www.hhs.gov/challenges/symposium/index.html (Accessed 05 February 2018).

Rudd RA, Seth P, David F, Scholl L, 2016 Increases in drug and opioid-involved overdose deaths - United States, 2010–2015. MMWR Morb. Mortal. Wkly. Rep. 65, 1445–1452. 10.15585/mmwr.mm655051e1. [PubMed: 28033313]

Ruhm CJ, 2017 Geographic variation in opioid and heroin involved drug poisoning mortality rates. Am. J. Prev. Med. 53, 745–753. 10.1016/j.amepre.2017.06.009. [PubMed: 28797652]

Young SD, 2014 Behavioral insights on big data: using social media for predicting biomedical outcomes. Trends Microbiol. 22, 601–602. 10.1016/j.tim.2014.08.004. [PubMed: 25438614]

Young SD, Heinzerling K, 2017 The Harnessing Online Peer Education (HOPE) intervention for reducing prescription drug abuse: a qualitative study. J. Subst. Use 22, 592–596. 10.1080/14659891.2016.1271039. [PubMed: 29551953]

Young SD, Torrone EA, Urata J, Aral SO, 2018 Using search engine data as a tool to predict syphilis. Epidemiology 29, 574–578. 10.1097/EDE.0000000000000836. [PubMed: 29864105]

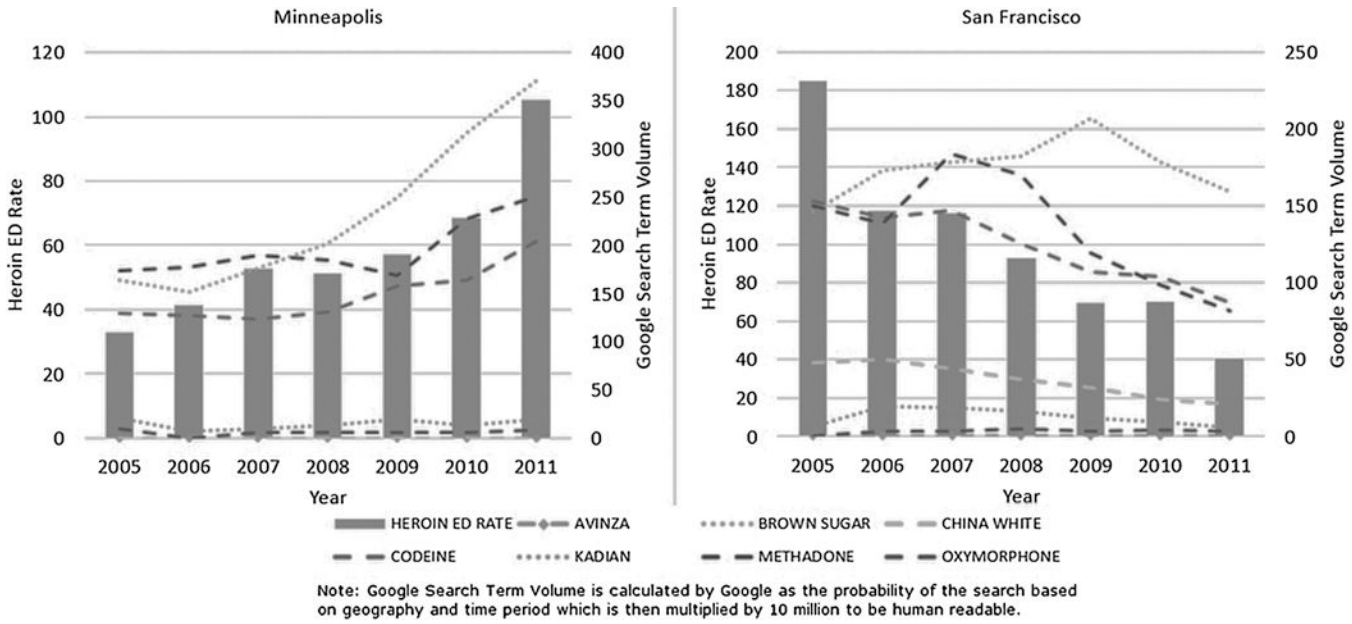

Fig. 1.

Google Trends opioid searches (January 1, 2004 to December 31, 2011) and following year heroin ED rate (per 100,000 people), example of 2 cities.

Table 1:

Multivariate Linear Mixed Model analysis of associations between search frequencies and Emergency Department heroin visits.

	Coefficient	Lower 95% CI	Upper 95% CI	Standard Error	p-value
Avinza	0.25	0.13	0.37	0.07	<.001
Brown Sugar	0.06	−0.04	0.16	0.05	0.27
China White	0.16	0.02	0.30	0.07	0.03
Codeine	0.13	0.05	0.21	0.04	0.004
Kadian	0.09	0.00	0.17	0.04	0.05
Methadone	0.19	0.09	0.29	0.05	<.001
Oxymorphone	−0.15	−0.21	−0.09	0.03	<.001