"Retweet to Pass the Blunt": Analyzing Geographic and Content Features of Cannabis-Related Tweeting Across the United States

RAMINTA DANIULAITYTE, PH.D.,[a,b,*] FRANCOIS R. LAMY, PH.D.,[a,b] G. ALAN SMITH, B.S.,[b] RAMZI W. NAHHAS, PH.D.,[c,d] ROBERT G. CARLSON, PH.D.,[a,b] KRISHNAPRASAD THIRUNARAYAN, PH.D.,[b] SILVIA S. MARTINS, M.D., PH.D.,[e] EDWARD W. BOYER, M.D., PH.D.,[f] & AMIT SHETH, PH.D.[b]

[a]*Center for Interventions, Treatment, and Addictions Research (CITAR), Department of Population and Public Health Sciences, Wright State University Boonshoft School of Medicine, Dayton, Ohio*

[b]*Ohio Center of Excellence in Knowledge-Enabled Computing (Kno.e.sis), Department of Computer Science and Engineering, Wright State University, Dayton, Ohio*

[c]*Department of Population and Public Health Sciences, Wright State University Boonshoft School of Medicine, Dayton, Ohio*

[d]*Department of Psychiatry, Wright State University Boonshoft School of Medicine, Dayton, Ohio*

[e]*Department of Epidemiology, Columbia University Mailman School of Public Health, New York, New York*

[f]*Department of Emergency Medicine, Brigham and Women's Hospital, Boston, Massachusetts*

ABSTRACT. Objective: Twitter data offer new possibilities for tracking health-related communications. This study is among the first to apply advanced information processing to identify geographic and content features of cannabis-related tweeting in the United States. **Method:** Tweets were collected using streaming Application Programming Interface (March–May 2016) and were processed by eDrugTrends to identify geolocation and classify content by source (personal communication, media, retail) and sentiment (positive, negative, neutral). States were grouped by cannabis legalization policies into "recreational," "medical, less restrictive," "medical, more restrictive," and "illegal." Permutation tests were performed to analyze differences among four groups in adjusted percentages of all tweets, unique users, personal communications only, and positive-to-negative sentiment ratios. **Results:** About 30% of all 13,233,837 cannabis-related tweets had identifiable state-level geo-information. Among geolocated tweets, 76.2% were personal communications, 21.1% media, and 2.7% retail. About 71% of personal communication tweets expressed positive sentiment toward cannabis; 16% expressed negative sentiment. States in the recreational group had significantly greater average adjusted percentage of cannabis tweets (3.01%) compared with other groups. For personal communication tweets only, the recreational group (2.47%) was significantly greater than the medical, more restrictive (1.84%) and illegal (1.85%) groups. Similarly, the recreational group had significantly greater average positive-to-negative sentiment ratio (4.64) compared with the medical, more restrictive (4.15) and illegal (4.19) groups. Average adjusted percentages of unique users showed similar differences between recreational and other groups. **Conclusions:** States with less restrictive policies displayed greater cannabis-related tweeting and conveyed more positive sentiment. The study demonstrates the potential of Twitter data to become a valuable indicator of drug-related communications in the context of varying policy environments. (*J. Stud. Alcohol Drugs, 78,* 910–915, 2017)

THERE IS A GROWING RECOGNITION that social media data can broaden the scope of existing substance use monitoring systems by enhancing their capacity for early identification of emerging trends (Corazza et al., 2013; Mounteney & Haugland, 2009). With changing cannabis legalization policies (Pacula et al., 2015; Room, 2014), public health professionals require timely information on cannabis use practices and trends.

Twitter, a microblogging service provider and social network platform, reports 310 million monthly active users (Twitter, 2016) who generate more than 500 million tweets per day (Internet Live Stats, 2016). Because of the availability of geo-identifiable information, Twitter data provide an opportunity to examine regional differences in terms of tweeting activity related to selected topics or substance use issues (Daniulaityte et al., 2015; Lamy et al., 2016).

The content of tweets, although brief and limited to 140 characters, can also be used to provide insights into user attitudes and behaviors related to drugs (Hanson et al., 2013; Thompson et al., 2015). Several prior studies used manual coding to classify substance use–related tweets by sentiment (Cavazos-Rehg et al., 2015a; Lamy et al., 2016; Shutler et al., 2015). Sentiment analysis seeks to identify positive, negative, and neutral attitudes toward cannabis expressed in tweet content. Prior research has also used manual coding to classify tweets by source/type of communication (Cavazos-Rehg et al., 2015b; Lamy et al., 2016). The ability to identify personal communications versus media and/or retail might help increase data quality for monitoring trends in communications of substance use (Kim et al., 2016; Yin et al., 2014). Although a few

Received: November 29, 2016. Revision: April 19, 2017.

This study was funded by National Institute on Drug Abuse Grant R01 DA039454 (Raminta Daniulaityte/Amit Sheth, principal investigators).

*Correspondence may be sent to Raminta Daniulaityte at the Center for Interventions, Treatment, and Addictions Research, Department of Population and Public Health Sciences, Wright State University Boonshoft School of Medicine, 3171 Research Blvd., Suite 124, Dayton, OH 45420-4006, or via email at: raminta.daniulaityte@wright.edu.

prior studies reported development of automated tools to process substance use–related tweets (Alvaro et al., 2015; Cole-Lewis et al., 2015; Myslín et al., 2013), most prior analyses of cannabis and other drug-related Twitter data were limited to relatively small samples of tweets.

This study is the first to deploy custom-developed advanced information processing techniques to analyze a complete sample of collected Twitter data. Further, the study is among the first to integrate geographic and content analysis features to better understand trends in substance use–related communications on Twitter. The key goals of the study are to identify type of communication and sentiment expressed in cannabis-related tweets and to analyze regional variability in relation to state-level cannabis legalization policies.

Method

Data collection

A Twitter data processing framework was available through the eDrugTrends platform (Daniulaityte et al., 2016; Sheth et al., 2014). Tweets were collected using Twitter's streaming Application Programming Interface that provides free access to 1% of all tweets (Internet Live Stats, 2016). eDrugTrends collects an average of 150,000 tweets per day, which is well below the allowable limit (1% constitutes about 5 million tweets per day, assuming 500,000,000 daily total volume). Thus, it is safe to assume that our data collection captures all or most tweets with relevant keywords selected for our study (Morstatter et al., 2013).

Data were collected March 1–May 31, 2016. Data collection was limited to English language content. Collected data included not only original tweets, but also retweets, because by retweeting, individuals actively participate in information sharing.

The following keywords were used to collect tweets: cannabis, weed, marijuana, spliff, ganja, kush, sativa, indica, chronic, blunt, hydro, dro, skunk, reefer, joint, pot, herb, loud, #420, THC, gravs, mj, jay. They were identified based on prior research (National Institute on Drug Abuse, 2014) and examination of Twitter and web forum discussions of cannabis-related content. Selected keywords were pre-tested. To increase accuracy of collected tweets, the following filtering strategies were used: (a) ambiguous keywords ("pot," "herb," "skunk," "loud," "hydro," "jay," "mj," "gravs") were modified by adding smoke/smoked/smoking (e.g., "smoke skunk"); (b) a number of blacklist words/phrases were used to filter out irrelevant tweets (e.g., "blunt statement").

The university institutional review board approved the study under Human Subjects Research Exemption 4 because it is limited to publicly available tweets. To protect anonymity, cited tweet content was modified slightly, and Global Positioning System (GPS) coordinates were converted into state-level information and analyzed in aggregate form.

Data processing

Geolocation information of tweets was processed by the eDrugTrends platform (Sheth et al., 2014). Twitter users may indicate geolocation information in their user profiles or enable their tweets to contain GPS coordinates via a mobile phone that supports the feature. Tweets that contained geolocation information indicating a state in the United States were extracted for further analysis.

To adjust for the different level of tweeting activity in each state, the eDrugTrends platform concurrently collects a general sample of tweets (tweet collection without use of any keywords, which results in a "default" random sample of 1% of tweets provided by the Twitter API). The general sample is processed by eDrugTrends to identify geolocation information of extracted tweets.

To analyze tweet content, the study used eDrugTrends automated source and sentiment analysis classifiers that were developed using supervised machine-learning algorithms (Daniulaityte et al., 2016). The source classifier automatically classifies tweets into one of three types or sources—personal communication, official/media related, and retail related—with fairly high accuracy. Sentiment classification is applied to personal communication tweets only. Tweets that express positive qualities/effects of the drug or indicate or encourage use of the drug are classified as positive tweets; in contrast, tweets that express negative qualities/effects or discourage usage are classified as negative. The "neutral/unidentifiable" group includes tweets that do not express an opinion or do not contain enough textual information to determine the sentiment (for more information on the development and accuracy of source and sentiment classifiers, see Daniulaityte et al., 2016).

Data analyses

To examine regional patterns of general cannabis-related tweeting (all tweets), raw numbers of cannabis-related tweets for each state were extracted. Next, tweet-volume-adjusted state-level percentages of cannabis-related tweets for the 3-month period were computed by dividing the number of cannabis-related tweets for each state by the number of general sample tweets for the same state during that period. These percentages were then rescaled by dividing each by the sum across states and multiplying by 100, resulting in adjusted state-specific percentages of cannabis-related tweets.

A permutation test with 10,000 replications was performed using R 3.3.1 (Team, 2016) to examine differences in the adjusted percentages of cannabis-related tweets among U.S. states with different cannabis legalization policies. We tested the null hypothesis of no difference in mean adjusted percentage between groups of states defined by legal status. Two-sided pairwise comparisons between four groups were adjusted for six multiple comparisons using the Hommel

Table 1. Adjusted proportion of cannabis-related tweets by states with different cannabis legalization policies

Groups of states by cannabis legal status	A. All tweets *M* % (*SD*)	B. Unique users *M* % (*SD*)	C. Personal communication tweets *M* % (*SD*)	D. Ratio of positive to negative sentiment tweets *M* (*SD*)
(1) Recreational AK, CO, DC, OR, WA	3.01 (1.06)	2.31 (0.35)	2.47 (0.57)	4.64 (0.58)
(2) Medical, less restrictive AZ, CA, HI, IL, MA, ME, MI, MT, NM, NV, RI	2.13 (0.26)	2.10 (0.15)	2.08 (0.25)	4.32 (0.29)
(3) Medical, more restrictive CT, DE, MD, MN, NH, NJ, NY, VT	1.79 (0.22)	1.95 (0.15)	1.84 (0.30)	4.15 (0.31)
(4) Illegal AL, AR, FL, GA, IA, ID, IN, KS, KY, LA, MO, MS, NC, NE, ND, OH, OK, PA, SC, SD, TN, TX, UT, VA, WI, WV, WY	1.75 (0.27)	1.84 (0.22)	1.85 (0.28)	4.19 (0.27)
Hommel-adjusted *p* values for pairwise comparisons	**1 vs. 2: .0204**	1 vs. 2: .2837	1 vs. 2: .1468	1 vs. 2: .3168
	1 vs. 3: <.0001	**1 vs. 3: .0473**	**1 vs. 3: .0055**	**1 vs. 3: .0490**
	1 vs. 4: <.0001	**1 vs. 4: .0008**	**1 vs. 4: .0033**	**1 vs. 4: .0340**
	2 vs. 3: .3616	2 vs. 3: .3130	2 vs. 3: .2848	2 vs. 3: .5760
	2 vs. 4: .1401	**2 vs. 4: .0175**	2 vs. 4: .2136	2 vs. 4: .6318
	3 vs. 4: .8538	3 vs. 4: .3130	3 vs. 4: .9129	3 vs. 4: .7466

Note: **Bold** indicates statistical significance at the family-wise .05 level of significance.

procedure (Hommel, 1988) via SAS PROC MULTTEST in SAS 9.4 (SAS Institute Inc., Cary, NC) to maintain a family-wise $\alpha = .05$ level of significance. As a sensitivity analysis to check the impact of the potential for some individuals to tweet far more than others, making them highly influential in the analysis and introducing dependence, the same analysis was repeated to compare mean adjusted percentages of unique users who posted cannabis-related content.

States' legal statuses were classified into four groups (Table 1). The "recreational" group comprised Alaska, Colorado, the District of Columbia, Oregon, and Washington. As of February 2016, they all had at least started implementation of recreational marijuana laws. We recognize that the group is not homogenous in terms of roll-out of recreational marijuana markets. For example, unlike other recreational states, the District of Columbia did not provision sales for recreational use. States that have medical marijuana programs were grouped into two categories based on their restrictiveness (Chapman et al., 2016) and medical/nonmedical orientation scores (Williams et al., 2016) as of February 2016. The "medical, less restrictive" group comprised 11 states classified as having less regulated programs. "Medical, more restrictive" comprised 8 states classified as having more restricted and "medicalized" programs. The "illegal" group was made up of the 27 states that had no comprehensive medical cannabis laws implemented as of February 2016 (Table 1).

For content analysis, eDrugTrends source classifier was used to process all cannabis-related tweets to identify personal communication, media, and retail-related tweets. Next, personal communication tweets were processed further to identify positive and negative sentiment.

To analyze regional variation in cannabis-related personal communication tweeting, numbers of cannabis-related personal communication tweets for each state were extracted and adjusted by the general sample numbers. To examine regional differences in sentiment expressed toward cannabis, numbers of personal communication tweets that were identified as positive and negative were extracted for each state. Positive-to-negative ratios were calculated for each state by dividing positive tweet numbers by negative tweet numbers. Permutation tests were conducted using the same methods as for the general cannabis-related tweeting.

Results

Regional variation in cannabis-related tweeting

Between March and May of 2016, the eDrugTrends platform collected 13,233,837 cannabis-related tweets; 3,948,402 (30%) of those tweets had identifiable state-level geolocation information and were posted by 965,610 unique users. Raw counts of cannabis-related tweets were the highest in California, Texas, Florida, and New York, which are also the most populous states. However, after we adjusted for the different levels of tweeting activity (obtained from the general sample), Colorado (4.65%) and Oregon (3.42%) had the highest adjusted proportion of cannabis-related tweets, whereas Wyoming (1.14%) and Arkansas (1.31%) had the lowest. Permutation tests revealed that the recreational group had a statistically significantly greater average adjusted percentage of cannabis-related tweets (3.01%) compared with the other three groups after adjusting for multiple comparisons (Table 1, column A). Analysis of adjusted proportions of unique Twitter users revealed similar differences between recreational and other states (Table 1); it also showed that the medical, less restrictive group (2.31%) was significantly greater than the illegal group (1.84%) (Table 1, column B).

Regional variation in cannabis-related tweeting by source/type

Of the total sample of 3,948,402 geolocated tweets, the majority (76.2%) were identified as personal communication tweets. They expressed personal opinions and experiences: "I need a blunt"; "Hit this weed cause it might calm you down." About 21.1% were identified as media-related communications ("Board of #health cautious on medical #marijuana benefits. News source: https://t.co/NzEfG6pj3W"), and 2.7% were retail-related, promoting cannabis products and businesses ("Check out our newly renovated space and receive $10 off your first purchase! #cannabis https://t.co/qxvLY9PGc").

States varied in the volume and percentage of media- and retail-related tweets. For example, the raw percentage of retail-related tweets, out of all cannabis-related tweets in each state, ranged from 1.3% in Mississippi to 5.6% in Washington. Similarly, percentages of media-related tweets ranged from 12.6% in Mississippi to 44.4% in Colorado.

Given such variability in media- and retail-related tweeting in each state, the next step was to examine differences among states in personal communication tweeting only, after removing media- and retail-related tweets. The adjusted percentages of personal communication cannabis-related tweets were the greatest in Colorado (3.17%), Oregon (2.77%), and Alaska (2.49%). Permutation tests revealed that the recreational group had the greatest average adjusted percentage of personal communication tweets (2.47%) in comparison with the other groups, significantly greater than the medical, more restrictive (1.84%) and illegal (1.85%) groups, but not significantly different than the medical, less restrictive group (Table 1, column C).

Regional variation in ratio of positive to negative cannabis-related tweets

Out of all personal communication tweets, the majority (70.9%; 2,132,720) were classified as expressing positive sentiment toward cannabis. For example, "Weed is the most important meal of the day"; "Smoke good weed with a bad bitch"; "You should start smoking weed"; "Companies that marijuana test are missing out on great employees"; "Legalize weed in California 2016"; "Leaving work. Today was so stressful. Need a blunt LOL a fat one."

About 16% (483,819) of personal communication tweets were classified as expressing negative sentiment. For example, "So tired of people's need to always smoke weed"; "Marijuana is supposed to relieve anxiety not make it worse"; "I don't even smoke weed"; "Do y'all know how hard it is to stop smoking weed, like the hardest." The remaining 13.1% (392,973) were classified as neutral/unidentifiable.

Overall, tweets expressed overwhelmingly positive sentiment toward cannabis, with positive-to-negative ratio of 4.4:1 for the country as a whole. Although all states showed mostly positive attitudes toward cannabis, there was some variation across states, ranging from the low of 3.61 in Delaware to the highest of 5.38 in Colorado. Permutation tests revealed that the recreational group had the greatest average positive-to-negative ratio (4.64) in comparison with the other groups, significantly greater than the medical, more restrictive and illegal groups (Table 1, column D).

Discussion

This is the first study to integrate content and geographic analysis features to better understand Twitter data about cannabis-related communications. Our analysis shows that states that allow recreational marijuana were significantly different from more restrictive states in terms not only of general tweeting but also of personal communication tweets and of positive sentiment expressed in tweet content.

This is the first study to apply automated content classification to process almost 4 million cannabis-related tweets. Most prior studies were limited to small, manually coded samples. Our automated content analysis revealed that the majority of tweets (76%) were personal communication tweets, whereas media comprised about 21%, and retail less than 3% of all tweets. Sentiment expressed in personal communication tweets was overwhelmingly positive—with about 71% expressing positive attitude toward cannabis and only about 16% conveying negative sentiment. Predominantly positive attitudes toward cannabis-related products have been also identified by prior studies (Cavazos-Rehg et al., 2015a; Lamy et al., 2016).

For all four measures of cannabis-related tweeting (all tweets, unique users, personal communication tweets, and positive-to-negative tweet ratio), Colorado, Oregon, and Washington consistently ranked among the very top states. Colorado and Washington were the first states to legalize recreational cannabis use and have established booming commercial markets for cannabis products (Kleiman, 2015). Oregon recreational cannabis laws became operational in 2015, and the state has one of the oldest medical marijuana programs in the United States (Oregon Health, 2015).

Similar regional differences were identified in relation to marijuana concentrates and marijuana edible-related tweeting activity, with legal recreational states ranking higher than states with more restrictive policies (Daniulaityte et al., 2015; Lamy et al., 2016). However, there was an even greater variability in marijuana concentrate–related tweeting (Daniulaityte et al., 2015), which was possibly attributable to their status as emerging products.

This study describes deployment of powerful new methods for monitoring social media communications related to substance use trends and demonstrates the potential of Twitter data for becoming a valuable indicator of drug-related communications in the context of varying policy environments, one that could be used to complement traditional epidemiologic indicators.

Several limitations are noted. First, our study is limited to English language content. Second, our source classification was based on the content of tweets. It is possible that some retailers may post messages that resemble personal communications. Further research should improve classification of Twitter data by integrating analysis of both tweet content and public user account metadata. Last, more research is needed to evaluate the relationship between drug-related communications via social media and actual drug use practices and prevalences. For example, one approach would include analyzing correlations between drug-related Twitter data and drug use data collected through national epidemiologic surveys.

Ongoing data collection by eDrugTrends will allow us to examine changes in tweeting activity over time, across regions, and within states that experience rapid changes in cannabis policies. Further development of automated information extraction methods will help conduct more powerful, in-depth analyses of Twitter and other social media data related to substance use communications, including assessment of regional and temporal patterns in reported adverse health experiences or motivations of use. It will also help identify emerging communications related to new cannabis-related products and trends that could provide timely alerts to inform policy makers and epidemiologists.

In the age of expanding use of digital communication technologies and growing importance of social media–based interactions, analysis of social media–based communications on substance use trends is an increasingly important task for early warning systems across the globe (Corazza et al., 2013; NDEWS, 2016). This study contributes important methodological advancement to this rapidly expanding but still very new field that is set to gain greater significance in the future.

References

Alvaro, N., Conway, M., Doan, S., Lofi, C., Overington, J., & Collier, N. (2015). Crowdsourcing Twitter annotations to identify first-hand experiences of prescription drug use. *Journal of Biomedical Informatics, 58,* 280–287. doi:10.1016/j.jbi.2015.11.004

Cavazos-Rehg, P. A., Krauss, M., Fisher, S. L., Salyer, P., Grucza, R. A., & Bierut, L. J. (2015a). Twitter chatter about marijuana. *Journal of Adolescent Health, 56,* 139–145. doi:10.1016/j.jadohealth.2014.10.270

Cavazos-Rehg, P. A., Krauss, M. J., Sowles, S. J., & Bierut, L. J. (2015b). "Hey everyone, I'm drunk." An evaluation of drinking-related Twitter chatter. *Journal of Studies on Alcohol and Drugs, 76,* 635–643. doi:10.15288/jsad.2015.76.635

Chapman, S. A., Spetz, J., Lin, J., Chan, K., & Schmidt, L. A. (2016). Capturing heterogeneity in medical marijuana policies: A taxonomy of regulatory regimes across the United States. *Substance Use & Misuse, 51,* 1174–1184. doi:10.3109/10826084.2016.1160932

Cole-Lewis, H., Varghese, A., Sanders, A., Schwarz, M., Pugatch, J., & Augustson, E. (2015). Assessing electronic cigarette-related tweets for sentiment and content using supervised machine learning. *Journal of Medical Internet Research, 17*(8), e208. doi:10.2196/jmir.4392

Corazza, O., Assi, S., Simonato, P., Corkery, J., Bersani, F. S., Demetrovics, Z., . . . Schifano, F. (2013). Promoting innovation and excellence to face the rapid diffusion of novel psychoactive substances in the EU: The outcomes of the ReDNet project. *Human Psychopharmacology: Clinical & Experimental, 28,* 317–323. doi:10.1002/hup.2299

Daniulaityte, R., Chen, L., Lamy, F. R., Carlson, R. G., Thirunarayan, K., & Sheth, A. (2016). "When 'bad' is 'good'": Identifying personal communication and sentiment in drug-related tweets. *JMIR Public Health and Surveillance, 2*(2), e162. doi:10.2196/publichealth.6327

Daniulaityte, R., Nahhas, R. W., Wijeratne, S., Carlson, R. G., Lamy, F. R., Martins, S. S., . . . Sheth, A. (2015). "Time for dabs": Analyzing Twitter data on marijuana concentrates across the U.S. *Drug and Alcohol Dependence, 155,* 307–311. doi:10.1016/j.drugalcdep.2015.07.1199

Hanson, C. L., Burton, S. H., Giraud-Carrier, C., West, J. H., Barnes, M. D., & Hansen, B. (2013). Tweaking and tweeting: Exploring Twitter for nonmedical use of a psychostimulant drug (Adderall) among college students. *Journal of Medical Internet Research, 15*(4), e62. doi:10.2196/jmir.2503

Hommel, G. (1988). A stagewise rejective multiple test procedure based on a modified Bonferroni test. *Biometrika, 75,* 383–386. doi:10.1093/biomet/75.2.383

Internet Live Stats. (2016). *Twitter usage statistics.* Retrieved from http://www.internetlivestats.com/twitter-statistics/

Kim, Y., Huang, J., & Emery, S. (2016). Garbage in, garbage out: Data collection, quality assessment and reporting standards for social media data use in health research, infodemiology and digital disease detection. *Journal of Medical Internet Research, 18*(2), e41. doi:10.2196/jmir.4738

Kleiman, M. A. R. (2015). *Legal commercial cannabis sales in Colorado and Washington: What can we learn?* Retrieved from https://www.brookings.edu/wp-content/uploads/2016/07/Kleiman-Wash-and-Co-final.pdf

Lamy, F. R., Daniulaityte, R., Sheth, A., Nahhas, R. W., Martins, S. S., Boyer, E. W., & Carlson, R. G. (2016). Those edibles hit hard: Exploration of Twitter data on cannabis edibles in the U.S. *Drug and Alcohol Dependence, 164,* 64–70. doi:10.1016/j.drugalcdep.2016.04.029

Morstatter, F., Pfeffer, J., Liu, H., & Carley, K. M. (2013). *Is the sample good enough? Comparing data from Twitter's streaming API with Twitter's firehose.* arXiv:1306.5204.

Mounteney, J., & Haugland, S. (2009). Earlier warning: A multi-indicator approach to monitoring trends in the illicit use of medicines. *International Journal on Drug Policy, 20,* 161–169. doi:10.1016/j.drugpo.2007.09.006

Myslín, M., Zhu, S. H., Chapman, W., & Conway, M. (2013). Using Twitter to examine smoking behavior and perceptions of emerging tobacco products. *Journal of Medical Internet Research, 15*(8), e174. doi:10.2196/jmir.2534

National Institute on Drug Abuse. (2014). *DrugFacts: Marijuana.* Retrieved from https://www.drugabuse.gov/publications/drugfacts/marijuana

NDEWS. (2016). *National Drug Early Warning System.* Retrieved from https://ndews.umd.edu/

Oregon Health Authority. (2015). *Oregon Medical Marijuana Program.* Retrieved from http://www.oregon.gov/oha/PH/DiseasesConditions/ChronicDisease/MedicalMarijuanaProgram/Pages/index.aspx

Pacula, R. L., Powell, D., Heaton, P., & Sevigny, E. L. (2015). Assessing the effects of medical marijuana laws on marijuana use: The devil is in the details. *Journal of Policy Analysis and Management, 34,* 7–31. doi:10.1002/pam.21804

Room, R. (2014). Legalizing a market for cannabis for pleasure: Colorado, Washington, Uruguay and beyond. *Addiction, 109,* 345–351. doi:10.1111/add.12355

Sheth, A., Jadhav, A., Kapanipathi, P., Lu, C., Purohit, H., Smith, G. A., & Wang, W. (2014). Twitris: A system for collective social intelligence. In R. Alhajj & J. Rokne (Eds.), *Encyclopedia of social network analysis and mining* (pp. 2240–2253). New York, NY: Springer. doi:10.1007/978-1-4614-6170-8_345

Shutler, L., Nelson, L. S., Portelli, I., Blachford, C., & Perrone, J. (2015). Drug use in the Twittersphere: A qualitative contextual analysis of Tweets about prescription drugs. *Journal of Addictive Diseases, 34,* 303–310. doi:10.1080/10550887.2015.1074505

Team, R. C. (2016). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org

Thompson, L., Rivara, F. P., & Whitehill, J. M. (2015). Prevalence of marijuana-related traffic on Twitter, 2012–2013: A content analysis. *Cyberpsychology, Behavior, and Social Networking, 18,* 311–319. doi:10.1089/cyber.2014.0620

Twitter. (2016). *About the company.* Retrieved from https://about.twitter.com/company

Williams, A. R., Olfson, M., Kim, J. H., Martins, S. S., & Kleber, H. D. (2016). Older, less regulated medical marijuana programs have much greater enrollment rates than newer medicalized programs. *Health Affairs (Project Hope), 35,* 480–488. doi:10.1377/hlthaff.2015.0528

Yin, P., Ram, N., Lee, W., Tucker, C., Khandelwal, S., & Salathe, M. (2014, May). *Two sides of a coin: Separating personal communication and public dissemination accounts in Twitter.* Paper presented at the 18th Pacific-Asia Conference on Knowledge Discovery and Data Mining, Tainan, Taiwan.