UC Irvine
UC Irvine Previously Published Works

Title
Regional variation in discussion of opioids on social media

Permalink
https://escholarship.org/uc/item/09z4j7hc

Journal
Journal of Addictive Diseases, 39(3)

ISSN
1055-0887

Authors
Flores, Lidia
Young, Sean D

Publication Date
2021-07-01

DOI
10.1080/10550887.2021.1874804

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

HHS Public Access
Author manuscript
J Addict Dis. Author manuscript; available in PMC 2022 July 01.

Published in final edited form as:
J Addict Dis. 2021 ; 39(3): 316–321. doi:10.1080/10550887.2021.1874804.

Regional Variation in Discussion of Opioids on Social Media: A Qualitative Study

Lidia Flores[+], **Sean D. Young, PhD, MS**[+,*]

[+]Department of Emergency Medicine, University of California, Irvine, California, United States of America

[*]University of California Institute for Prediction Technology, Department of Informatics, University of California, Irvine, California, United States of America

Abstract

Background: New data sources and analysis methods are urgently needed to improve opioid surveillance and prevent potential overdose. Social media data is one potential data source that might be used and integrated to address this issue.

Objective: This study explored opioid-related topics discussed across geographical regions of varying population sizes to determine whether social media data might inform opioid surveillance.

Methods: Between March 17th to July 17th, 2020, we collected tweets (N= 19,721) mentioning opioid-related keywords across 7 cities within the United States.

Results: Results found that opioid-related keywords were distributed as follows: New York (29%), Los Angeles (23%), Chicago (18%), Atlanta (18%), San Francisco (8%), Iowa (3%), and Orange County, CA (1%). We also found regional differences in the types of opioids and topics mentioned.

Conclusions: Findings suggest the feasibility of using opioid-related social media data to inform surveillance efforts, as well as potential regional and time-varying differences in topics discussed.

Keywords

real-time data; public health surveillance; opioids; drug consumption; geographical locations; Twitter data

INTRODUCTION

Opioid overdose prevention is one of the top public health concerns in the United States. The Centers for Disease Control and Prevention estimates that 128 people die of an opioid overdose every day. In 2018, there were 46,802 fatal overdoses involving opioids across the nation [1]. Of these overdoses, approximately one third involved heroin, a highly addictive

Corresponding Author Sean D. Young, PhD, MS, University of California, Irvine, SOM – Emergency Medicine, Irvine, CA 92697-4064, *Telephone*: 714-456-5239; *Fax*: 714-456-3719; syoung5@hs.uci.edu, **First Author** Lidia Flores, University of California, Irvine, SOM – Emergency Medicine, Irvine, CA 92697-4064, *Telephone*: 949-510-2049; floreslc@hs.uci.edu.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

opioid derived from morphine. Fatality reports on heroin reveal a constantly changing trend in overdoses geographically. Between 2017 and 2018, rates in heroin overdoses in the Midwest declined as rates in the West increased [2]. Trends also display shifting patterns in the use of prescription and synthetic opioids over time, with synthetics like Fentanyl on the rise [3]. As high rates of overdoses and opioid misuse persist, health departments are diligently working towards solutions. Strategies to mitigate the epidemic have involved a number of approaches including improving access to treatment, promoting overdose-reversing drugs, and refining public health surveillance [1]. Constantly evolving trends in drug use highlight the importance of closely monitoring patterns to better prepare for overdose response and to improve community treatment.

Monitoring and predicting opioid use to prevent overdoses has been a challenge for health departments across the nation. To improve public health surveillance and data quality, health departments have utilized emergency department data to report non-fatal overdoses and the State Unintentional Drug Overdose Reporting System (SUDORS) for fatal overdoses [4]. However, these data reports are reported in yearly increments, limiting the utilization to inform timely interventions. Acquiring timely reports continues to be a conspicuous obstacle in addressing the epidemic due to the coordinated effort required across local health departments, emergency departments, and substance use treatment providers. For example, one approach to tracking heroin usage in the United States is through government-issued surveys which may not provide credible real-time responses. Without accurate and real-time data, measures to intervene on this epidemic are delayed [6]. Innovative solutions to achieve more timely data are needed to adapt to the quickly changing pace of this crisis.

Social big data, including data from social media sites and search engines, have been utilized as a method of public health surveillance and may provide near real-time data to track trends and make predictions [7 – 11]. High-risk groups utilizing social media for interactions related to drug use, may be a reliable source for understanding and predicting opioid-related behaviors. [12 – 13] The use of social media to discuss the misuse of prescription opioids has shown promising correlations to actual substance use. Through comparing national survey data and geographical variations in social media, researchers have found connections in high-risk groups utilizing social media to discuss their opioid usage. Additionally, the use of internet search data in correlation to heroin-related ED visits also show promising ties. [12] Similarly, social big data from platforms such as Twitter, may be explored to extract real-time data on opioid-related ED visits [14]. For example, researchers have utilized big data from social media to determine behavioral patterns related to prescription drug abuse. Through exploring user characteristics as markers for patterns of substance use-related communications on social media, researchers may be able to better understand correlations between social media usage and drug abuse [15]. Furthermore, researchers have used Google search engine data to monitor a variety of public health issues and model predictions. Google Trends has been used to track search topics like “influenza”; suggesting that increases in the number of searches have promising correlations to the amount of influenza cases that the CDC reports [16]. Autoregression models constructed with google search data have become exceedingly accurate at estimating influenza patterns [17]. Subsequently, internet search trend data have maintained notoriety for maintaining promising exhibitions of accessibility and capacity to improve public health surveillance

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

[18]. Such methods have also been used in monitoring and predicting HIV outbreaks. For instance, search volumes in Google Trends relating to sexual-risk behaviors at the state level show promising correlations to reported HIV cases. Similarly, online social media platforms that yield geotagged data, such as Twitter, could potentially be utilized in monitoring drug-related conversations and sexual-risk behaviors. Utilizing big data that contains geotagged information can be helpful in explaining the differences in opioid overdoses and deaths by each region. Identifying regions in which patterns occur, can help inform treatment and outreach services on how to best intervene in a particular region. The purpose of surveillance through monitoring social media tweets would be to provide public health departments with more insight on potential increases in opioid usage on the internet and determining if these correlate to overdoses. Applying these methods to the opioid crisis may significantly expedite the timing of reports, allowing for quicker interventions from health departments among communities in need. Subsequently, big data might help expand on knowledge of overdose risks and increase the effectiveness of treatment by targeting specific regions. Innovatively, big data has the potential to provide low-cost public data that may serve as a surveillance tool in predicting future opioid-related hospitalizations and fatalities [19]. The purpose of this preliminary study is to seek to identify whether and how near real-time social media data might be used to track regional differences in opioid use and ways that opioids are discussed online.

METHODS

Twitter is a social media platform which allows users to microblog experiences and thoughts through posts called “tweets”. Using Twitter’s application programming interface (API), we retrieved tweets related to opioids (brand, generic, and street names) from March 17th to July 17th, 2020, including Oxycontin, Vicodin, and Morphine. Street names of opioids gathered from the 2018 Drug Enforcement Administration (DEA) Intelligence report were also included into the search query containing keywords such as “percs”, “roxis”, and “hillbilly heroin” [20]. To ensure the Twitter mentions collected pertained to drug usage, tweets were sampled for accuracy and compared to the DEA’s classifications. We required that opioid-related keywords be within 10 characters of verbs that implied drug use (e.g., “popped”, “taking”, and “using”). This search included tweets from geographical locations of diverse population sizes such as the states of New York and Iowa, the cities of Chicago, Illinois and Atlanta, Georgia, and the counties of Los Angeles, San Francisco, and Orange, California. Specific locations were utilized in order to encompass different regions in which street names for opioids may differ. Twitter users have the ability to display their tweets publicly or turn on privacy settings when posting a tweet. As a part of our search criteria, we requested only tweets that have geo-locations turned on and accounts that allow public view of their content. In our analysis, the results that were analyzed were from users who chose to publicly post their content with geo coordinates on. Tweets were then categorized as relevant or irrelevant, with irrelevant tweets being tweets that did not directly relate to opioid usage. Upon collecting the tweets that matched our search criteria, we evaluated the geo coordinates of each tweet and classified them based on location for comparison. Our analysis looked at the frequency in which a keyword was mentioned within each specific

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

location. The UC Irvine Institutional Review Board ruled this study as non-human subject research, as the data collected were publicly available.

RESULTS

After removing irrelevant and duplicate tweets (N = 22,400 tweets (10.7% removed)), the search identified 19,721 tweets. Of these results, the distribution across regions was the following: 29% New York, 23% Los Angeles, 18% Chicago, 18% Atlanta, 8% San Francisco, 3% Iowa, and 1% Orange County. We used an embedded feature from a software application to view the frequency of top themes relative to the prevalence in which they were mentioned. Upon assessing the themes across tweets that matched our search criteria, the data depicted variations in keywords across geographical locations. The deviations found through these themes suggested the potential for differences in the type of drugs that were mentioned and the methods of drug consumption.

Geographical locations showed perceived variations in mentions of natural and synthetic drugs. As illustrated in Table 1, tweets within Los Angeles did not contain mentions of the drug methadone. However, tweets in New York displayed a higher amount of mentions with methadone comprising 4.6% of mentions. In addition, both locations depicted perceived differences in the mention of opioids such as Vicodin, which yielded more mentions in Los Angeles as opposed to New York. Based on the data assessed from the themes, mentions of Vicodin comprised 3.4% of mentions, however the frequency of mentions in New York were too low to detect as a theme. Tweets in Orange County also had high mentions of Vicodin similar to Los Angeles but differing from the county of San Francisco. Furthermore, top themes in Iowa in comparison with other geographical locations showed high mentions of Vicodin (5.2%) similar to Orange County and San Francisco. Although San Francisco had lower mentions of Vicodin, drugs such as Methadone and Fentanyl exceeded those in Los Angeles and Orange County. Furthermore, the opioid tramadol had the highest mentions in Iowa in comparison to other locations.

Across geographical locations there were also perceived differences in the mention of words that implied drug consumption. As displayed in Table 2, average tweets in the West Coast within the counties of Los Angeles, Orange County, and San Francisco contained more mentions of the base word “shoot” as opposed to the average mentions in regions in the East Coast and Midwest. In San Francisco, tweets differed and had higher mentions of the term “needle”, which were absent in tweets from Los Angeles and Orange County. These results contest more perceived granular differences down to the county level. Meanwhile, tweets on the East Coast within New York contained higher mentions of key base words such as “take”. Chicago reflected similar patterns to New York containing the same keywords, but with additional phrases such as “pop”. Atlanta exhibited higher mentions of the base word “pop” indicative of potential opioid consumption. Based on the assessed results, geographical locations in the West Coast contained higher mentions of “shoot” than on the East Coast and Midwest. In addition, the top themes illustrated varying mentions of drug consumption methods down at the county level.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

DISCUSSION

Public health departments and researchers have faced challenges in monitoring opioid overdoses through public health surveillance reports. Reports reveal a substantial lag time in receiving information on deaths due to reports being generated in yearly increments [21]. Results from this exploratory analysis disclose the timeliness in which opioid data may be retrieved through social media. The data analyzed conveyed potential differences in the type of drugs and methods of drug consumption that were mentioned across different locations within the United States. We used software features that track mentions from social media, such as word clouds that display top themes, to assess changing trends in opioid usage. Twitter results observed from this search query displayed perceived differences between mentions in geographical locations on the East and West coast and more granular differences at the county level. Differences ranged from variations in the frequency in which an opioid was mentioned and the methods of consumption most commonly referenced. Identifying geographical variations in the populations of opioid users is especially important in monitoring usage patterns in rural regions that have previously not been well-documented.

As observed in Atlanta, mentions of methadone, often used as a medication to treat opioid use disorder, reveal regions that have an absence of methadone mentions. This suggests the potential for further exploration of the reason for the lack of mentions, for example, whether it would be related to reduced supply or reduced demand for this medication. This information may be used to allocate or increase funding of methadone maintenance treatment centers in regions where low mentions of methadone are present.

This approach to public health surveillance has limitations. The first of which is that we are unable to verify whether Twitter users who tweet phrases implying drug use are actual drug users. For instance, a twitter user who tweets “poppin percs” might not be a drug user but, for example, could be referencing a song lyric instead. In addition, our search query was run on a limited time frame from March 17th to July 17th, 2020. In order to adequately determine the accuracy of the patterns we assessed in this analysis a larger time frame would be needed. Furthermore, this analysis presents data from seven different regions, limiting the ability to generalize throughout the nation. Comparison between cities and states as described in our analysis, does not take into account the differences in homogeneity and sample size. In addition, data are limited to twitter users who have their geolocation on. Twitter data from opioid users who have their privacy settings set to restrict geolocation are not presented in this data. It is also important to note that comparisons across different types of opioids (e.g., opioids for pain versus medication treatments; illicit and prescription opioids) is a limitation to this study, as social media interactions vary depending on the type of opioid. Future research can better explore differences across types of opioids before implementing big data strategies for surveillance. Lastly, our analysis removed 10.7% of tweets that we classified as not related to opioid usage which may have biased the sampling.

In addition to providing near real time data, big data from social media has the potential to be a low-cost method of public health surveillance and may save health departments millions of dollars in spending [4]. Big data has the advantage of being easily accessible and could potentially contribute to the elimination of lag times in retrieving reports and distributing them to researchers and public health officials. Employing innovative approaches to the opioid crisis through the use of public social media data might ameliorate current issues faced in public health surveillance, such as lag times in reporting, and allow health departments to better prepare for opioid overdoses.

ACKNOWLEDGEMENTS

We thank the National Institutes of Health for funding support.

FUNDING AND SUPPORT

This work was supported by the NATIONAL INSTITUTE OF ALLERGIES AND INFECTIOUS DISEASES, under grant 5R01AI132030–05, as well as NATIONAL INSTITUTE ON DRUG ABUSE, NATIONAL INSTITUTE OF MENTAL HEALTH, and NATIONAL CENTER FOR COMPLEMENTARY AND INTERGRATIVE HEALTH.

REFERENCES

1. Abuse NI on D. Opioid Overdose Crisis https://www.drugabuse.gov/drug-topics/opioids/opioid-overdose-crisis (accessed Aug 15, 2020).
2. Heroin Overdose Data | Drug Overdose | CDC Injury Center https://www.cdc.gov/drugoverdose/data/heroin.html (accessed Aug 15, 2020).
3. Wilson N. Drug and Opioid-Involved Overdose Deaths — United States, 2017–2018. MMWR Morb Mortal Wkly Rep 2020, 69. 10.15585/mmwr.mm6911a4.
4. Enhanced State Opioid Overdose Surveillance | Drug Overdose | CDC Injury Center https://www.cdc.gov/drugoverdose/foa/state-opioid-mm.html (accessed Aug 15, 2020).
5. Chary M; Genes N; Giraud-Carrier C; Hanson C; Nelson LS; Manini AF; Epidemiology from Tweets: Estimating Misuse of Prescription Opioids in the USA from Social Media. Journal of Medical Toxicology, 13, 278–286(2017). 10.1007/s13181-017-0625-5. (accessed Dec 11, 2020). [PubMed: 28831738]
6. Kilmer B; Caulkins JP Hard Drugs Demand Solid Understanding https://www.rand.org/blog/2014/03/hard-drugs-demand-solid-understanding.html (accessed Aug 15, 2020).
7. Young SD; Heinzerling K; The Harnessing Online Peer Education (HOPE) intervention for reducing prescription drug abuse: A qualitative study. Journal of Substance Use 2017, 592–596. 10.1080/14659891.2016.1271039. [PubMed: 29551953]
8. Young SD (2015). A “big data” approach to HIV epidemiology and prevention. Preventive Medicine, 70, 17–18. 10.1016/j.ypmed.2014.11.002 [PubMed: 25449693]
9. Young SD, Yu W, & Wang W. (2017). Toward Automating HIV Identification: Machine Learning for Rapid Identification of HIV-Related Social Media Data. JAIDS Journal of Acquired Immune Deficiency Syndromes, 74, S128–S131. 10.1097/QAI.0000000000001240 [PubMed: 28079723]
10. Young SD, Torrone EA, Urata J, & Aral SO (2018). Using Search Engine Data as a Tool to Predict Syphilis: Epidemiology, 29(4), 574–578. 10.1097/EDE.0000000000000836 [PubMed: 29864105]
11. Garett R, Liu S, & Young SD (2018). The relationship between social media use and sleep quality among undergraduate students. Information, Communication & Society, 21(2), 163–173. 10.1080/1369118X.2016.1266374
12. Young SD; Koussa M; Lee S; Perez H; Gill N; Gelberg L; Heinzerling K; Feasibility of a social media/online community support group intervention among chronic pain patients on opioid therapy. Journal of Addictive Diseases 2019, 37(8):1–6. 10.1080/10550887.2018.1557992.
13. Young SD, Lee S-J, Perez H, Gill N, Gelberg L, & Heinzerling K. (2020). Social media as an emerging tool for reducing prescription opioid misuse risk factors. Heliyon, 6(3), e03471. 10.1016/j.heliyon.2020.e03471
14. Young SD; Zheng K; Chu LF; Humphreys K. Internet Searches for Opioids Predict Future Emergency Department Heroin Admissions. Drug and Alcohol Dependence 2018, 190, 166–169. 10.1016/j.drugalcdep.2018.05.009. [PubMed: 30036853]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

15. Kim JS; Marsch LA; Hancock JT; Das AKScaling Up Research on Drug Abuse and Addiction Through Social Media Big Data. https://www.jmir.org/2017/10/e353/ (accessed Dec 3, 2020).
16. Cook S; Conrad C; Fowlkes AL; Mohebbi MH Assessing Google Flu Trends Performance in the United States during the 2009 Influenza Virus A (H1N1) Pandemic. PLOS ONE 2011, 6 (8), e23610. 10.1371/journal.pone.0023610.
17. Yang S; Santillana M; Kou SC Accurate Estimation of Influenza Epidemics Using Google Search Data via ARGO. PNAS 2015, 112 (47), 14473–14478. 10.1073/pnas.1515373112.
18. Milinovich GJ; Avril SMR; Clements ACA; Brownstein JS; Tong S; Hu W. Using Internet Search Queries for Infectious Disease Surveillance: Screening Diseases for Suitability. BMC Infectious Diseases 2014, 14(1), 690. 10.1186/s12879-014-0690-1. [PubMed: 25551277]
19. Young SD; Zhang Q. Using Search Engine Big Data for Predicting New HIV Diagnoses. PLOS ONE 2018, 13(7), e0199527. 10.1371/journal.pone.0199527.
20. Slang Terms and Code Words: A Reference for Law Enforcement Personnel | Drug Enforcement Administration (DEA). Published July 2018. (accessed Aug 15, 2020).
21. CA Opioid Dashboard https://skylab.cdph.ca.gov/ODdash/ (accessed Aug 15, 2020).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1.

Frequency of medication names and slang terms, N = 19,721.

	New York	Los Angeles County, California	Chicago, Illinois	Atlanta, Georgia	San Francisco County, California	Iowa	Orange County, California
Themes	***Total (n= 5,776)***	***Total (n = 4,556)***	***Total (n = 3,594)***	***Total (n = 3,573)***	***Total (n = 1,642)***	***Total (n = 580)***	***Total (n = 165)***
Methadone	4.7%	0.0%	0.0%	1.3%	6.0%	0.00%	2.4%
Fentanyl	16.7%	13.3%	15.9%	10.0%	34.8%	16.5%	20.0%
Percs	8.7%	11.1%	13.6%	26.2%	6.0%	3.1%	2.4%
Vicodin	0.0%	3.4%	2.1%	0.8%	0.0%	5.3%	5.4%
Tramadol	0.0%	0.0%	0.0%	0.0%	0.0%	3.2%	0.0%

Note. Percentages do not aggregate to 100%, key terms may appear more than once in the same tweet.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 2.

Base words associated with methods of drug consumption. N = 19,721.

Themes	New York	Los Angeles County, California	Chicago, Illinois	Atlanta, Georgia	San Francisco County, California	Iowa	Orange County, California
	Total (n = 5,776)	***Total (n = 4,556)***	***Total (n = 3,594)***	***Total (n = 3,573)***	***Total (n = 1,642)***	***Total (n = 580)***	***Total (n = 165)***
pop	0.0%	0.0%	1.6%	4.3%	0.0%	0.0%	0.0%
take	1.8%	1.4%	1.4%	0.0%	0.0%	0.0%	0.0%
shoot	0.9%	1.7%	1.2%	0.0%	3.2%	0.0%	1.2%
needle	1.2%	0.0%	0.0%	0.0%	1.5%	1.03%	0.0%

Note. Percentages do not aggregate to 100%, key terms may appear more than once in the same tweet. Percentages displayed include past, present, and future tenses of the base words presented.