CrossMark

ORIGINAL RESEARCH ARTICLE

Social Media Mining for Toxicovigilance: Automatic Monitoring of Prescription Medication Abuse from Twitter

Abeed Sarker[1] · Karen O'Connor[1] · Rachel Ginn[1] · Matthew Scotch[1,2] ·
Karen Smith[3] · Dan Malone[4] · Graciela Gonzalez[1]

Abstract

Introduction Prescription medication overdose is the fastest growing drug-related problem in the USA. The growing nature of this problem necessitates the implementation of improved monitoring strategies for investigating the prevalence and patterns of abuse of specific medications.

Objectives Our primary aims were to assess the possibility of utilizing social media as a resource for automatic monitoring of prescription medication abuse and to devise an automatic classification technique that can identify potentially abuse-indicating user posts.

Methods We collected Twitter user posts (tweets) associated with three commonly abused medications (Adderall®, oxycodone, and quetiapine). We manually annotated 6400 tweets mentioning these three medications and a control medication (metformin) that is not the subject of abuse due to its mechanism of action. We performed quantitative and qualitative analyses of the annotated data to determine whether posts on Twitter contain signals of prescription medication abuse. Finally, we designed an automatic supervised classification technique to distinguish posts containing signals of medication abuse from those that do not and assessed the utility of Twitter in investigating patterns of abuse over time.

Results Our analyses show that clear signals of medication abuse can be drawn from Twitter posts and the percentage of tweets containing abuse signals are significantly higher for the three case medications (Adderall®: 23 %, quetiapine: 5.0 %, oxycodone: 12 %) than the proportion for the control medication (metformin: 0.3 %). Our automatic classification approach achieves 82 % accuracy overall (medication abuse class recall: 0.51, precision: 0.41, *F* measure: 0.46). To illustrate the utility of automatic classification, we show how the classification data can be used to analyze abuse patterns over time.

Conclusion Our study indicates that social media can be a crucial resource for obtaining abuse-related information for medications, and that automatic approaches involving supervised classification and natural language processing hold promises for essential future monitoring and intervention tasks.

✉ Abeed Sarker
abeed.sarker@asu.edu

1   Department of Biomedical Informatics, Arizona State University, Scottsdale, AZ, USA

2   Center for Environmental Security, Biodesign Institute, Arizona State University, Tempe, AZ, USA

3   Rueckert-Hartman College for Health Professions, Regis University, Denver, CO, USA

4   Department of Pharmacy Practice and Science, University of Arizona, Tucson, AZ, USA

Key Points

Monitoring prescription medication abuse, which is a rapidly growing medication-related problem in the USA, is of paramount importance to public health.

Social media postings can be used to detect patterns and intents of abuse and also to estimate the prevalence of abuse for a drug.

Natural language processing and machine learning can be applied to automatically detect posts indicating prescription medication abuse, allowing interested agencies to perform real-time monitoring and analysis of medication abuse information.

△ Adis

1 Introduction

Prescription medication abuse and overdose have become the fastest growing medication-related problems in the USA, reaching epidemic proportions [1–3]. According to the Drug Abuse Warning Network (DAWN) report [3] there were 5.1 million drug-related (prescription and illicit) emergency department visits in the year 2011, with approximately half of these attributed to misuse or abuse. The negative health consequences of prescription medication abuse are many, ranging from nausea to disorientation, paranoia, seizures, and even death [2]. A recently published Centers for Disease Control (CDC) report [4] shows that in the year 2013 there were 43,982 deaths due to drug overdose, of which 22,767 were attributable to prescription medications such as opioids and benzodiazepines (some were due to co-ingestion [5]). The number of deaths due to drug overdose has more than tripled in the USA since 1991 [6]. In the year 2007, opioid abuse alone amounted to an estimated total cost of US$55.7 billion [7] and later estimates suggest that medication misuse costs up to US$72.5 billion annually [8–10]. Despite such widespread abuse of prescription medication, there is no current well-established source of data to monitor abusers' attitudes, methods, and patterns of abuse. Because of the severe nature of this problem, the Office of National Drug Control Policy released a drug abuse prevention plan [11], which puts monitoring as one of the primary areas of focus.

Current prescription medication abuse monitoring strategies are aimed primarily at distributors and licensed practitioners. The Drug Enforcement Agency (DEA) requires that wholesalers have monitoring programs in place to identify suspicious orders. For licensed, prescribing health practitioners, most states have monitoring programs that are supported by the National Association of State Controlled Substance Authorities and pharmacies are required to report the patients, prescribers, and specific medications dispensed for controlled substances. These data are used by prescribers and law enforcement agencies to identify and limit medication abuse. However, existing control measures lack critical information such as the patterns of usage of various medications and the demographic information of the users. For example, advertisements to deter prescription medication abuse might be more successful if broadcast during high abuse periods, if that information were available. Thus, there is a strong motivation to discover newer monitoring sources and methods.

We take the first steps towards operationalizing an automated, social media-based, medication abuse monitoring system. Social media has evolved into a crucial source of communication and it offers a range of possibilities for establishing multi-directional interaction, as well as acting as a resource for monitoring public sentiment and activity [12]. Twitter [13] is one of the most popular resources, with 289 million monthly active users, 58 million tweets per day, and 2.3 billion search queries per day [14]. Social media is currently being used as a resource for various tasks ranging from customized advertising [15] and sentiment analysis [16] to tasks specific to the public health domain (e.g., monitoring influenza epidemics [17], sexual health [18], pharmacovigilance [19], and drug abuse [3]). Our aims are to (1) verify that Twitter posts contain identifiable information about prescription medication abuse; (2) annotate abuse-indicating posts and create an annotation guideline for large-scale future research; and (3) verify the future applicability of automatic systems for social media-based medication abuse monitoring. We take a systematic approach towards our end goal following on from our past work on social media-based pharmacovigilance [20–22].

2 Methods

Figure 1 summarizes our methodological pipeline. The figure is divided into four parts, each indicating the subsection in which the associated study component is discussed. From a high level, the whole process may be grouped into four steps: (1) data collection; (2) annotation; (3) classification; and (4) analysis. We outline each subprocess in the rest of this section.

2.1 Data Collection and Spelling Variant Generation

We selected three abuse-prone medications (APMs) based on their reported abuse potential. Two, Adderall® (amphetamine-mixed-salts) and oxycodone, are well-known, while the third, quetiapine, is less widely known as an abused medication but is gaining popularity [23]. Table 1 presents key information about these three medications. We collected data associated with these medications by querying the Twitter Streaming API, which provides real-time access to a subset of all tweets being posted. For Adderall®, we used the trade name for querying. For quetiapine and oxycodone, we used the generic names and also their popular trade names OxyContin® and Seroquel®, respectively. Since Twitter users frequently misspell medication names, we applied a phonetic spelling generator to generate common misspellings [24]. We collected data from March 2014 to June 2015 and included a control medication for comparison, the oral diabetes drug metformin, as it does not have potential for abuse.
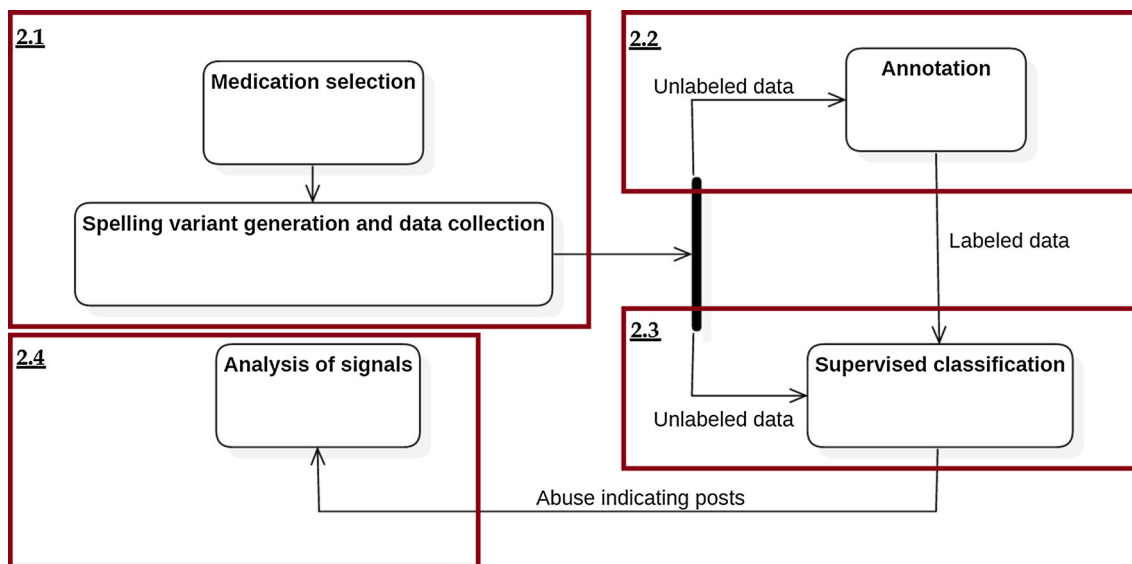
Fig. 1 Pipeline for monitoring prescription medication abuse signals from Twitter. The four components of the pipeline are discussed in Sects. 2.1, 2.2, 2.3 and 2.4

Table 1 The three abuse-prone drugs included in the study, their types, and details of common abuse

Drug	Type	Abuse details
Adderall® (amphetamine-mixed-salts)	Schedule II controlled substance. A combination psychostimulant drug used for the treatment of ADHD and narcolepsy. Produces euphoria, alertness, and increased concentration	Adderall® abuse includes intake for weight loss, increased focus (e.g., for studying), athletic performance, and recreational use
Oxycodone (OxyContin®)	Schedule II controlled substance (opioid family). It is an opioid agonist that produces analgesia through its effect on the μ-receptor and contributes to addiction through its effect on dopamine receptors	In 2012, it is estimated that narcotic pain relievers were abused by 2.1 million people in the USA [50]
Quetiapine (Seroquel®)	An atypical antipsychotic generally used for the treatment of schizophrenia or bipolar disorder. It may produce euphoria in addition to its anxiolytic properties	Known to be less prone to abuse than some other drugs, but has been named in the top 10 list of abused prescription medications [23]

Note that tweets for Adderall® were collected using the trade name only, while both generic and trade names were used for oxycodone and seroquel

ADHD attention-deficit hyperactivity disorder

2.2 Annotation

We randomly selected 2600 user posts mentioning Adderall®,[1] 1600 each for oxycodone and quetiapine, and 600 mentioning metformin for annotation. Two annotators performed binary classification of 6400 tweets to indicate abuse or non-abuse. Both intent to abuse and actual abuse were considered to be abuse for the purposes of annotation and only personal experiences, rather than generic statements, were considered. Because of the short nature of the posts on Twitter, it was often difficult to determine if a tweet indicated abuse or prescription use. Tweets were annotated as abuse-indicating if they did not mention

clinical diagnoses but did describe potential use for non-medical benefits. Table 2 provides some examples of the annotated tweets.

We used Cohen's kappa [25] to measure the inter-annotator agreement, and obtained substantial agreement ($\kappa = 0.78$). Figure 2 illustrates the final distributions for the proportions of abuse-indicating versus non-abuse tweets for each of the four medications. Adderall® has a much higher proportion of abuse-related Tweets than the other medications. In contrast, the control medication, metformin, had only two annotated abuse mentions. These were judged to be noise during our review of the annotations, enabling us to estimate the amount of noise in non-abuse data. Annotator disagreements were resolved by the study's pharmacology expert (KS). Because of the

---

[1] A larger number of Adderall® tweets were chosen for annotation because of the significantly larger number of tweets available for it.

Table 2 Examples of abuse and non-abuse-indicating tweets from our dataset

Non-medical use/abuse-indicating tweets	Non-abuse tweets
about to be cracked on adderall to survive today	Seroquel is prescribed. i use valerian root sometimes too. mostly i don't sleep
i'm just gonna shower and overdose on Seroquel so I'll sleep until morning	a prescription for adderall should come with my college acceptance paper
popped Adderall tonight hahahah let's finish this 100 page paper	speaking of oxycodone .. i need to take mine. This pain is ridiculous
an oxycodone high from snorting lasts for one hour, if it is swallowed, your looking at three hour high	
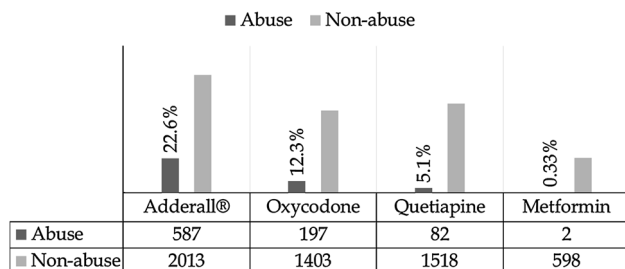
Fig. 2 Distributions of abuse/non-abuse tweets for the four drugs. The numbers and percentages of abuse-indicating tweets for each drug are also shown

relatively complex decision making required to decide if a tweet contains abuse information or not, and in the interest of comparative research, our evolving annotation guideline for this study is available to view [26].[2]

Following the annotation, in order to ascertain if the proportions of abuse tweets for the APMs are different to the proportion among the metformin tweets, we performed Chi squared tests (H_0: $p_1 = p_2$, i.e., the proportions of abuse tweets are equal for each APM and metformin). In all cases, the null hypothesis was rejected ($p < 0.05$), verifying that the APMs have significantly higher proportions of abuse-indicating tweets.

2.3 Supervised Classification

We model the task of detecting potentially abuse-indicating tweets as a supervised classification problem and use a number of features to address it. From the perspective of machine learning, classification is the problem of assigning a category, from a finite set of categories, to an observation. In our work, each tweet is an observation and there are two possible categories. Since we have a set of annotated tweets, we use them to train several machine learning algorithms. In this manner, we provide supervision to our algorithms. Given an unclassified observation, our

approach generates a set of features from the observation and, using the model built on the training data, assigns a category to the new observation. We describe our feature sets in the following subsections.

2.3.1 Word n-Grams

A word n-gram is a sequence of contiguous n words in a text segment. Prior to generating the n-grams (1-, 2-, and 3-grams), we preprocessed the tweets to remove usernames (e.g., @xxxx) and hyperlinks, stemmed (i.e., converted words to their base forms: controlled, controlling → control) the tweets using the Porter stemmer [27] and lowercased them. We used a specialized tokenizer and part-of-speech tagger [28] to break down posts into words and identify their part-of-speech tags. We used the generated word tokens to build the n-grams. Because of the short nature of the tweets, the number of n-grams generated is relatively small. Hence, we attempted to generate a set of additional features from the texts, which encapsulate deeper semantic information.

2.3.2 Abuse-Indicating Terms

We collected terms from Hanson et al. [3, 29] that were used to search for indications of alternative motive, co-ingestion, frequent doses, and alternative routes of admission. As they may indicate medication abuse, we generated two features based on these terms—presence and counts of occurrences.

2.3.3 Drug–Slang Lexicon

We collected a set of colloquial phrases and terms that are used to discuss drug-related content on the internet [30]. We again used presence and counts as features, leaving out terms that are five characters or less in length (to ignore ambiguous terms: e.g., A, ace).

2.3.4 Synonym Expansion

Our analysis of the tweets during annotation suggested that users often express polarized sentiments when mentioning

---

[2] The guideline also contains a set of ambiguous tweets identified by our pharmacology expert, which require contextual analysis in the future.

Table 3 Tenfold cross-validation results showing *F* scores for the two classes and the overall accuracies

Classifier	Abuse *F* score	Non-abuse *F* score	Accuracy (%)
Naïve Bayes	0.39	0.84	75
Weighted support vector machine (wSVM)	0.45	0.89	81
Maximum entropy	0.24	0.85	75
J48	0.22	0.92	85
Stacking	0.46	0.89	82

drug abuse. Therefore, we introduced a feature that has been used in the past for sentiment analysis and polarity classification [31]. This feature is generated by identifying all synonyms of all nouns, verbs and adjectives using WordNet [32]. Synonym expansion enables the generation of a large number of sentiment-indicating terms from a short post.

2.3.5 Word Clusters

Recent text classification tasks have benefited from the use of word clusters, which are broad semantic groupings of words learnt automatically from large amounts of unlabeled data, as features [33]. Therefore, we used the Twitter word clusters provided by Nikfarjam et al. [22]. The clusters were learnt from medication-related chatter collected for monitoring adverse drug reactions and each cluster attempts to encapsulate terms that are generally used in similar contexts. For each term, we used its cluster number as a nominal feature.

2.3.6 Classification

We used four off-the-shelf supervised classification algorithms to assess the performance of automatic detection, namely: Naïve Bayes, Support Vector Machines (SVMs), Maximum Entropy (ME), and a decision tree-based classifier (J48). We used the publicly available LibSVM [34] and Weka [35] tools for these classifiers. Following the initial classification, we focused on exploring optimization techniques to improve the classification performance particularly for the positive (abuse) class, which is especially difficult due to the class imbalance (i.e., the number of negative examples is much greater than the number of positive examples). To bias our classifiers towards choosing the positive class, we applied weights to give preference to this class [34]. Finally, all four classifiers were combined via stacking [36], a technique where the predictions from the different classifiers are combined and another algorithm is trained to make a final decision based on the individual predictions.

Using the best classification system, we performed two additional sets of experiments to (1) analyze the feature contributions; and (2) estimate how more data will affect performance. Results are presented in Sect. 3.

2.4 Analysis of Signals

To verify the utility of automatic classification and possible future applications, we performed an analysis of abuse patterns over time for Adderall® and oxycodone. We first classified all of our collected tweets using the stacking-based classifier and then visualized the distributions of tweets over time. Details of this analysis are presented in Sect. 4.

3 Results

We collected 119,809 tweets for Adderall®, 4082 for oxycodone, 4505 for quetiapine, and 1052 for metformin. Our final annotated dataset contained 5799 annotated tweets (excluding the 600 metformin tweets, and one Adderall® tweet due to character encoding issues): 869 indicating abuse and 4930 with no indication of abuse.[3]

Table 3 presents the results of our automatic classification experiments using stratified tenfold cross-validations over the whole data and Fig. 3 presents the equations used to compute the scores.[4] The table shows that weighted SVMs have the highest *F* score for abuse among the individual classifiers, but marginally better results are observed when stacking is introduced.

In Table 4, we present the results of our two feature analysis experiments. The single-feature scores indicate how each feature performs individually, while the leave-out-feature scores indicate their contributions when combined with other features by illustrating changes in performance when they are removed from the combination. The most useful feature set is n-grams, as depicted by the two sets of results. Importantly, none of the single-feature scores are able to achieve the performance of all of the features combined. Synonym expansion and the drug–slang lexicon perform poorly when applied as single features. Despite the relative importance of word clusters in the single-feature

---

[3] A sample of the finalized annotations will be made available to the research community at http://diego.asu.edu/Publications/DrugAbuse_DrugSafety.html.

[4] Note that the overall accuracy is primarily driven by the classification performance on the larger class (i.e., non-abuse class) because of the significantly larger number of instances.

$$p = \frac{tp}{tp + fp} \qquad r = \frac{tp}{tp + fn}$$

$$f = \frac{2 \times p \times r}{p + r} \qquad a = \frac{tp + tn}{tp + fn + fp + tn}$$

key:
p: precision, r: recall,
f: F-score, a: overall accuracy,
tp: # true positives, fp: # false positives,
fn: # false negatives, tn: # true negatives

Fig. 3 Equations for the classifier evaluation metrics. Accuracy is the combined accuracy for the two classes, while the other three scores are computed per class. *a* overall accuracy, *f* F score, *fn* number of false negatives, *fp* number of false positives, *p* precision, *r* recall, *tn* number of true negatives, *tp* number of true positives

Table 4 Single-feature and leave-out-feature experiments showing the impact of each of the five feature sets on classification *F* scores (abuse class)

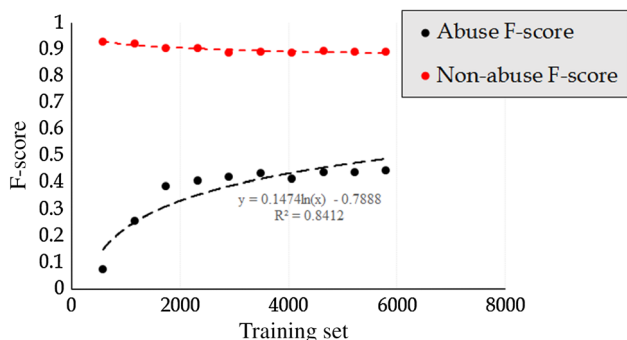| Feature	Single-feature *F* score	Leave-out-feature *F* score
N-grams	0.42	0.37
Abuse-indicating terms	0.30	0.41
Drug–slang lexicon	0.07	0.44
Syn-sets	0.09	0.44
Word clusters	0.35	0.43

Fig. 4 Classification performances for training data of different sizes

classifications, the performance of the classifier when this feature is left out does not drop significantly.

Figure 4 illustrates how the classification performance varies with the training set size. The key information illustrated by the figure is that the abuse *F* score improves, as training data are added, in an approximately logarithmic fashion. Based on the logarithmic trend-line, one can predict that if 10,000+ annotated training instances are used, *F* scores of over 0.55 can be achieved for the abuse class without the use of any additional features. The non-abuse *F* scores and the accuracy scores are not significantly affected by the size of the training set, which is unsurprising considering the data imbalance issue.

4 Discussion

Our goals in this study were to verify that social media contains information indicating prescription medication abuse, prepare an annotation guideline for the annotation of

Twitter data, and use a moderate sample of annotations to investigate if automatic classification approaches can be used to automatically detect information about prescription medication abuse from Twitter. Our experiments strongly suggest that social networks such as Twitter may provide valuable information about medication abuse. While the task of annotation of social network data is time-consuming and ambiguous, the availability of annotated data allows the training of systems that can identify abuse-related information from large volumes of social media data. Our trained system is capable of classifying abuse and non-abuse tweets with fair accuracy and its performance is likely to improve significantly with the availability of larger amounts of annotated data.

In this section we first outline some related work in order to put our contributions into context. We then verify the utility of automatic classification approaches for medication abuse monitoring. Finally, we present a brief error analysis to identify essential future improvements and discuss some of the limitations of this study.

4.1 Related Work

Toxicovigilance involves the active detection, validation, and follow-up of clinical adverse reactions related to toxic exposures, which may be caused by the non-medical use of prescription medications. Traditional tools for toxicovigilance include federally sponsored surveys and reports from the National Poisoning Data System (NPDS) [37], the US Food and Drug Administration (FDA) [38], and hospital networks such as DAWN. The NPDS provides data about calls that are placed to poison centers nationwide and, according to their website, the information may be used to

track the risks of prescription medication abuse. The FDA MedWatch program [39] provides information about adverse drug events but does not monitor patterns of abuse actively (although there is much discussion around opioids and, in the last year, hydrocodone has moved to a schedule II drug). In contrast, DAWN primarily reports emergency department visit data. Although these systems all provide valuable data, there is a gap in determining prevalence (no denominator) and monitoring patterns of and attitudes towards abuse, which is one contribution social media can provide.

The first effort to propose the use of social media for monitoring the prevalence of drug abuse was the Psychonaut project [40] and recent studies have verified that it operates as a channel for exchanging prescription medication abuse information [41]. Chary et al. [6] discussed the benefits of using social media for toxicovigilance, relative to traditional tools, and suggested the use of text mining techniques, but did not report conducting any evaluations using these techniques. Recent efforts have attempted to develop web-based platforms and ontologies for social media mining for toxicovigilance [42]. Also, social media data have been used to perform targeted toxicovigilance tasks such as analyzing the effects of drug reformulation [43] or the phases of drug abuse recovery [44] from specialized forums.

Some recent studies have focused specifically on utilizing data from Twitter. Cavazos-Rehg et al. [45] demonstrated that social media can influence the behavior of young people, necessitating the need for surveillance efforts to monitor content on Twitter. Scott et al. [46] and Hanson et al. [29] presented similar conclusions and suggested that Twitter serves as a potential resource for not only surveillance but also for studying attitudes towards prescription drug abuse. Hanson et al. [3] studied the use of Adderall®, concluding that through the analysis of Twitter posts it can be verified that the medication is frequently abused as a study aid by college students. Similarly, Shutler et al. [47] qualitatively studied opioid abuse using a small set of annotated data. Very recently, Coloma et al. [48] studied the potential of social media in drug safety surveillance. The study concluded, following an elaborate manual analysis, that further research is required for establishing social media as a reliable resource. Importantly, this and other recent studies on this topic do not utilize the most attractive property of social media—massive volumes of data and automatic processing.

4.2 Utility of Automatic Monitoring

We classified all Adderall® tweets collected from March 2014 to March 2015 and all oxycodone tweets from June 2014 to June 2015.[5] Figure 5 presents the distribution of all tweets and abuse-indicating tweets for the two medications, and also the proportions of abuse-indicating tweets. Monthly usage patterns are different for the two medications, the only similarity being peaks around the holiday season (December). Furthermore, the distribution for Adderall® agrees with the manual analysis in Hanson et al. [3] and clearly illustrates the high number of tweets close to traditional exam times (i.e., November/December and April/May). Interestingly, Fig. 5b suggests that the general trend for proportions is increasing for Adderall® and decreasing for oxycodone, although further analysis is essential before drawing such conclusions. Analysis over longer periods of time may reveal other interesting insights.

These results verify that automatic social media-based monitoring of prescription medication abuse can play a role in identifying patterns of prescription medication usage. First, the numbers and proportions of abuse-indicating tweets over a period of time give us an indication of the extent of abuse for a specific medication. In addition, deeper analysis, manual or otherwise, of a set of classified tweets may reveal potentially decisive information. Other information, such as demographics of users, can also be extracted via automatic techniques for deeper analysis.

4.3 Error Analysis

In Table 5, we show that there are a large number of false positives for Adderall®, but the reverse is true for the other medications. Our analysis identified some key reasons behind these numbers. For Adderall®, a large number of tweets are impersonal, without any direct association to the user. Although these were annotated as non-abuse, due to the n-gram patterns and the presence of keywords, the classifiers are often unable to accurately identify them. Tweets mentioning other medications face the same issue, but to a lesser extent. For quetiapine and oxycodone, the high numbers of false negatives are caused by the lack of sufficient abuse-indicating tweets in our training sample. This problem can easily be addressed by incorporating more data. In many cases, the tweets contained too many non-standard spellings, which resulted in misclassifications (generally, these contributed to false negatives). Non-standard spellings also impact the generation of deep linguistic features (e.g., synonyms). We identified these three reasons (impersonal tweets, small training set, and non-standard spellings) to be the primary causes of misclassifications. In addition, we also observed during the analysis that a significant number of correctly classified tweets contained evidence of co-ingestion (e.g., with marijuana,

---

[5] Slightly different periods are used for the two medications as the collection of oxycodone tweets commenced a little later than the other medications.
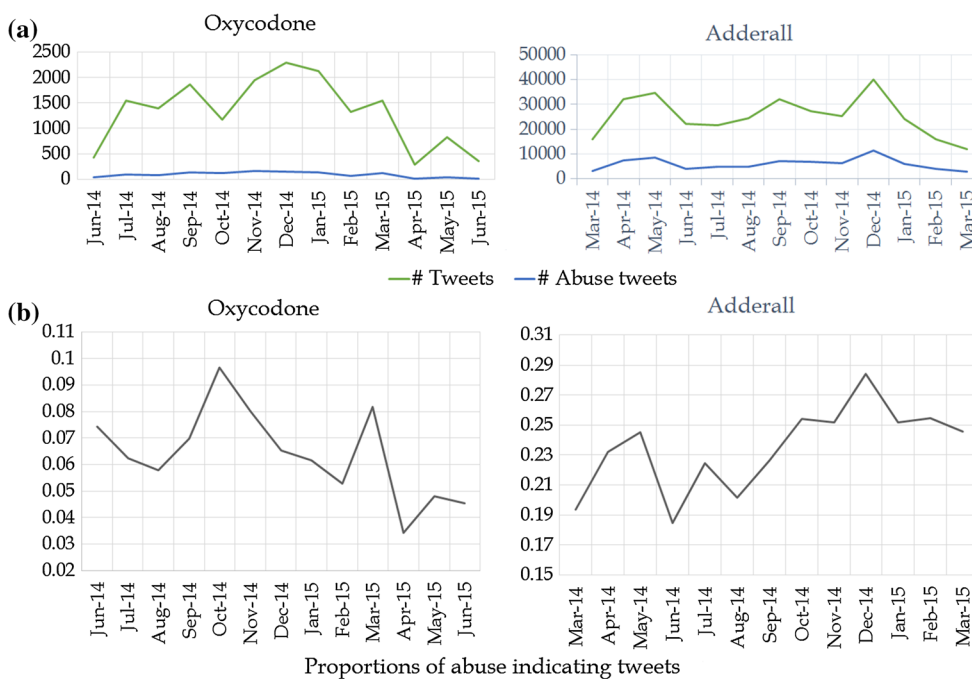
Fig. 5 a Distributions of all collected tweets and automatically detected abuse-indicating tweets for Adderall® and oxycodone and **b** the proportions of abuse-indicating tweets over the same time periods

Table 5 False negative and positive tweets for abuse classification with the best classification system (stacking based)

Drug	False positive	False negative
Adderall®	559	197
Quetiapine	28	63
Oxycodone	95	108

coffee, and/or prescription medications). With Twitter data indicating that there are significant levels of co-ingestion-related abuse of these medications, the effects of such

intake should perhaps be studied in greater detail. Table 6 presents some sample misclassified tweets and some correctly classified tweets indicating co-ingestion.

4.4 Limitations

As already mentioned, the biggest limitation of our study, from the perspective of supervised learning, is the lack of sufficient training data. This is particularly true for the abuse class, which has less than 1000 instances. Ambiguity in tweets and the lack of context hinders both annotation and automatic classification. A number of tweets were

Table 6 Examples of tweets that are difficult to classify (false positives and false negatives) and examples of co-ingestion

False positives	False negatives	True positives (suggestions of co-ingestion)
they need to make armpit tampons for adderall abusers	if she's craving ecstasy, oxycodone to, and is tuliao you're her, she, and is pending mine for one day. maybe	john picked us up from the airport with medicinal, blunts, adderall and booze. he was actually sent from the heavens
why would you want to take seroquel recreationally ***	i got some oxycodone who tryna buy it off me? $6 a pill	adderall to stay focused, xanax to take the edge off, pot to mellow me out, cocaine to wake me back up and morphine..well because its awesome
these are actually (half) sober tweets and i haven't slept yet thanks Adderall	took adderall thinking it'd make work go by faster now i'm not tired #paper-view	time for my daily afternoon relaxation ritual of smoking weed, taking 2 mgs of clonazepam, and 400 mg of seroquel xr
		hello i am looking to speedball some cocaine and adderall so i can complete a large online project if u have any leads hmu

Offensive terms have been censored

regarded to be ambiguous during annotations and these were resolved by our pharmacology expert. These ambiguities leave a gray zone in the binary classification process and this limitation will persist until future annotation guidelines are able to specify more fine-tuned annotation rules.

Because we use social media to collect abuse-related information, our study is limited to the population group that is adept at using social media. All public health monitoring studies that rely on social media data are faced with this limitation. However, as the rapid growth of social media continues and as current users grow older, the impact of this limitation will inevitably decrease.

5 Conclusion and Future Directions

We investigated the potential of social media as a resource for monitoring medication abuse patterns. We first verified that social media contains identifiable abuse-indicating information for three commonly abused medications and that these are significantly more prominent than potential noise associated with non-abuse medications. We prepared annotation guidelines and annotated a moderate-sized dataset. Using supervised classification, we obtained classification accuracies of 82 % (positive class F score: 0.46). We identified three key factors that are likely to improve classification performance: identification of personal versus impersonal tweets, normalization of tweets (e.g., spelling correction), and the utilization of more training data. To show the utility of automatic approaches for toxicovigilance, we performed an analysis of usage patterns and showed that our conclusions agree with past manual analyses.

Based on the promising results obtained, we will focus strongly on the following three research problems in the future:

1. Annotation. We will significantly increase the size of our annotated dataset and perform iterative annotations to reduce ambiguities.
2. Visualization and real-time monitoring. Figure 5 is an example of how medication abuse information can be visualized. We intend to implement web-based visualization tools for real-time monitoring of the social media sphere.
3. Natural language processing-oriented improvements. To improve the reliability of our methods, we will explore natural language processing techniques to automatically identify non-personal tweets. We will also attempt to perform lexical normalization as a preprocessing step to correct spelling errors. While some work has been done on social media text normalization [49], there are no available techniques that are customized to specific types of social media text (e.g., health-related social media text).

We believe that our automatic approaches will have significant importance for various toxicovigilance tasks, including, but not limited to, determining the prevalence of abuse, studying medication abuse patterns, and identifying the impacts of control measures.

Compliance with Ethical Standards

Conflict of interest Abeed Sarker, Rachel Ginn, Karen O'Connor, Matthew Scotch, Karen Smith, Dan Malone, and Graciela Gonzalez have no conflicts of interest that are directly relevant to the content of this study.

Ethical approval Not applicable.

Informed consent Not applicable.

References

1. Young S. White House launches effort to combat soaring prescription drug abuse. CNN 2011. http://www.cnn.com/2011/HEALTH/04/19/drug.abuse/. Accessed 15 Nov 2015.
2. The DAWN Report. Highlights of the 2011 Drug Abuse Warning Network (DAWN) findings on drug-related emergency department visits. 2013. http://archive.samhsa.gov/data/2k13/DAWN127/sr127-DAWN-highlights.htm. Accessed 15 Nov 2015.
3. Hanson CL, Cannon B, Burton S, Giraud-Carrier C. An exploration of social circles and prescription drug abuse through twitter. J Med Internet Res. 2013;15(9):e189.
4. Centers for Disease Control and Prevention: National Vital Statistics System. Mortality data. 2015. http://www.cdc.gov/nchs/deaths.htm. Accessed 15 Nov 2015.
5. Centers for Disease Control and Prevention. Prescription drug overdose in the United States: fact sheet. 2015. http://www.cdc.gov/drugoverdose/index.html. Accessed 15 Nov 2015.
6. Chary M, Genes N, McKenzie A, Manini AF. Leveraging social networks for toxicovigilance. J Med Toxicol. 2013;9(2):184–91.
7. Birnbaum HG, White AG, Schiller M, Waldman T, Cleveland JM, Roland CL. Societal costs of prescription opioid abuse, dependence, and misuse in the United States. Pain Med. 2011;12(4):657–67.

8. Centers for Disease Control and Prevention. Prescription pain-killer overdoses in the US. 2015. http://www.cdc.gov/vitalsigns/painkilleroverdoses/. Accessed 15 Nov 2015.

9. Jena AB, Goldman DP. Growing Internet use may help explain the rise in prescription drug abuse in the United States. Health Aff (Millwood). 2011;30(6):1192–9.

10. Wu LT, Pilowsky DJ, Patkar AA. Non-prescribed use of pain relievers among adolescents in the United States. Drug Alcohol Depend. 2008;94(1–3):1–11.

11. Office of National Drug Control Policy (ONDCP). Responding to America's prescription drug abuse crisis. 2011. https://www.whitehouse.gov/sites/default/files/ondcp/issuescontent/prescription-drugs/rx_abuse_plan.pdf. Accessed 15 Nov 2015.

12. Capurro D, Cole K, Echavarría MI, Joe J, Neogi T, Turner AM. The use of social networking sites for public health practice and research: a systematic review. J Med Internet Res. 2014;16(3):e79.

13. Twitter. Company. https://about.twitter.com/company. Accessed 21 May 2015.

14. Statistic Brain Research Institute. Twitter statistics. http://www.statisticbrain.com/twitter-statistics/. Accessed 19 Nov 15.

15. Kaplam AM, Haenlein M. Users of the world, unite! The challenges and opportunities of Social Media. Bus Horiz. 2010;53(1):59–68.

16. Pang B, Lee L. Opinion mining and sentiment analysis. Found Trends® Inf Retr. 2008;2(1–2):1–135.

17. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. Nature. 2009;457(7232):1012–4.

18. Gold J, Pedrana AE, Sacks-Davis R, Hellard ME, Chang S, Howard S, et al. A systematic examination of the use of online social networking sites for sexual health promotion. BMC Public Health. 2011;11:583.

19. Freifeld CC, Brownstein CC, Menone CM, Bao W, Felice R, Kass-Hout T, et al. Digital drug safety surveillance: monitoring pharmaceutical products in Twitter. Drug Saf. 2014;37(5):343–50.

20. Sarker A, Gonzalez G. Portable automatic text classification for adverse drug reaction detection via multi-corpus training. J Biomed Inform. 2015;53:196–207.

21. Sarker A, Ginn R, Nikfarjam A, O'Connor K, Smith K, Jayaraman S, et al. Utilizing social media data for pharmacovigilance: a review. J Biomed Inform. 2015;54:202–12.

22. Nikfarjam A, Sarker A, O'Connor A, Ginn R, Gonzalez G. Pharmacovigilance from social media: mining adverse drug reaction mentions using sequence labeling with word embedding cluster features. J Am Med Inform Assoc. 2015;22(3):671–81.

23. Sansone RA, Sansone LA. Is seroquel developing an illicit reputation for misuse/abuse? Psychiatry (Edgmont). 2010;7(1):13–6.

24. Pimpalkhute P, Patki A, Nikfarjam A, Gonzalez G. Phonetic spelling filter for keyword selection in drug mention mining from social media. AMIA Jt Summits Transl Sci Proc. 2014;2014:90–5.

25. Carletta J. Assessing agreement on classification tasks: the kappa statistic. Comput Linguist. 1996;22(2):249–54.

26. O'Connor K, Ginn R, Smith K, Sarker A. Toxicovigilance from social media: annotation guidelines. Arizona State University. Version 1.1. July 2015. http://diego.asu.edu/guidelines/DrugAbuseAnnotationGuideline1.1.pdf. Accessed 14 Dec 2015.

27. Porter MF. An algorithm for suffix stripping. Program. 1980;14(3):130–7.

28. Owoputi O, O'Connor B, Dyer C, Gimpel K, Schneider N, Smith NA. Improved part-of-speech tagging for online conversational text with word clusters. HLT-NAACL Proc. 2013;2013:380–90.

29. Hanson CL, Burton SH, Giraud-Carrier C, West JH, Barnes MD, Hansen B. Tweaking and tweeting: exploring Twitter for non-medical use of a psychostimulant drug (Adderall) among college students. J Med Internet Res. 2013;15(4):e62.

30. NoSlang.com Drug Slang Translator. Drug slang dictionary. http://www.noslang.com/drugs/dictionary.php. Accessed 7 May 2015.

31. Sarker A, Mollá D, Paris C. Automatic evidence quality prediction to support evidence-based decision making. Artif Intell Med. 2015;64(2):89–103.

32. Princeton University. WordNet: a lexical database for English. https://wordnet.princeton.edu/. Accessed 19 Nov 15.

33. Turian J, Ratinov L, Bengio Y. Word representations: a simple and general method for semi-supervised learning. ACL Proc. 2010;2010:384–94.

34. Chang CC, Lin CJ. LIBSVM: a library for support vector machines. ACM Trans Intell Syst Technol. 2011;2(3):1–27.

35. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH. The WEKA data mining software: an update. SIGKDD Explor. 2009;11(1):10–8.

36. Wolpert DHDH. Wolpert. Stacked generalization. Neural Netw. 1992;5(2):241–59.

37. National Poisoning Data System. http://www.aapcc.org/data-system/. Accessed 16 Dec 2015

38. US Food and Drug Administration (FDA). http://www.fda.gov/. Accessed 16 Dec 2015

39. US FDA. MedWatch: the FDA safety information and adverse event reporting program. http://www.fda.gov/Safety/MedWatch/. Accessed 8 Aug 2015.

40. Schifano F, Deluca P, Baldachhino A, Peltoniemi T, Scherbaum N, Torrens M, et al. Drugs on the web; the Psychonaut 2002 EU project. Prog Neuro-Psychopharmacol Biol Psychiatry. 2006;30(4):640–6.

41. Mackey K, Liang BA, Strathdee SA. Digital social media, youth, and nonmedical use of prescription drugs: the need for reform. J Med Internet Res. 2013;15(7):e143.

42. Cameron D, Smith GA, Daniulaityte R, Sheth AP, Dave L, Chen L, et al. PREDOSE: a semantic web platform for drug abuse epidemiology using social media. J Biomed Inform. 2013;46(6):985–97.

43. McNaughton C, Coplan PM, Black RA, Weber SE, Chilcoat HD, Butler SF. Monitoring of internet forums to evaluate reactions to the introduction of reformulated oxycontin to deter abuse. J Med Internet Res. 2014;16(5):e119.

44. MacLean D, Gupta S, Lembke A, Manning C, Heer J. Forum77: an analysis of an online health forum dedicated to addiction recovery. In: Proceedings of the 18th ACM conference on Computer-Supported Cooperative Work and Social Computing. New York: ACM; 2015. pp. 1511–1526.

45. Cavazos-Rehg P, Krauss M, Grucza R, Bierut L. Characterizing the followers and tweets of a marijuana-focused Twitter handle. J Med Internet Res. 2014;16(6):e157.

46. Scott KR, Nelson L, Meisel Z, Perrone J. Opportunities for exploring and reducing prescription drug abuse through social media. J Addict Dis. 2015;34(2–3):178–84.

47. Shutler L, Nelson LS, Portelli I, Blackford C, Perrone J. Drug use in the Twittersphere: a qualitative contextual analysis of tweets about prescription drugs. J Addict Dis. 2015;34(4):303–10.

48. Coloma PM, Becker B, Sturkeboom MC, van Mulligen EM, Kors JA. Evaluating social media networks in medicines safety surveillance: two case studies. Drug Saf. 2015;38(10):921–30.

49. Han B, Cook P, Baldwin T. Lexical normalization for social media text. ACM Trans Intell Syst Technol. 2013;1(1):1–27.

50. Substance Abuse and Mental Health Services Administration. Results from the 2012 National Survey on Drug Use and Health: summary of national findings. 2013.http://www.samhsa.gov/data/sites/default/files/NSDUHresultsPDFWHTML2013/Web/NSDUHresults2013.pdf. Accessed 15 Nov 2015.