Distribution Agreement

In presenting this Thesis as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my Thesis in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this Thesis. I retain all ownership rights to the copyright of the Thesis. I also retain the right to use in future works (such as articles or books) all or part of this Thesis.

Kirsten Yates Date

AN EXAMINATION OF THE PORTRAYAL OF
PRESCRIPTION OPIOID USE AND MISUSE
ON TWITTER

BY

Kirsten M. Yates
Degree to be awarded: M.P.H.
Executive MPH

Rita Noonan, PhD | Date
Committee Chair

Sara Lewis, MPH | Date
Field Advisor

Melissa Alperin, MPH, CHES | Date
Chair, Executive MPH Program

AN EXAMINATION OF THE PORTRAYAL OF
PRESCRIPTION OPIOID USE AND MISUSE
ON TWITTER

BY

Kirsten M. Yates
M.P.H., Emory University, 2015
B.A., Vanderbilt University, 2007

Thesis Committee Chair: Rita Noonan, PhD

An abstract of
A Thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements of the degree of
Master of Public Health in the Executive MPH program
2015

Abstract

AN EXAMINATION OF THE PORTRAYAL OF
PRESCRIPTION OPIOID USE AND MISUSE
ON TWITTER

BY
Kirsten M. Yates

Deaths resulting from drug overdoses are the leading cause of injury death in the United States and are due in part to nonmedical use or misuse of prescription opioids. While strategies exist to combat this epidemic, current monitoring and surveillance systems in place do not provide real-time data. Research shows that substantial changes have taken place over time in the geographic distribution of opioid-related mortality. A need exists for a public health platform to help public health workers predict where in the United States a rise in misuse or overdose may be occurring, so that evidence-informed strategies may be implemented as quickly as possible.

Methods: This study analyzed how Twitter users describe prescription opioid use and misuse on Twitter, whether there were discernible state-specific trends in the quantity of tweets, and whether the variations in tweets by state reflected variations in estimated prescribing rates, overdose rates, and nonmedical use rates.

A search was conducted for all tweets mentioning keywords related to prescription opioids during a six-month period. The tweets were filtered to include only those originating from individuals discussing individual prescription opioid use at least five times, indicating potential for misuse. The tweets were sorted by user-identified geographic location and qualitatively analyzed for content indicative of misuse according to four categories: overdose, dependence, co-ingestion, and seeking. Results were compared to state-specific prescribing rates, nonmedical use rates, and overdose rates.

Results: Results demonstrated that Twitter users are engaging in conversation about prescription opioid use (n=809), and the majority (76.9%) mention behaviors that are indicative of misuse. A modest, statistically significant correlation (r=0.303) with state-specific estimates of nonmedical use of prescription pain relievers was found, if one data point considered an outlier was removed. No statistically significant correlation was found for estimated prescribing and overdose rates.

Conclusion: The findings indicate that individuals are discussing their prescription opioid use on Twitter, and there is a wealth of information available to analyze within the content of these discussions. Further research is needed to determine the potential to use Twitter as an early predictor of prescription opioid misuse and overdose.

AN EXAMINATION OF THE PORTRAYAL OF
PRESCRIPTION OPIOID USE AND MISUSE
ON TWITTER

BY

Kirsten M. Yates
M.P.H., Emory University, 2015
B.A., Vanderbilt University, 2007

Thesis Committee Chair: Rita Noonan, PhD

A Thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements of the degree of
Master of Public Health in the Executive MPH program
2015

ACKNOWLEDGEMENTS

I would like to thank my thesis committee chair, Dr. Rita Noonan, and my field advisor, Sarah Lewis, for their time, assistance, and encouragement throughout the conduct of this project.

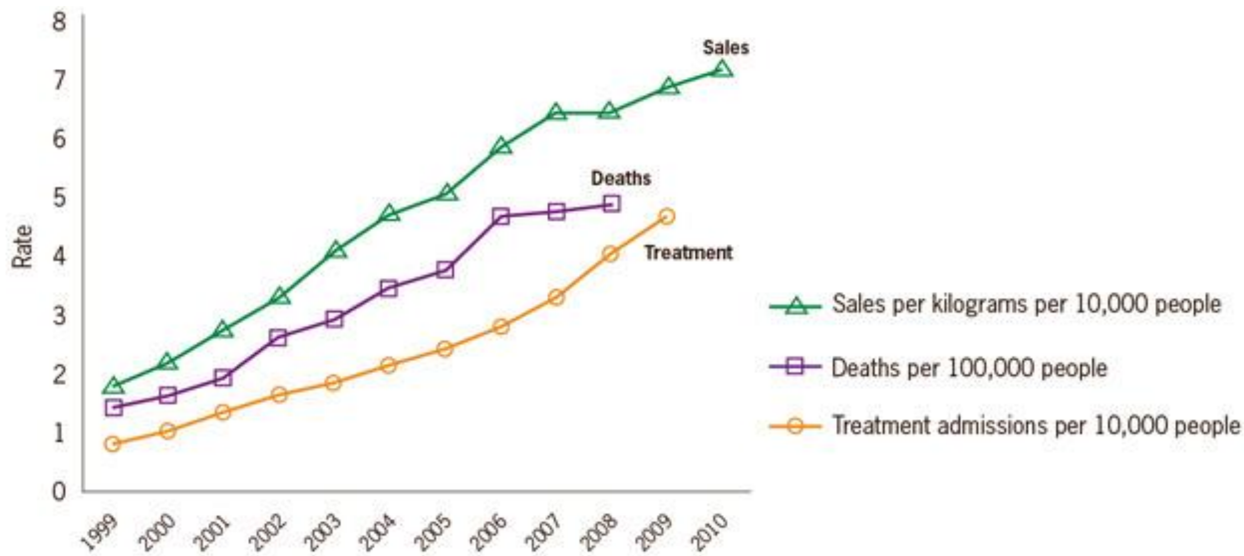
TABLE OF CONTENTS

Chapter I – Introduction and Rationale

I. Introduction

Deaths resulting from drug overdoses have become the leading cause of injury death in the United States (CDC, 2014). Unintentional overdoses from prescription opioids in particular have reached epidemic levels and have caused more deaths than overdoses from heroin and cocaine combined (CDC, 2011).

Figure 1: Rates of prescription painkiller sales, deaths and substance abuse treatment admissions (1999-2010)
Source: Centers for Disease Control and Prevention, 2012



Prescription painkiller overdoses killed more than 16,000 people in the United States in 2013 (CDC, 2015). For every unintentional overdose death related to an opioid analgesic:

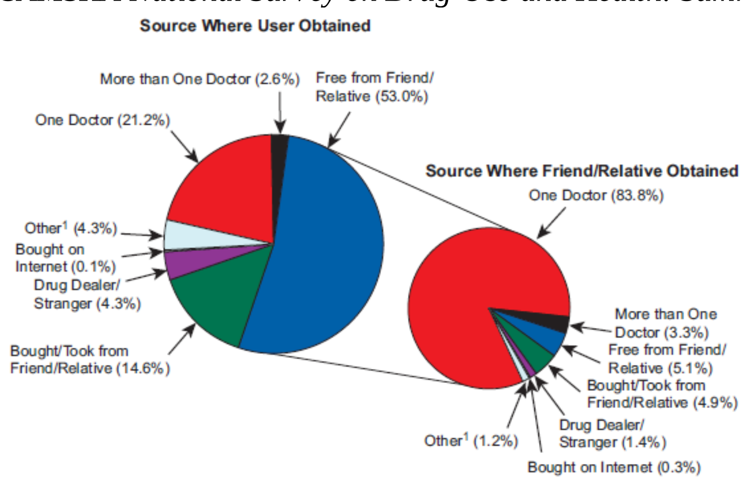
- Nine people are admitted for substance abuse treatment;
- 35 visit emergency departments;
- 161 report drug abuse or dependence; and
- 461 report nonmedical use of opioid analgesics (CDC, 2012; King et al., 2014).

This problem costs the United States about $55 billion annually, according to the most recent estimates. Of this amount, 46% was attributable to workplace costs such as lost productivity, 45% to healthcare costs including substance abuse treatment, and 9% to criminal justice costs (Birnbaum et al., 2011).

Problematic prescribing practices are a leading contributor to epidemic. Safe and informed prescribing practices and the following sensible prescribing guidelines can help stop it (CDC, 2015). Additionally, the increasing numbers of emergency department visits and deaths from overdoses have been attributed in part to an increase in the nonmedical use of prescription drugs (ACOG, 2012). Individuals may misuse prescription drugs by using them intentionally without a prescription, in a way other than as prescribed, or for the feeling that it causes.

In 2013, there were an estimated 6.5 million nonmedical users of prescription drugs, 4.5 million of whom used prescription pain relievers. Of those who used pain relievers nonmedically in the past 12 months who were aged 12 years or older, 53% got the drug they used most recently from a friend or relative for free, and 10.6% bought the drug from a friend or relative. Another 21.2% reported that they got the drug through a prescription from one doctor. (SAMHSA, 2014).

Figure 2: Source Where Pain Relievers Were Obtained for Most Recent Nonmedical Use among Past Year Users Aged 12 or Older: 2012-2013
Source: 2013 SAMSHA National Survey on Drug Use and Health: Summary of Findings



[1]The Other category includes the sources "Wrote Fake Prescription," "Stole from Doctor's Office/Clinic/Hospital/Pharmacy," and "Some Other Way."

Note: The percentages do not add to 100 percent due to rounding.

Men are more likely to abuse prescription drugs than women, and estimates suggest that abuse is highest among young adults aged 18 to 25 (NIH, 2014b).

II. Problem Statement

The Centers for Disease Control and Prevention (CDC) has published recommended evidence-informed prevention strategies to reduce prescription drug overdose. They comprise insurance restrictions to prevent "doctor shopping," improving legislation and enforcement of existing laws against "doctor shopping," improving prescribing practices, and secondary and tertiary prevention strategies, such as harm reduction and substance abuse treatment programs (CDC, 2012). Specific recommendations include implementation of:

- **Prescription Drug Monitoring Programs (PDMPs):** Higher prescribing of opioids is associated with more overdose deaths. PDMPs are electronic databases maintained by states that can track the prescribing and dispensing of controlled prescription drugs. This information helps officials identify patients at highest risk for abuse and overdose in terms of dosage, number of prescriptions, and number of prescribers. It also helps officials identify prescribers who are deviating from accepted medical practice by high opioid prescribing rates and high number of identified "doctor shoppers" as patients. This is particularly important, as CDC has found that an increase in painkiller prescribing is a key driver of the increase in prescription overdoses.

- **Patient Review and Restriction Programs:** CDC recommends that states implement programs that will allow them to monitor prescription claims information and PDMP data (if available) for signs of inappropriate use of controlled prescription drugs. If the review indicates a patient is using multiple providers for controlled prescription drugs and the use of multiple providers cannot be justified on medical grounds, then states could consider only agreeing to reimburse claims from one designated physician and pharmacy. This can help promote appropriate access and care for patients at risk for overdose.

- **Legislation to Prevent Abuse and Diversion:** CDC recommends that states enact and enforce laws to prevent "doctor shopping" and "pill mills," as well as laws that will reduce diversion and abuse. CDC recommends laws be evaluated to ensure they safeguard patients who have a legitimate need for prescription opioids for pain management.

- **Accessible Substance Abuse Treatment Programs:** CDC recommends that states promote accessible substance abuse treatment programs, which can help prevent overdose in those struggling with addiction (CDC, 2013).

Figure 3: Cumulative number of states authorizing prescription drug abuse-related laws by type of law, United States, 1970-2010
Source: Centers for Disease Control and Prevention, 2014

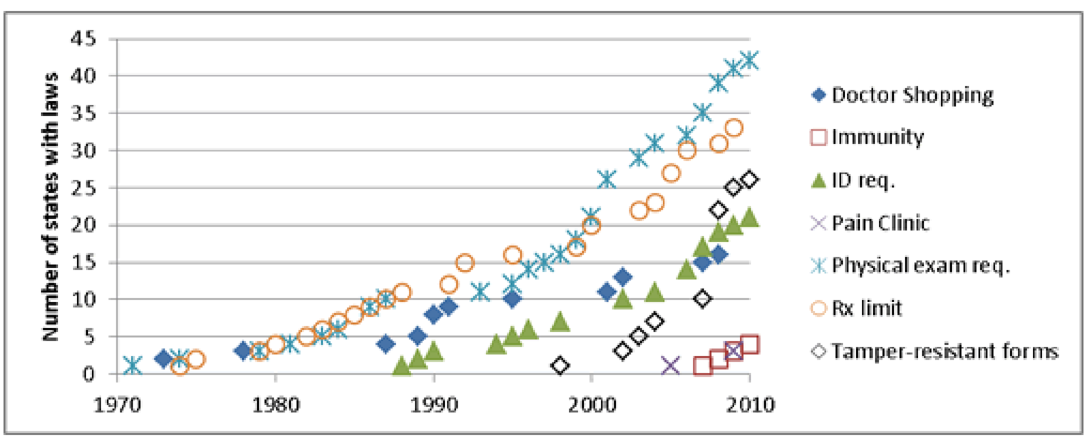


Figure 4. State Successes: Reduced Opioid Prescribing Results
Source: NY, TN: PDMP Center of Excellence at Brandeis University, 2014. FL: Vital Signs Morbidity and Mortality Weekly Report, July 1, 2014.

For such strategies to have the greatest public health impact, it is imperative to understand where the greatest need exists. Public health practitioners continue to operate with limited resources, and funding to implement prevention strategies is often based on where the problem is most severe, according to the available data. However, researchers have noted a number of challenges to obtaining accurate and timely data on prescription opioid misuse and overdose.

Attempting to identify the extent of diversion and "doctor shopping" in a given geographic area is particularly difficult because it generally requires interviewing individuals or proxies. Additionally, determining the exact cause of death in an overdose is difficult, as there is some inconsistency in *International Classification of Diseases* codes for drug poisoning, as well as a lack of standardization in drug categorization and terminology in coroners' reports (King et al., 2014). Furthermore, documented research shows substantial changes over time in the geographic distribution of opioid-related mortality (Paulozzi & Xi, 2008). For example, in 1999, large metropolitan areas had the highest rates of opioid-related deaths, and non-metropolitan areas had the lowest rates; however, by 2004, non-metropolitan areas had the highest rates and had seen the largest relative increase during that time period (Paulozzi & Xi, 2008). However, available incidence data is, at best, two to three years behind the present day. The most recent surveillance data available to demonstrate the current distribution of opioid-related mortality is from 2010 (CDC, 2012). It could take multiple years before surveillance data on new prescription opioid overdose deaths in a certain region are released, making it difficult to rapidly respond to new overdose and misuse "hot spots."

The availability of real-time data to serve as an indicator for a potential uptick in opioid abuse in a specific geographic area would help public health practitioners allocate resources more effectively and respond more efficiently to this growing issue.

III. Purpose Statement

This study examined the extent to which content of Twitter conversation (or "tweets") related to prescription opioid use reflected larger behavioral misuse patterns and evidence-informed prevention strategies, as documented by prescribing and mortality data available from CDC, as well as prevalence estimates of nonmedical use available from the Substance Abuse and Mental Health Services Agency (SAMHSA). The specific research questions this study attempted to answer were:

1. How do Twitter users describe prescription opioid misuse? What language, phrases, and references are made with respect to opioid misuse?
2. Are there discernible state-specific trends in tweets about prescription opioid misuse? If so, do these trends reflect documented behavioral patterns in those states?
3. Do variations in tweets by state reflect the variations in prescribing rates by state?

IV. Theoretical Framework

This study is based on the social exchange theory, which posits that social behavior is based on an individual's cost-benefit assessment of participation in a social exchange. Specifically, an individual enters into a social exchange with another individual because of the returns he or she expects to receive from that exchange (Nord, 1969). These returns can be in the form of social rewards, such as opportunity, prestige, conformity, or acceptance. For example, people may engage in social exchanges for an expected gain in reputation and influence or an expected reciprocity on the part of others (Nord, 1969; Pan & Crotts, 2014).

The elements of the social exchange theory are played out in social media platforms, which provide online channels for individuals to build relationships and engage in social exchanges. These social exchanges may influence attitudes and behaviors (Hanson et al., 2013). Indeed, social media applications, such as Twitter and Facebook, are platforms for people to meet and network. Individuals often use these platforms to share information about themselves and find others who are like-minded and who will reinforce their own attitudes and behaviors. Their desire to exchange information in this manner, particularly surrounding prescription opioid use, serves as the basis for this study.

V. Significance Statement

Social media platforms have become novel sources of health-related data, providing public health practitioners with real-time data on a range of health issues that span from influenza infection "hot spots" to natural disasters to postpartum depression (Hill et al., 2013). Twitter, in particular, is becoming a more common tool to use to mine public health data, due to its immediacy and the fact that by default, "tweets" are public.

While recent studies have started exploring the use of Twitter as a surveillance tool for prescription drug misuse in general, to the best of my knowledge, no research has focused on prescription opioid misuse, particularly as it relates to documented geographic surveillance patterns. If results of this study indicate that Twitter trends related to discussion about prescription opioid use and misuse may reflect actual trends occurring in certain geographic areas, public health officials may be able to use Twitter as a predictor of misuse activity, allowing them to intervene quickly and potentially save lives by reducing deaths from overdose.

VI. Definition of Terms

The following terms used in this study will convey the definitions outlined below throughout the entirety of this document.

- **Misuse:** Individuals who misuse prescription drugs use them intentionally without a prescription, in a way other than as prescribed, or for the experience or feeling that it causes.

- **Nonmedical use**: Nonmedical use of prescription drugs involves the taking of prescription drugs, whether obtained by prescription or otherwise, other than in the manner or for the reasons or time period prescribed, or by a person for whom the drug was not prescribed.

- **Prescription Opioid:** A type of medication used to relieve pain. Medications that fall within this class include hydrocodone (e.g., Vicodin), oxycodone (e.g., OxyContin, Percocet), morphine (e.g., Kadian, Avinza), codeine, and related drugs.

- **Prescription Painkiller**: The term "prescription painkillers," as used by the Centers for Disease Control and Prevention (CDC) and referenced in this study when citing their data, refers to opioid or narcotic pain relievers, including drugs such as Vicodin (hydrocodone), OxyContin (oxycodone), Opana (oxymorphone), and methadone.

- **Prescription Pain Reliever:** The term "prescription pain relievers," as used by the Substance Abuse and Mental Health Services Administration and referenced in this study when citing their data, refers to opioid or narcotic pain relievers, including Vicodin, Lortab, Norco, Zohydro ER, Hydrocodone, OxyContin, Percocet, Percodan, Roxicet, Roxicodone, Oxycodone, codeine, and morphine.

- **Overdose Rate:** The rate of individuals ingesting a lethal amount of a prescription drug, resulting in death.

- **Self-identified location (as provided by Twitter users):** The geographic location a Twitter user has entered as part of his or her user profile.

- **Tweet:** A 140-character message posted by an individual user on Twitter.

Chapter II – Literature Review

I. Introduction

This chapter will review the current literature related to the use of Twitter as tool to monitor and predict public health issues. Firstly, it is important to review general research and summaries exploring whether Twitter is a platform that can be used for public health surveillance with validity (and reliability). This requires consideration of potential or documented limitations related to using it for this purpose. It is also important to review available research on the feasibility of using Twitter to accurately track potential public health issues in the United States by state, using available Twitter location data.

Secondly, this chapter will review the results of studies that have analyzed tweets related to specific public health issues, ranging from chronic issues such as obesity to acute issues such as influenza. Finally, this review will move into an analysis of research that specifically examines the portrayal of prescription drug misuse on Twitter.

II. The Potential of Twitter Data for Public Health Analysis

With the growing acceptance and use of social media platforms as mechanisms for networking and social interaction, interest into their potential use for public health purposes has increased greatly over the past five years. Researchers are attempting to understand what public health information can be learned from Twitter in particular, given its public status, the immediate availability of the content being shared, and users' tendencies to disclose information about their emotional and physical well-being.

Shawndra Hill et al. reviewed the strengths and limitations of the use of Twitter for public health surveillance in an article published in 2013 in *Big Data*. The authors note that the platform shows promise as a public health surveillance tool because of its immense scale, the opportunity for the content of tweets to be systematically searched, and its immediacy. For example, Twitter's immediacy has "permitted real-time assistance in the case of natural disasters," by helping public health officials understand where the greatest needs are for those most impacted. Additionally, the authors demonstrate the usefulness of Twitter's immediacy by referencing the 2013 Boston Marathon bombings. Hill et al. note that emergency departments in Boston learned of the bombings through Twitter *before* being informed through more conventional sources.

Hill et al. also point out various limitations in using Twitter for public health surveillance. These limitations include:

- Ethical issues related to expectations of privacy and duty to intervene if a situation appears dire;
- The lack of context given in relation to the data, as the tweets do not divulge health history or medical outcomes, and casual claims about specific behaviors can be difficult to substantiate;
- The lack of distinction between genuine and spurious information;
- Concerns about the representativeness of Twitter users as sample populations; and
- The ambiguity of search words, as people mention diseases without necessarily experiencing them.

The authors conclude that more information is needed to determine whether or not Twitter can serve as an accurate surveillance tool for some public health issues. Researchers should be careful to refine data mining strategies to correct for the biases and issues outlined above (Hill et al., 2013).

During a 2011 conference on social media, Michael J. Paul and Mark Dredze of Johns Hopkins University presented a study that also attempted to determine the different ways that

Twitter could potentially impact research on a range of public health issues. They posit that Twitter can become a source of geographic syndromic surveillance for specific public health issues, which will help public health officials improve medical resource allocation, health policy, and education (Paul & Dredze, 2011).

To test this hypothesis, Paul and Dredze apply their Ailment Topic Aspect Model (ATAM) to analyze disease information from tweets. ATAM models how users express their illnesses and ailments in tweets, assuming that for each health-related tweet there is a latent ailment identified using a list of key phrases. The key phrases selected for the model were identified from articles for the general public that were written about specific ailments.

When applied nationally, results demonstrated that the probability of a mention of the chosen ailment (influenza in this case) on Twitter for any given week between August 2009 and October 2010 correlated with the influenza rate in the United States during that week, as measured by CDC. When applied geographically, results demonstrated that the probability of a mention of the chosen ailment (seasonal allergy was chosen for this comparison) on Twitter in a specific geographic region between August 2009 and October 2010 correlated with already known seasonal patterns of allergy.

These results suggest that tweets about some health ailments correlate with public health data according to quantitative and geographic analyses. Using this knowledge, we may consider that Twitter has the potential to become a new informatics tool for geographic syndromic surveillance. However, the authors point out some limitations. The capability to analyze tweets geographically was limited, as only 12% of the tweets included in the study indicated a geographic location. More sophisticated geocoding techniques could address this issue. The authors also note, similar to Hill et al. (2013), that Twitter user demographics will limit research, as Twitter users tend to skew younger, which does not provide researchers with a nationally representative sample (Paul & Dredze, 2011).

Scott Burton et al., in a study published in the *Journal of Medical Internet Research* in 2012, looked further into the accuracy of Twitter location information. The objective of the study was to determine what proportion of Twitter user accounts contain verifiable geographic information. The authors acknowledge that Twitter location information is not well understood, and that this has limited its public health utility.

Twitter users who wish to provide location information can do so in four different ways:

- They can share exact GPS coordinates associated with a tweet when tweeting from smartphones or other GPS-enabled devices;
- They can share GPS coordinates of a place (e.g. a city) when tweeting from their computers or other devices that do not have GPS services;
- They can add their location to the location information listed in their public profile descriptions; or
- They can share information about their general geographic location according to their time zone.

The first two options utilizing GPS location information requires opting in by the user (the setting is disabled by default). The last option related to time zones occurs automatically, because Twitter automatically infers a time zone when an account is created.

The authors note limitations to each option. While the GPS information may provide the most reliable data about a Twitter user's location, the setting is automatically disabled and the majority of users do not opt in (in this study, only 2.02% of the tweets analyzed contained GPS information). The user-supplied location information in the profile allows users to enter their locations in any format, leaving room for spelling errors or humorous/sarcastic entries in place of actual locations. Finally, the time zone inferred by Twitter only provides information by time zone region, which doesn't allow for state- or city-specific data.

Despite these limitations, the study authors found that the location indicators offered in

Twitter, when taken together, can offer “a sizable sample of individuals whose location can be accurately inferred.” Of particular note, results of the study found that user-supplied location information matched available GPS data in 87.69% of cases (in the United States). When combining these results with the percentage of Twitter users who supply location information in their profiles, one could reliably use location information for between 15.35% and 17.13% of all Twitter users. Burton et al. conclude, “whereas it was beyond the purview of the current study to assess the validity of social media data, for public health researchers and communicators to dismiss such data sources without further consideration would be premature because it may miss an opportunity to observe, reach, and communicate with people in unprecedented ways” (Burton et al., 2012).

Connie St. Louis and Gozde Zorlu further sum up the potential for Twitter to be used for public health surveillance in their feature in the *British Medical Journal* in 2012, in which they interview a group of public health leaders from reputable agencies around the world. They note that traditional methods of surveillance, while very accurate, can take a long time—and time is critical when trying to prevent spread of a disease. Many public health workers are beginning to acknowledge the growing number of informal sources of information—like Twitter—that can provide a much faster picture of serious public health issues.

St. Louis and Zorlu cite recent examples of Twitter’s potential utility in predicting disease outbreaks. For example, an analysis of tweets from 2009 demonstrated that the 2009 H1N1 flu outbreak could have been identified via Twitter one week before it emerged in general records from health care provider reports (Szomszor et al., 2010). Feedback from the public health officials interviewed focuses on the usefulness of this kind of speed in outbreak detection, as an extra week or two can be incredibly important in preparing resources for a response and allowing faster communication and more timely intervention with the public. .

However, these same individuals note that tools like Twitter cannot and will not replace traditional surveillance systems, as neither option can "do it perfectly." The "background noise" posted on platforms like Twitter can dilute the valid health information, and more models are needed to filter and validate the data from the noise (St. Louis & Zorlu, 2012).

III. Twitter Data Analyses Related to Specific Public Health Issues

Researchers are beginning to examine the validity of Twitter as a surveillance and data-mining tool for specific public health issues, ranging from chronic diseases to mental health issues to infectious disease outbreaks. The body of literature on the use of Twitter for individual ailments is growing every day, which is summarized below.

Many of the articles and studies mentioned use of the platform to predict or track influenza outbreaks. Alessio Signorini, Alberto Maria Segre, and Philip M. Polgreen examined the use of Twitter specifically during the H1N1 pandemic in the United States to track levels of disease activity. The results of their study were published in *PLoS ONE* in 2011.

Signorini et al. collected and stored a sample of tweets according to a set of pre-specified search terms related to H1N1 that suggested influenza-like illness (ILI). The tweets included a time stamp, as well as the author's self-declared home location (if provided in the user's profile). Tweets that were tagged as originating outside of the United States, tweets from users with a non-U.S. time zone, and tweets not written in English were excluded. Tweets with less than 5 characters and tweets sent from a client identifying as "API" were also excluded (those identified as "API" are typically generated by a computer and likely "spam").

The authors used the resulting tweets to 1) determine the accuracy of national quantitative estimates of ILI values, using CDC-reported ILI values across the United States as the standard; and 2) determine the accuracy of real-time estimates of ILI activity in a single CDC region by comparing tweets from users providing a GPS-location via Twitter to CDC region ILI estimates. The results demonstrated that the estimates calculated using the sample tweets were fairly accurate when compared to CDC-reported values. The Twitter estimates of national ILI values had an average error of 0.28%, with a standard deviation of 0.23%. The regional model was somewhat less accurate, with an average error of 0.37% and a standard deviation of 0.26%, which could be a result of the significantly smaller sample of tweets. Regardless, the results demonstrated that Twitter could be used to estimate influenza disease activity in real time, which is 1-2 weeks faster than current practice allows.

Justin C. Bosley et al. examined the use of Twitter to identify trends related to a noninfectious acute health issue: cardiac arrest. Different from the study conducted by Signorini et al. outlined above, this study did not attempt to track cardiac arrest rates in real time; instead, researchers were interested in analyzing the content, dissemination, and temporal trends of tweets to identify knowledge deficiencies and understand how the issue was being discussed online. Results of the study were published in *Resuscitation* in 2012.

Similar to Signorini et al., Bosley et al. obtained a sample of tweets by filtering them according to a specific date range (April 19 – May 26, 2011) and according to specific key words (cardiac arrest, CPR, AED, resuscitation, heart arrest, sudden death, and defib). The tweets were then reviewed and categorized as either related or unrelated to cardiac arrest/resuscitation. Tweets considered to be related were further categorized as (1) cardiac arrest [personal or information sharing]; (2) CPR [personal or information sharing]; (3) AED [personal or information sharing]; (4) cardiac arrest/CPR/AED [information seeking]; or (5) resuscitation education/research/news media.

Results demonstrated that the general public is using Twitter to both seek and share information about cardiac arrest and resuscitation. Of the tweets referencing cardiac arrest, 29%

referenced personal sharing, and 68% of those referencing CPR/AED represented personal sharing. Of all the tweets analyzed, 2% (270) were categorized as information seeking. Bosley et al. note that the information seeking tweets may be used by cardiac health organizations and agencies to understand gaps in knowledge among individuals and pave the way for further outreach (Bosley et al., 2012).

Debarchana Ghosh and Rajarshi Guha also used a noninfectious public health issue—obesity—to examine whether topic and spatial modeling could accurately identify a relevant health issue on Twitter. The results of their study were published in *Cartography and Geographic Information Science* in 2013.

Ghosh & Guha aimed to determine how to effectively identify, mine, and spatially analyze public health related topics discussed in tweets. They gathered a geographically representative sample of tweets using Twitter's optional GPS location information. Obesity-related search terms were also chosen based on knowledge of obesity and its risk factors in the United States. The researchers displayed the tweets via point density maps to show the spatial distribution of obesity-related tweets for all search terms, as well as the terms 1) "obesity"; 2) "childhood AND obesity"; and 3) "McDonalds AND obesity." These terms were developed to indicate common obesity-related issues and risk factors, such as childhood obesity and food deserts. These search terms were also used to generate topic models, which allowed for the algorithmic identification of topics within a collection of words. The topic models identified were used to group the individual tweets based on the words in the tweet.

The results demonstrated a higher density of tweets around bigger cities, which is likely because people from cities have higher rates of using Twitter from GPS-enabled devices, such as smartphones. When the researchers evaluated the statistical correlation between obesity prevalence among U.S. adults and the percentage of obesity-related tweets at the state level, the correlation was negative. However, Ghosh and Guha note that this could have been expected, as Twitter users are typically educated and likely aware of the growing concern of the rising rates of obesity. Several studies have shown that obesity is negatively correlated with education, and many of the obesity-related tweets revolve around forms of prevention and risk factors.

The topic modeling resulted in the identification of three primary themes related to how Twitter users discuss obesity-related topics: childhood obesity and schools, obesity prevention, and obesity and food habits. These themes were also spatially distributed in point density maps to help the researchers understand where the conversations around each theme were taking place. Ghosh and Guha state that results suggest, "the combined use of topic modeling and [geocoding] demonstrated the potential to extract themes from large datasets of conversational or textual data," such as data from tweets (Ghosh & Guha, 2013).

IV. Twitter Data Analyses Related to Prescription Drug Misuse

As the knowledge supporting Twitter's potential as a public health analysis platform grows, researchers are beginning to examine its ability to help public health officials better understand, prevent, and treat substance abuse. Indeed, the National Institutes of Health (NIH) just awarded more than $11 million over the next three years to support research exploring the use of social media to advance the scientific understanding, prevention, and treatment of substance use and addiction. The NIH press release announcing the awards on October 17, 2014, states, "researchers can analyze social media interactions to gain insights into patterns of use, risk factors, and behaviors associated with substance abuse…social media may also enhance screening, prevention, and treatment of substance use and addiction" (NIH, 2014b). To date, to the best of my knowledge, only a small handful of peer-reviewed, published studies have analyzed Twitter data in relation to prescription drug misuse in particular.

Carl Hanson et al. published the results of a novel study that examined Twitter data for evidence of nonmedical use of Adderall among college students in the *Journal of Medical Internet Research* in April 2013. The research questions that guided the study were:

1. When do Twitter users typically tweet about Adderall?
2. To what extent do tweets about Adderall abuse differ among various college and university clusters in the United States?
3. What, if any, substances do Twitter users tweet about commonly abusing in combination with Adderall?
4. What common side effects are mentioned?

To address these questions, Carl Lee Hanson et al. collected all tweets between November 29, 2011, to May 31, 2012, that included the keyword “Adderall.” Tweets from users whose screen-names included “Adderall” or “pharm” were excluded because they were not representative of typical users. Tweets were further categorized according additional keywords that indicated co-ingestion with other drugs, alternative motives, and possible side effects.

Of the tweets included in the sample, 2.8% provided GPS data. Hanson et al. further analyzed these tweets to determine those who were most likely to be students by searching them for student-related terms, such as “homework” and “class.” The geolocated tweets were then fit to pre-established clusters of colleges with a student population of 10,000 or more, according to geographic location.

The results of the study demonstrated that tweets related to Adderall peaked during December and May, which aligns with college students’ exam schedules. The geographic analysis revealed a concentration of tweets about Adderall along the northeastern portion of the United States and in some southern states. The rates of Twitter users tweeting about Adderall per 100,000 students in the east and south indicated greater use and abuse of the drug, which is consistent with findings from previous studies that examined the nonmedical use of prescription stimulants. Sleep deprivation (5%) and loss of appetite (2.6%) were the most common side effects discussed in tweets. The most common co-ingested substances mentioned were alcohol-related (4.8%) and stimulant-related, such as coffee or Red Bull (4.7%).

Hanson et al. conclude that their Twitter-based surveillance methodology produced similar findings to traditional survey designs related to Adderall use and abuse. They note that additional research is needed to understand the reasons for geographical variations in use of Adderall, and it may be useful to compare the geographic "hot spots" with those states that have active prescription drug monitoring programs. Furthermore, they acknowledge that ongoing Twitter conversations about Adderall abuse could end up exacerbating the problem, as the behavior could be portrayed as the "norm" among college students (Hanson et al., 2013a).

In a separate study, Carl Lee Hanson et al. examine the networks of individuals on Twitter who show signs of prescription drug abuse to determine whether they are connecting with others who are reinforcing this behavior. The results of this study were published in the *Journal of Medical Internet Research* in September 2013.

Hanson et al. explored three hypotheses in the study:

1. People discuss prescription drug abuse on Twitter.
2. People who discuss prescription drug abuse on Twitter belong to social circles that engage with each other about prescription drug abuse.
3. Social engagement about prescription drug abuse varies across social circles of those who discuss it, and higher engagement correlates with higher levels of abuse.

The researchers collected tweets mentioning prescription drug terms from November 29, 2011, through November 14, 2012. The tweets were further filtered to identify potential prescription drug abusers. This was done by examining Twitter users who had mentioned prescription drugs in at least 10 tweets but less than 100—a level shown through previous research to be most indicative of regular use. Twenty-five users from this sample were then chosen for further analysis.

Hanson et al. analyzed the 25 index users' social circles according to the levels of interactions the users had with their followers. Social circles of the same size were examined for each index user, and the most recent tweets of each Twitter user in the social circle were obtained for content analysis.

The results of the study demonstrated that index users and their social circles typically tweeted about similar drugs and indicated similar levels of abuse, according to pre-established keywords indicative of abuse. The levels of abuse were strongly correlated with the number of Twitter users interacting with others about prescription drugs. This suggests that Twitter users who tweet about prescription drug abuse are in an environment that potentially supports their behavior. Hanson et al. conclude that Twitter is being used "as a platform for discussion about prescription drug abuse within social circles…as such, [it] provides an additional 'access point' to groups of individuals who are abusing prescription drugs" (Hanson et al., 2013b).

V. Conclusion

The research reviewed above demonstrates that Twitter may be considered an effective public health monitoring and surveillance tool for *some* conditions. Analyzing tweets in the midst of an infectious disease outbreak or a natural disaster has proven immensely useful, providing officials with critical information quickly in situations of extreme urgency. Conversely, the analysis of tweets related to chronic conditions, such as obesity, does not appear to provide an accurate portrayal of the prevalence of these conditions. This is likely due to the potentially stigmatizing or personal nature of many chronic conditions. For instance, it is highly unlikely that the majority of obese people would tweet about their own obesity; rather, tweets from individuals about obesity are more often related to recent statistics or in reference to other individuals (Ghosh & Guha, 2013).

Researchers are just starting to attempt to determine whether Twitter can be used to better understand and prevent substance misuse and abuse; however, results to date have shown promise. In regard to prescription drug abuse in particular, it is apparent from the recent research outlined above that people do discuss their prescription drug misuse on Twitter and that misusers are likely to interact with other misusers via Twitter. This information, coupled with the relative accuracy of Twitter users' self-reported geographic location, suggests that the platform has potential to become a surveillance and monitoring tool for prescription drug misuse. Indeed, experts are already starting to realize its potential, as evidenced by the recent $11 million in funding awarded for further research on the topic.

Chapter III – Methodology

I. Introduction

Primary methodologies for this study included the collection and analysis of secondary data in the form of publicly available tweets, which were examined and filtered according to search terms and frequencies indicative of prescription opioid misuse. The data were analyzed quantitatively and qualitatively. A geographical analysis was also conducted to facilitate a comparison, whereby the extent to which trends in Twitter content correlate with documented patterns of behavior and prescribing rates was determined.

II. Population and Sample

Twitter is an online social media platform, through which users post tweets that are limited to 140 characters. Tweets are by default sent publicly, and users agree to Twitter's privacy policy when registering with the site. Users also populate their own public profiles, which include an optional geographic location field. The public nature of tweets makes Twitter an easily accessible database that can be used for the analysis of trends and insights related to a range of topics.

According to Twitter demographic research published in 2014, 19% of the entire U.S. adult population (aged 18 years or older) uses Twitter, and 23% of U.S. adult Internet users use Twitter. The demographic group most likely to use Twitter is male, between the ages of 18 – 29, college-educated, and has an income of $50,000 or more (Duggan et al., 2015). This demographic group happens to align with the group most at risk for prescription drug abuse, which is also male and between the ages of 18 and 25 (NIH, 2014b).

Table 1: Twitter User Demographics
(Percentages marked with an asterisk represent a significant change from 2013)
Source: Pew Research Center's Internet Project September 2014 Combined Omnibus Survey

Twitter users

Among online adults, the % who use Twitter

	2013	2014
All internet users	*18%*	*23%**
Men	17	24*
Women	18	21
White, Non-Hispanic	16	21 *
Black, Non-Hispanic	29	27
Hispanic	16	25
18-29	31	37
30-49	19	25
50-64	9	12
65+	5	10*
High school grad or less	17	16
Some college	18	24
College+ (n= 685)	18	30*
Less than $30,000/yr	17	20
$30,000-$49,999	18	21
$50,000-$74,999	15	27*
$75,000+	19	27*
Urban	18	25*
Suburban	19	23
Rural	11	17

Tweets about personal prescription opioid use from U.S. Twitter users were used in this study; any user that indicated he or she was under the age of 18 was excluded. Any tweets not in English or from a user that appeared to be from a pharmaceutical company or representing a third party were also excluded.

III. Research Design

Public tweets that were posted in the United States between July 1, 2014, and December 31, 2014, were collected and filtered according to a set of pre-specified search terms related to prescription opioids. The terms purposefully included slang words for prescription opioids, as well as common misspellings.

Table 2: Prescription Opioid Search Terms

Search Terms
• Percocet • Percs • Oxycontin • Oxycotton • Oxycet • Oxys • Opes • Vicodin • Vicoden • Vicadin • Hydrocodone • Hydros • Lortab

Additional filters were used to help eliminate users that were likely representing pharmaceutical companies or other third parties; this included eliminating tweets that included the words "pharma" or "pharmaceuticals," as well as tweets that included "http," which is indicative of a web link for more information and would likely be used by a company representative or the media.

Evaluation in a previous study on the portrayal of prescription drug misuse on Twitter revealed that Twitter users that mentioned prescription drugs in at least 10 tweets but less than 100 over the course of a year were most likely regular users (as opposed to users representing drug sales or automated feeds, or those who used the drugs sparingly) (Hanson et al., 2013b). Using this research as a guide, the tweets analyzed in this study were further filtered to only include tweets originating from users who tweeted about prescription opioids in at least 5 tweets but less than 50 over the 6-month period being analyzed.

The remaining sample of tweets was then analyzed quantitatively according to abusive behaviors. The tweets were coded according to the four abusive behavior categories outlined in Table 3. Keywords related to each abusive behavior category were used to guide the coding, but the coding was not limited to these keywords. For example, a tweet that read, "Gimme some of those percs" would have been coded as "Seeking" even though the keyword "gimme [sic]" is not outlined in Table 3.

Table 3: Keywords for Abusive Behaviors

Abusive Behaviors	Keywords
Overdose	• Too many • Two • Three • Double • Couple • Overdose
Co-ingestion	• Alcohol • Beer • Wine • Shot • Shots • Booze • Margarita • Xanax • Whiskey • Scotch • Gin • Drink
Dependence	• Sleep • Stress • Addicted/addiction • Stressful • Relax • High
Seeking	• Want • Wanting • Need • Needing • Wish • Give • Buy • Sell

During the quantitative coding process, any remaining tweets that were found to be irrelevant to this analysis (i.e., referencing news stories, quoting song lyrics, etc.) were also removed.

For the second portion of the analysis, tweets that originated from users who had provided a self-identified location in their user profile were analyzed geographically to determine a state-specific rate for the number of Twitter users tweeting about prescription opioid per 100,000 Internet users, according to 2010 census data. Ideally, this rate would have been calculated using the number of known Twitter users per state; however, because all Twitter users do not provide their geographic location, estimates of Twitter users by state are sparse and their validity has not been confirmed. To date, the most referenced estimates come from a social media monitoring platform called Twellow, which does not disclose its algorithm for estimating state-specific use numbers (Twellow, 2015).

These results were compared to:

1. The most recent state-specific opioid prescription drug prescribing rates available from CDC, (for 2012), as high prescribing rates are strongly associated with high rates of overdose (CDC, 2014b);
2. The most recent state-specific drug overdose mortality rates available from CDC, which were from 2010 (CDC, 2012b); and
3. The most recent estimated prevalence of nonmedical use of opioid prescription drugs by state available from the Substance Abuse and Mental Health Services Agency, which were from 2013 (SAMSHA, 2014).

IV. Instruments

The data for this analysis was collected through use of the online monitoring platform, Sysomos MAP. MAP is a database of billions of social media conversations and online news stories dating back up to one year, and it includes a full history of tweets. Fresh content is constantly added and stored by actively indexing Twitter. MAP is part of the Twitter Certified Products Program with access to the full Twitter fire hose.

A search query was constructed using a Boolean expression that would capture tweets related to prescription drug overdose according to the preselected search terms (see Table 1). The Boolean expression included a filter to reduce the unwanted tweets, as noted in the exclusion criteria in Section II of this Chapter. A geographic filter was also applied to limit the results to content produced in the United States.

V. Plans for Data Analysis

After the tweets were obtained through Sysomos MAP, they were exported into a Microsoft Excel spreadsheet for analysis. Frequencies, percentages, means, quartiles, and rates were calculated to describe the portrayal of prescription opioid misuse on Twitter. The Pearson product-moment correlation coefficient was used to measure the linear correlation between the rate of Twitter users tweeting about prescription opioid use per state and states' documented prescribing rates, overdose rates, and nonmedical prescription pain reliever use rates, according to the most recently available data outlined above. A critical value table for Pearson's correlation coefficient was used to determine whether the results were statistically significant with a p-value of 0.5.

VI. Limitations and Delimitations

There are several inherent limitations in this study design. Firstly, there is no way to determine whether the information being relayed in the tweets is genuine. For example, someone may tweet about wanting to "pop an oxy," but there is no way to determine if that person was joking or trying to project a certain image, without actually engaging in abusive behavior in reality. Secondly, considering that only 19% of the U.S. population uses Twitter, there are inherent concerns about its representativeness as a public health sample (Duggan et al., 2015). Furthermore, some states have a higher percentage of residents using Twitter than others; however, as referenced above, the validity of current estimates of the number of Twitter users per state has not been confirmed. Furthermore, some Twitter users do not include a location in their user profiles, and other users may not provide an accurate location, which means that the sample that was examined for this study may not be representative of all the Twitter users tweeting about prescription opioid misuse in specific geographic areas.

Several delimitations were imposed to produce a data set that was manageable to analyze for a single researcher. Only six months of tweets were examined, as opposed to a larger sample of tweets (up to one year, for example). Additionally, the search terms used to identify tweets about prescription opioids were limited in scope. General terms such as "pain reliever" or "pain killer" were omitted from the search terms, as they resulted in an overly large data set (more than 59,000) that included tweets more often referring to news stories or policy issues, as opposed to individual use. While this omission resulted in a more manageable dataset for one researcher to code, it may also have resulted in the elimination of some relevant tweets, which could have affected the study outcome. Other terminology or slang may also be used to refer to the issue on Twitter that was not captured in this analysis.

Chapter IV – Results

I. Introduction

The query entered into Sysomos MAP returned a total of 59,835 tweets, before any of the filters were applied, suggesting there is a high level of conversation on Twitter related to prescription opioids. After all filters were applied and the data was cleaned up by hand, the final data set comprised a total of 6,494 tweets about individual prescription opioid use originating from 809 unique users. The mean number of tweets per user was 8.17.

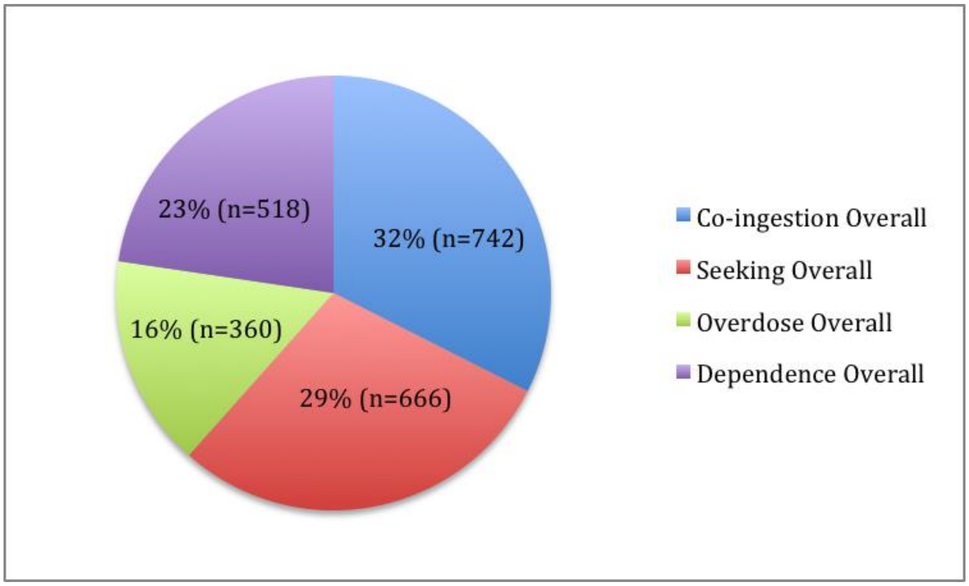
II. Findings

Of the 6,494 total tweets, 2,069 tweets (31.86%), originating from 622 unique users (76.9%), were indicative of abusive behaviors, according to the qualitative coding process used for this study. The abusive behavior most frequently mentioned overall was co-ingestion (n=742 tweets), followed by seeking, overdose, and dependence, respectively. Two hundred sixteen (216) tweets contained content that fell into more than one of the behavioral categories.

Table 4. Number of Tweets Indicating Prescription Opioid Misuse Behaviors By Category

Abusive Behaviors Mentioned	Number of Tweets
Seeking	605
Co-ingestion	573
Dependence	417
Overdose	258
Overdose & Co-ingestion	70
Overdose & Dependence	27
Co-ingestion & Dependence	54
Seeking & Co-ingestion	40
Seeking & Dependence	16
Seeking & Overdose	4
Overdose, Dependence, & Co-ingestion	4
Seeking, Co-ingestion, Overdose	1
Total	2,069

Figure 4. Percentage of Tweets Indicative of Prescription Opioid Misuse by Category



After reviewing Twitter users' self-identified location information and removing any ambiguous entries (e.g., multiple locations or slang words), state-specific locations were identified for 745 (92%) of the Twitter users. There was at least one user representing every state except for three: North Dakota, Vermont, and Utah. The mean rate was 0.27 tweets about prescription opioid use per 100,000 internet users, and rates ranged from 0.057 (West Virginia) to 2.58 (Washington, DC). The first and third quartiles were 0.14 and 0.29, respectively. The rate calculated for Washington, DC, was more than 1.5 interquartile ranges above the third quartile and will be considered an outlier for the purposes of this analysis.

Table 5: Twitter Users Tweeting about Prescription Opioid Use by State

State	Number of Twitter Users Tweeting about Opioid Use	Number of Internet Users	Rate per 100,000 Internet Users
West Virginia	1	1753000	0.057045066
Idaho	1	1468000	0.068119891
Arkansas	2	2743000	0.072912869
New Hampshire	1	1270000	0.078740157
Maine	1	1254000	0.079744817
New Mexico	2	1899000	0.105318589
Montana	1	920000	0.108695652
Alabama	5	4503000	0.111037086
South Dakota	1	763000	0.131061599
Iowa	4	2843000	0.140696447
Oklahoma	5	3505000	0.142653352
Mississippi	4	2789000	0.143420581
Kentucky	6	4067000	0.147528891
Alaska	1	660000	0.151515152
Hawaii	2	1210000	0.165289256
Illinois	22	12248000	0.179621163
Florida	33	17688000	0.186567164
North Carolina	17	8901000	0.190989776
Wyoming	1	521000	0.19193858
Indiana	12	6139000	0.195471575
Tennessee	12	6068000	0.197758734
Minnesota	10	5001000	0.199960008
South Carolina	9	4310000	0.208816705
Wisconsin	12	5401000	0.222181078
Kansas	6	2649000	0.226500566
Nebraska	4	1695000	0.235988201
Nevada	6	2528000	0.237341772
Connecticut	8	3364000	0.237812128
California	89	35181000	0.252977459
Virginia	19	7418000	0.256133729
New Jersey	22	8269000	0.266053936
Arizona	17	6340000	0.268138801
Ohio	31	11000000	0.281818182
Missouri	16	5625000	0.284444444
New York	54	18549000	0.291120815
Michigan	28	9473000	0.295576903
Massachusetts	19	6389000	0.297386132
Washington	19	6373000	0.298132748
Pennsylvania	36	11981000	0.300475753
Colorado	15	4836000	0.310173697

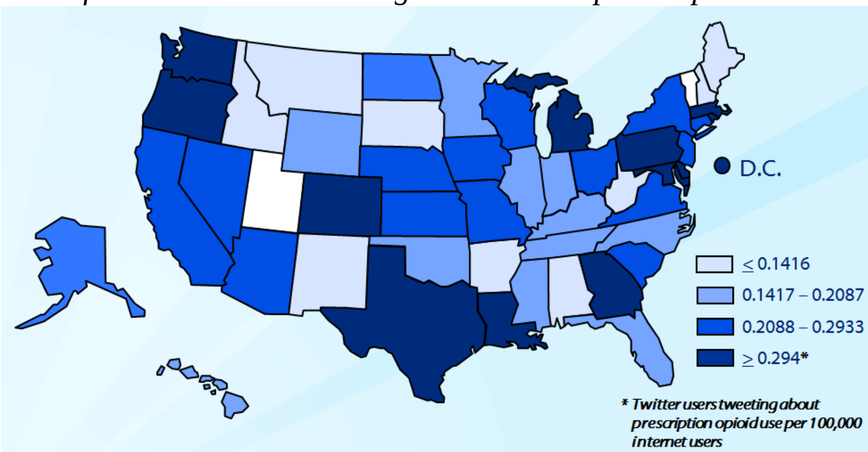
(continued)

State	Number of Twitter Users	Number of Internet Users	Rate per 100,000 Internet Users
Georgia	30	9296000	0.322719449
Texas	83	23481000	0.35347728
Delaware	3	842000	0.356294537
Louisiana	16	4272000	0.374531835
Oregon	15	3695000	0.405953992
Maryland	24	5431000	0.441907568
Rhode Island	5	994000	0.503018109
DC	15	581000	2.581755594

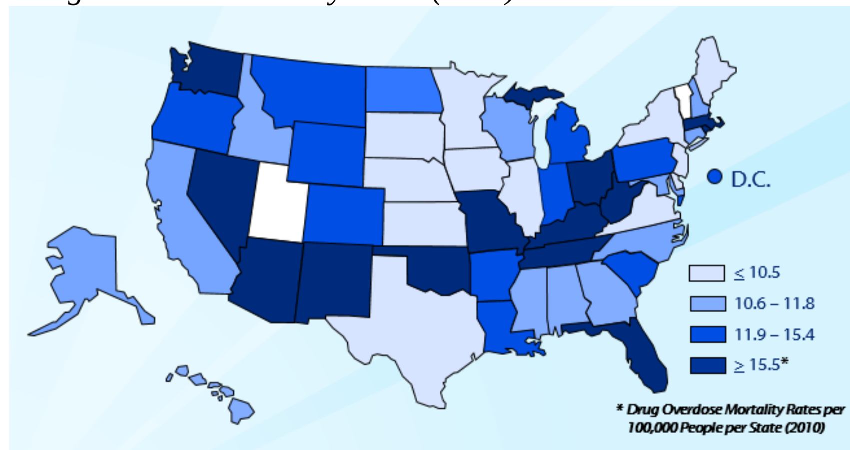
The rate of Twitter users tweeting about prescription opioid use by state was compared to the most recently available opioid prescribing rates, prescription drug overdose rates, and rates of nonmedical use of prescription pain relievers (see Table 6 and Figure 5). When calculating the correlation with the Washington, DC, rate, no correlation was identified between rates of Twitter users and any of the comparison data. However, with the Washington, DC, rate removed as an outlier, a modest correlation with prevalence estimates of nonmedical prescription pain reliever use in persons 12 years and older was identified ($r=0.303$), which was statistically significant ($p>0.5$). No correlations were identified for overdose and prescribing rates with the outlier removed.

Figure 5. Geographic Comparison of Rates of Internet Users Tweeting about Prescription Opioid Use, Overdose Rates, Prescribing Rates, and Nonmedical Use Prevalence Estimates

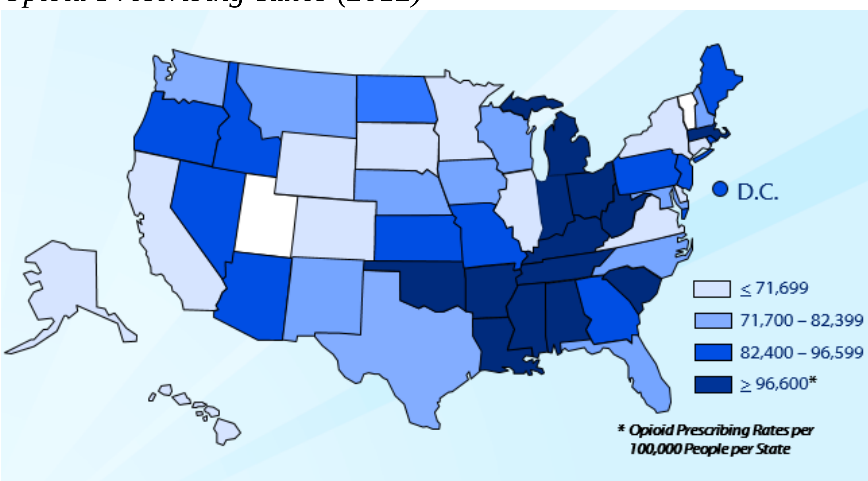
Rate of Internet Users Tweeting About Prescription Opioid Use



Drug Overdose Mortality Rates (2010)



Opioid Prescribing Rates (2012)



Nonmedical Use of Prescription Pain Relievers Prevalence (2013)

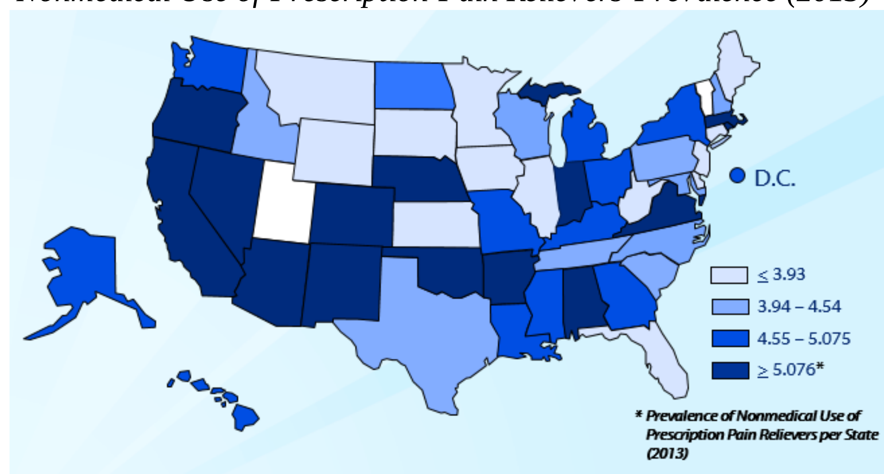


Table 6: Comparison of Rates of Internet Users Tweeting About Opioid Use with Prescription Drug Overdose Rates, Opioid Prescribing Rates, and Estimated Prevalence of Nonmedical Use of Prescription Pain Relievers, per 100,000 People

State	Rate of Internet Users Tweeting about Opioid Use	State Prescribing Rate (2012)	Overdose Mortality Rate (2010)	Rate of Nonmedical Use of Prescription Opioids - Ages 12 or Older (2013)
Alabama	0.111037086	142900	11.8	5.42
Alaska	0.151515152	65100	11.6	4.91
Arizona	0.268138801	82400	17.5	5.82
Arkansas	0.072912869	115800	12.5	5.38
California	0.252977459	57000	10.6	5.2
Colorado	0.310173697	71200	12.7	5.08
Connecticut	0.237812128	72400	10.1	3.56
DC	2.581755594	85700	12.9	4.62
Delaware	0.356294537	90800	16.6	4.87
Florida	0.186567164	72700	16.4	3.61
Georgia	0.322719449	90700	10.7	4.58
Hawaii	0.165289256	52000	10.9	4.54
Idaho	0.068119891	85600	11.8	4.45
Illinois	0.179621163	67900	10	3.57
Indiana	0.195471575	109100	14.4	5.28
Iowa	0.140696447	72800	8.6	3.93
Kansas	0.226500566	93800	9.6	3.65
Kentucky	0.147528891	128400	23.6	4.61
Louisiana	0.374531835	118000	13.2	4.85
Maine	0.079744817	85100	10.4	3.8
Maryland	0.441907568	74300	11	4.18
Massachusetts	0.297386132	70800	11	3.71
Michigan	0.295576903	107000	13.9	4.77
Minnesota	0.199960008	61600	7.3	3.59
Mississippi	0.143420581	120300	11.4	4.66
Missouri	0.284444444	94800	17	4.71
Montana	0.108695652	82000	12.9	3.93
Nebraska	0.235988201	79400	6.7	5.27
Nevada	0.237341772	94100	20.7	4.27
New Hampshire	0.078740157	71700	11.8	3.96
New Jersey	0.266053936	62900	9.8	5.22
New Mexico	0.105318589	73800	23.8	3.89
New York	0.291120815	59500	7.8	5.07
North Carolina	0.190989776	96900	11.4	4.07
Ohio	0.281818182	100100	16.1	4.9
Oklahoma	0.142653352	127800	19.4	5.43

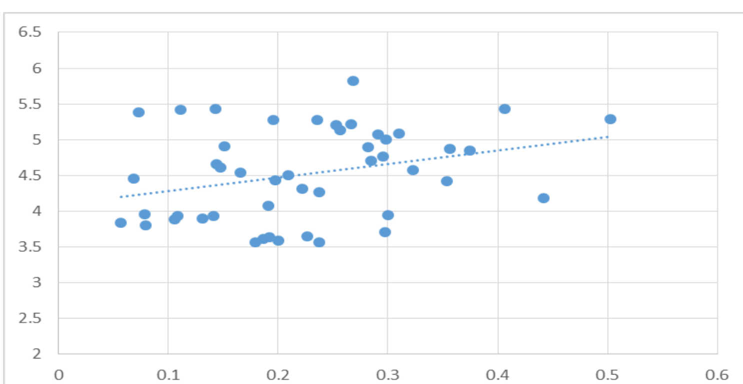
Continued

State	Rate of Internet users tweeting about Opioid use	State Prescribing Rate (2012)	Overdose Mortality Rate (2010)	Rate of Nonmedical Use of Prescription Pain Relievers - Ages 12 or Older (2013)
Oregon	0.405953992	89200	12.9	5.43
Pennsylvania	0.300475753	88200	15.3	3.94
Rhode Island	0.503018109	89600	15.5	5.29
South Carolina	0.208816705	101800	14.6	4.5
South Dakota	0.131061599	66500	6.3	3.9
Tennessee	0.197758734	142800	16.9	4.43
Texas	0.35347728	74300	9.6	4.42
Virginia	0.256133729	67400	6.8	5.13
Washington	0.298132748	77300	13.1	5
West Virginia	0.057045066	137600	28.9	3.84
Wisconsin	0.222181078	76100	10.9	4.31
Wyoming	0.19193858	69600	15	3.63

Table 7: Correlation of Rate of Internet Users Tweeting About Opioid Use with Prescription Drug Overdose Rates, Opioid Prescribing Rates, and Estimated Prevalence of Nonmedical Use of Prescription Pain Relievers

Comparison Data	Correlation Coefficient With DC	Correlation Coefficient Without DC
Nonmedical use of prescription pain relievers	0.107661	0.303144
Prescribing Rates	-0.06482	-0.18138
Overdose Rates	-0.04864	-0.13806

Figure 6: Scatterplot and Trendline Indicating Correlation between Internet Users Tweeting about Prescription Opioid Use and Prevalence of Nonmedical Use of Prescription Pain Relievers by State



In consideration of the likely wide range of Twitter users per state, the data was further analyzed to examine only states estimated to have high rates of Twitter users, according to data from Twellow (Twellow, 2015). This was done to determine whether the states estimated to have a relatively low percentage of the population using Twitter were influencing the correlation results. States with low Twitter usage rates could be less likley to have residents discussing prescription opioid use on Twitter, regardless of actual use trends, because they are less likely to be using Twitter in the first place. When comparing rates of just the top ten states estimated to have the highest rate of overall Twitter users, the correlation between the rate of Internet users tweeting about prescription opioid use and the estimated prevalence of nonmedical use of prescription pain relievers increased (r=0.5). However, there was still no correlation found with overdose and prescribing rates.

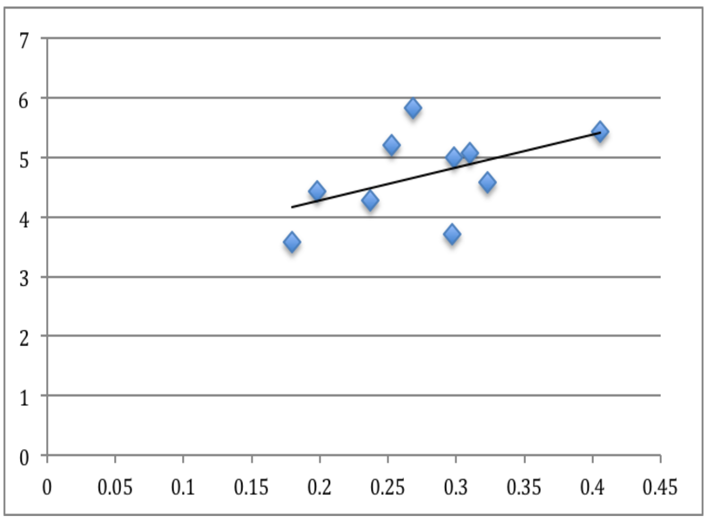
Table 8: Comparison of Rates of Internet Users Tweeting About Opioid Use with Prescription Drug Overdose Rates, Opioid Prescribing Rates, and Estimated Prevalence of Nonmedical Use of Prescription Pain Relievers, per 100,000 People for 10 States Estimated to Have the Highest Percentage of Twitter Users

State	Rate of Internet Users Tweeting about Prescription Opioid Use	Opioid Prescribing Rate	Overdose Rate	Nonmedical Use of Pain Reliever Prevalence
Washington	0.298132748	77300	13.1	5
Massachusetts	0.297386132	70800	11	3.71
Oregon	0.405953992	89200	12.9	5.43
Nevada	0.237341772	94100	20.7	4.27
Colorado	0.310173697	71200	12.7	5.08
Illinois	0.179621163	67900	10	3.57
Georgia	0.322719449	90700	10.7	4.58
California	0.252977459	57000	10.6	5.2
Arizona	0.268138801	82400	17.5	5.82
Tennessee	0.197758734	142800	16.9	4.43

Table 9: Correlation of Rate of Internet Users Tweeting About Opioid Use with Prescription Drug Overdose Rates, Opioid Prescribing Rates, and Estimated Prevalence of Nonmedical Use of Prescription Pain Relievers in 10 States Estimated to Have the Highest Percentage of Twitter Users

Comparison Data	Correlation Coefficient
Prescribing Rate	-0.193217991
Overdose Rate	-0.223292122
Nonmedical Use Rate	0.496606768

Figure 6: Scatterplot and Trendline Indicating Correlation between Internet Users Tweeting about Prescription Opioid Use and Prevalence of Nonmedical Use of Prescription Pain Relievers for 10 States Estimated to Have the Highest Percentage of Twitter Users



III. Other Findings

While a qualitative analysis of tweets related to the illegal selling and purchasing of prescription drugs was outside the scope of this study, it was observed that multiple tweets from indivduals (as opposed to organizations or companies) included content soliciting prescription opioids for sale. More research may be useful to determine whether Twitter could serve as a platform to help identify illegal drug traffickers in specific states, especially considering many individuals tweeting about buying and selling prescription drugs also included a self-identified location.

IV. Summary

The findings demonstrated positive responses to two out of the three research questions developed for this study—if the Washington, DC, rate is considered an outlier and removed. The majority of Twitter users tweeting about prescription opioid use used words indicative of misuse. There was a modest correlation between state rates of Twitter users tweeting about prescription opioid use and state prevalence estimates of nonmedical prescription pain reliever use. There was no correlation identified between state rates of Twitter users and opioid prescribing rates or prescription drug overdose rates.

Chapter V – Conclusions, Implications, and Recommendations

I. Introduction

The following chapter provides a summary and analysis of the findings outlined in Chapter IV. It also includes proposed explanations for portions of the research questions that were rejected, as well as potential implications of the findings on future public health work regarding prescription opioid misuse and overdose.

II. Summary of Study

Deaths resulting from drug overdoses are the leading cause of injury death in the United States and are often a result of nonmedical use or misuse of prescription opioids (CDC, 2014). Unintentional overdoses from prescription opioids have caused more deaths than overdoses from heroin and cocaine combined (CDC, 2011). In 2013, an estimated 4.5 million people over the age of 12 in the United States used prescription pain relievers for nonmedical purposes (SAMSHA, 2014).

While strategies exist to combat this epidemic, current monitoring and surveillance systems in place do not provide officials with real-time data. Research shows that substantial changes have taken place over time in the geographic distribution of opioid-related mortality (Paulozzi LJ et al., 2008). A need exists for a public health platform to help public health workers predict where in the United States a rise in misuse or overdose may be occurring, so that evidence-informed strategies may be implemented as quickly as possible.

This study examined the extent to which content of Twitter conversation (or "tweets") related to prescription opioid use reflects larger behavioral patterns and documented prescribing, overdose, and nonmedical use rates. It was designed to analyze how Twitter users describe prescription opioid use and misuse on Twitter, whether there were discernible state-specific trends in the quantity of tweets, and whether the variations in tweets by state reflected variations in prescribing rates.

A search was conducted for all tweets mentioning keywords related to prescription opioids during a six-month period. The tweets were filtered to include only those originating from individuals discussing individual prescription opioid use and who could potentially be abusing the drugs, as evidenced by the number of tweets sent. The tweets were sorted by user-identified geographic location and qualitatively analyzed for content indicative of misuse according to four categories: overdose, dependence, co-ingestion, and seeking. Rates were calculated to determine the number of Twitter users tweeting about prescription opioid use per 100,000 Internet users in each state. The rates were compared to state-specific prescribing rates, nonmedical use rates, and overdose rates.

Results demonstrated that Twitter users are engaging in conversation about prescription opioid use, and the majority mention behaviors that are indicative of misuse. A modest, statistically significant correlation between the state-specific rate of Twitter users mentioning prescription opioid use and the estimated state-specific prevalence of nonmedical use of prescription opioids was found, if one data point considered an outlier was removed. If the analysis was conducted using only the top ten states estimated to have the highest percentage of Twitter users, the correlation with nonmedical use prevalence grew stronger. No correlation was identified with prescribing or overdose rates.

III. Conclusion

The findings of this study indicate that Twitter users are discussing their prescription opioid use and misuse through tweets, and there is a wealth of information available to analyze within the content of these discussions. Abusive behaviors, especially co-ingestion (32%) and seeking (29%) (see Figure 4), were discussed at length. Users tweeted about ending days with a Vicodin and wine, for example, and others asked questions about how much alcohol they could mix with prescription opioids. Still others discussed the perfect “high” achieved by combining a prescription opioid with central nervous system depressants, such as Xanax.

It is important to note that coding the tweets according to the pre-established keywords (Table 3) alone would not have resulted in accurate proportions of tweets in the data set that were actually indicative of misuse. A plethora of slang words and shortened words or phrases were used to indicate abusive behaviors that could not have been captured in a pre-established keyword bank. For example, a user may tweet, “Where the percs at?” This suggests seeking behavior, even though the tweet did not including any of the seeking keywords in Table 3. Furthermore, many different slang words were used to indicate highs being felt as a result of prescription opioids, such as “hopped up,” “loaded,” and “lifted.” These were not anticipated beforehand. Therefore, hand coding was critical to achieving accurate results. However, as this study was conducted by a single researcher, the coding accuracy was not verified by other researchers coding the same dataset, which may be considered a limitation.

The high proportion of Twitter users tweeting about prescription opioid use who provided self-identified locations in their user profiles was also unexpected. While past research has suggested that only around 12% of Twitter users provide a self-identified location in their profiles (Burton et al., 2012), 92% of users in this analysis did. It is unclear why this may be so or if this indicates that Twitter users who are using prescription drugs are more likely to share location data; however, the high proportion of location information proved useful for the geographic analysis portion of this study. It is worth noting that the location field provided in a Twitter user's profile does not require a standardized entry, so users indicated their location in a variety of formats (e.g., "NYC," "Brooklyn," "NY,NY," or "the Brnx," etc.). Therefore, it was critical to clean up and format the location information by hand before moving forward with data analysis.

The representation of states in the data set was robust (47 states plus Washington, DC), which helped improve the comprehensiveness of the sample. Still, lack of data for Vermont, North Dakota, and Utah should be considered a limitation, especially as in the past decade, Utah has experienced a more than 400% increase in deaths associated with misuse and abuse of prescription drugs (Utah Department of Health, 2013). Utah's traditionally conservative culture and strict family environments may have contributed to the lack of Twitter conversation about the issue in the state, or it may be that Twitter users from Utah (and potentially Vermont and North Dakota) are less likely to provide their location information than users from other states.

There were also seven states in which only one Twitter user was identified to be discussing prescription opioid use according to the pre-specified search terms, and one of the states (West Virginia) is known to have highest overdose rate of all states, as of 2010. As evidence has proven that prescription opioid misuse is an indicator for overdose, the lack of correlation between the data points for this state begs further analysis. The population most likely to use prescription opioids in West Virginia may skew older than the relatively young (18-29 year old) population most likely to use Twitter, or West Virginia's culture may promote an environment in which individuals are less likely to discuss personal issues, such as prescription opioid use and misuse, on Twitter. People in West Virginia may also be less likely to use Twitter at all when compared to other states, according to the available estimates (Twellow, 2015).

Additional research into Twitter use and characterization of prescription opioid use in West Virginia may be needed to more fully understand this discrepancy and the future utility of Twitter data from this state.

Additionally, the unusually high rate of Twitter users tweeting about prescription opioid use who identified themselves as living in Washington, DC—when compared to rates for the rest of the states—suggests the Washington, DC, data point may not be valid. It is possible that users who in fact live in the suburbs surrounding Washington, DC, prefer to identify themselves as living in metropolitan DC because it indicates a certain status or class. Others may work and socialize in Washington, DC, and self-identify as living there, even though their physical address is just outside the city, in neighboring states such as Maryland or Virginia. Regardless the reason, because the rate for Washington, DC, was more than 1.5 interquartile ranges above the third quartile and could be considered an outlier, it was important to calculate results without that data point when examining correlation between data sets.

No correlation was found between rates of Twitter users tweeting about prescription opioid use and the most recent prescription opioid prescribing rates (CDC, 2012) or prescription drug overdose rates (CDC, 2010). There are multiple factors that may have contributed to this result. First, these data sets are three and five years old, respectively. As noted previously, the landscape of prescription drug use and misuse can change significantly over time; indeed, so can prescribing rates, as new state prescription drug monitoring laws are implemented. The three-year data gap between the data sets being compared may have contributed to the negative results.

Furthermore, this study was designed to identify individuals who were likely not only to be using prescription opioids, but also misusing them. The filter that excluded any users who had tweeted about prescription opioid use less than five times in the 6-month period likely excluded many users who were prescribed opioids, but were either unaware theat they were misusing them, were prescribed too high of a dosage by their health care provider, or did not consider themselves addicts or misusers. These types of individuals are still at risk for overdose and would be impacted by prescribing rates in their states. However, the filters set for this study may have excluded them for analysis. They may be more likely to refer to their prescription opioids in general terms, such as "pain killers" or "pain relievers," which were not used as search terms in this study. These individuals may also be more likely to tweet less about their use, but inclusion of tweets from users who had tweeted between one and 4 times was outside the scope of this study. Future research should take these limitations into consideration, depending on the target audience and outcomes being studied.

When analyzing the data without Washington, DC, the correlation between rates of Twitter users and prevalence estimates of nonmedical prescription pain reliever use by state—although modest—was promising. The prevalence data was slightly more recent (2013) than the data from the other two comparison sets, and it related most to the filters and search terms applied for this analysis. The findings indicate that with additional related research, Twitter may have the potential to become an early warning detection system for spikes in state-level prescription opioid misuse, particularly in states with highly-populated metropolitan cities and with high estimated percentages of Twitter users, as evidenced by the increase in correlation strength when focusing on only these geographic locations.

IV. Implications

Findings from this study have implications for both professionals involved in the prevention and treatment of prescription opioid misuse, as well as professionals who are monitoring the U.S. prescription drug epidemic using varying surveillance platforms. Results indicate that individuals are using Twitter to discuss their prescription opioid misuse, which makes the platform an entry point for those who are hoping to understand abusive behaviors and potentially intervene, if warranted. The primary Twitter demographic (ages 18 – 29 years) is still very vulnerable to and influenced by peer behavior and social norms, which makes them more likely to be negatively influenced by descriptions of abusive behavior on Twitter (Hanson et al., 2013).

Additionally, the results of the geographic analysis of Twitter users discussing prescription opioid use adds to the existing knowledge of the emerging field of infodemiology, a new field of study where the Internet, including social media platforms, provides channels for researchers to use to explore the distribution and determinants of information (Eysenback, 2006). The immediacy of Twitter offers a great potential for officials to identify, track, and prevent prescription opioid misuse in real time, which could eventually help reduce overdose rates. The results of this study indicate that Twitter may become a surveillance platform to help identify potential upticks in prescription opioid misuse, particularly in states known to have high proportions of residents active on Twitter. Further research is needed to understand its utility in states with lower levels of Twitter activity and in accurately indicating real-time prescribing rates or overdose rates.

V. Recommendations

Future studies could address the limitations identified in this study in a number of ways. First, the gap in time between the data sets being compared could be addressed by comparing historical Twitter data, if available, with the corresponding year's prescribing, overdose, and misuse rates. This would provide insight into the affect the time gap between data sets had on results. Furthermore, a larger research team could analyze a larger number of tweets from a larger representation of users, spanning a year or more and including a range of tweets per user that was more inclusive than the range used in this study. A larger team could also help develop a more robust coding system to qualitatively analyze the tweets and differentiate between those that are geniuine versus spurious, helping to isolate "the signal from the noise."

The effects of the differing overall rates of Twitter usage in each state could also be examined more closely. As Twitter's database of users' verifiable geographic information (provided via GPS coordinates sent from smart devices) grows, researchers may be able to use goespatial mapping to hone in on users tweeting from specific metropolitan counties. This would provide a refined set of data with less room for potential confounders related to differing Twitter usage estimates across states.

References

American College of Obstetricians and Gynecologists, Committee on Health Care for Underserved Women. (2012). Nonmedical use of prescription drugs. Committee Opinion No. 538. Obstetrics and Gynecology, 120(4), 977–82-977–82. Retrieved February 2015 from http://www.ncbi.nlm.nih.gov/pubmed/22996128

Birnbaum HG, White AG, Schiller M, Waldman T, Cleveland JM, and Roland CL. Societal costs of prescription opioid abuse, dependence, and misuse in the United States. *Pain Medicine* 2011; 12: 657-667.

Bosely JC, Zhao NW, Hill S, Shofer F, Asch DA, Becker LB, Merchant RM. (2012). Decoding Twitter: Surveillance and trends for cardiac arrest and resuscitation communication. Resuscitation. 84(2013) 206-212. doi: 10.1016/j.resuscitation.2012.10.017.

Burton SH, Tanner KW, Giraud-Carrier CG, West JH, Barnes MD. (2012). "Right Time, Right Place" Health Communication on Twitter: Value and Accuracy of Location Information. Journal of Medical Internet Research. 14(6): e156. Published online November 2012. Doi: 10.2196/jmir.2121.

Centers for Disease Control and Prevention. (2015). National Vital Statistics System mortality data. Available from http://www.cdc.gov/nchs/deaths.htm.

Centers for Disease Control and Prevention. (2014). Web-based Injury Statistics Query and Reporting System (WISQARS) [online]. Available from http://www.cdc.gov/injury/wisqars/fatal.html.

Centers for Disease Control and Prevention. (2014b). Variation among states in prescribing of opioid pain relievers and benzodiazepines – United States, 2012. MMWR 2014; 63(26): 563-568. Retrieved February 2015 from http://www.cdc.gov/mmwr/preview/mmwrhtml/mm6326a2.htm?s_cid=mm6326a2_w.

Centers for Disease Control and Prevention. (2013). Policy Impact: Prescription Painkiller Overdoses. Available from http://www.cdc.gov/HomeandRecreationalSafety/rxbrief/index.html.

Centers for Disease Control and Prevention. (2012a). CDC Grand Rounds: Prescription Drug Overdoses — a U.S. Epidemic. MMWR 2012; 61: 10-13. Retrieved February 2015 from http://www.cdc.gov/mmwr/preview/mmwrhtml/mm6101a3.htm

Centers for Disease Control and Prevention, National Center for Health Statistics. (2012b). Multiple Cause of Death 1999-2010 on CDC WONDER Online Database, released 2012. Data are from the Multiple Cause of Death Files, 1999-2010, as compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program. http://wonder.cdc.gov/mcd-icd10.html.

Centers for Disease Control and Prevention. (2011). Vital Signs: Overdoses of Prescription Opioid Pain Relievers—United States, 1999-2008. MMWR 2011; 60: 1-6. Available from http://www.cdc.gov/mmwr/preview/mmwrhtml/mm6043a4.htm?s_cid=mm6043a4_w

Duggan M, Ellison N, Lampe C, Lenhart A, Madden M. (2015). Social Media Update 2014. Pew Research Center. Retrieved February 2015 from http://www.pewinternet.org/2015/01/09/social-media-update-2014/.

Eysenback G. (2006). Infodemiology: tracking flu-related searches on the web for syndromic surveillance. AMIA Annual Symposium Proceedings. 2006:244-8.

Ghosh D, Guha R. (2013). What are we 'tweeting' about obesity? Mapping tweets with topic modeling and geographic information system. Cartography and Geographic Information Science. 40(2): 90-102. doi: 10.1080/15230406.2013.776210.

Hanson CL, Burton SH, Giraud-Carrier C, West, JH, Barnes MD, Hansen B. (2013a). Tweaking and tweeting: exploring Twitter for nonmedical use of a psychostimulant drug (Adderall) among college students. Journal of Medical Internet Research. 15(4): e62. Published online April 2013. doi: 10.2196/jmir.2503.

Hanson CL, Cannon B, Burton S, Giraud-Carrier C. (2013b). An Exploration of Social Circles and Prescription Drug Abuse Through Twitter. Journal of Medical Internet Research. 15(9): e189. Published online September 2013. Doi: 10.2196/jmir.2741.

Hill S, Merchant R, Ungar L. (2013). Lessons learned about public health from online crowd surveillance. Big Data. 1(3). Published online September 2013. Doi: 10.1089/big.2013.0020

King NB, Fraser V, Boikos C, Richardson R, & Harper S. (2014). Determinants of Increased Opioid-Related Mortality in the United States and Canada, 1990–2013: A Systematic Review *American Journal of Public Health*. 2014 August; 104(8): e32–e42. Published online 2014 August. doi: 10.2105/AJPH.2014.301966

National Institutes of Health. (2014a). Using social media to better understand, prevent, and treat substance abuse [Press release]. Retrieved February 2015 from http://www.nih.gov/news/health/oct2014/nida-17.htm.

National Institutes of Health, National Institute on Drug Abuse. (2014b). Research Report Series: Prescription Drug Abuse. NIH Publication Number 15-4881. Retrieved February 2015 from http://www.drugabuse.gov/sites/default/files/prescriptiondrugrrs_11_14.pdf

Nord, WR. (1969). Social exchange theory: an integrative approach to social conformity. Psychological Bulletin. Mar;71(3):174-208.

Pan B & Crotts J. (2014). Theoretical models of social media with marketing implications. In Web 2.0 in Travel, Tourism and Hospitality: Theory, Practice and Cases. Ashgate Publishing.

Paul M, Dredze M. (2011). You are what you tweet: Analyzing Twitter for public health. 5th International AAAI Conference on Weblogs and Social Media; July 17-21, 2011; Barcelona, Spain. 2011.

Paulozzi LJ & Xi Y. (2008). Recent changes in drug poisoning mortality in the United States by urban-rural status and by drug type. Pharmacoepidemiology and Drug Safety. 17(10):997–1005. Retrieved February 2015 from http://www.ncbi.nlm.nih.gov/pubmed/18512264.

Signorini A, Segre AM, Polgreen PM. (2011). The use of Twitter to track levels of disease activity and public concern in the U.S. during the Influenza A H1N1 pandemic. PLoS ONE. 6(5): e19467. Published online May 2011. doi: 10.1371/journal.pone.0019467.

St Louis C. & Zorlu G. (2012). Can Twitter predict disease outbreaks? British Medical Journal. 344: e235. Published online May 2012. Retrieved February 2015 from http://www.bmj.com.proxy.library.emory.edu/content/344/bmj.e2353.

Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality. (2014). Results from the 2013 National Survey on Drug Use and Health: Summary of national findings (HHS Publication No. 14–4863, NSDUH Series H–48).

Szomszor M, Kostkova P, de Quincey E. (2010). #swineflu: Twitter predicts swine flu outbreak in 2009. 3rd International ICST Conference on Electronic Healthcare for the 21st Century (eHealth2010), 13-15 December 2010, Casablanca, Morocco. Retrieved February 2015 from www.mendeley.com/research/swineflu-twitter-predicts-swine-flu-outbreak-2009/#page-1.

Twellow. (2015). Twellowhood for the World. Available from https://www.twellow.com/twellowhood.

Utah Department of Health. (2013). Prescription Drug Overdoses. Violence and Injury Prevention Program. Available from http://www.health.utah.gov/vipp/topics/prescription-drug-overdoses/.