Detecting Signals of Adverse Drug Reactions from Health Consumer Contributed Content in Social Media

Christopher C. Yang
College of Information Science and Technology
Drexel University
Philadelphia, PA 19104
Chris.Yang@drexel.edu

Haodong Yang
College of Information Science and Technology
Drexel University
Philadelphia, PA 19104

Ling Jiang
College of Information Science and Technology
Drexel University
Philadelphia, PA 19104

Xuning Tang
College of Information Science and Technology
Drexel University
Philadelphia, PA 19104

ABSTRACT
Adverse drug reactions are causing a substantial amount of hospital admissions and deaths, which cannot be underestimated. Although a great effort has been put on the pre-marketing review during pharmaceutical product development, it cannot identify all possible adverse drug reactions. Currently, post-marketing surveillance is conducted through centralized volunteering reporting systems. However, the reporting rate is low, which makes it difficult to detect the adverse drug reactions signals in a timely manner. With the advance of Web 2.0 technologies and the popularity of social media, many health consumers are discussing and exchanging health-related information with their peers. Many of this online discussion involve adverse drug reactions. In this work, we propose to mine the associations between drugs and adverse reactions from the user contributed content in social media. We have conducted an experiment using five drugs and five adverse drug reactions. The FDA alerts are used as the gold standard to test the performance of the proposed techniques. The result shows that the proposed technique is promising to detect the adverse drug reactions reported by FDA, such as diarrhea, heart condition, depression, and suicidal thoughts. However, adverse drug reaction such as cancer cannot be detected effectively.

Categories and Subject Descriptors
H.2.8 [Database Management]: Database applications – Data mining; H.3.1 [Information Storage and Retrieval]: Content Analysis and Indexing – Linguistic processing; H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval; H.5.4 [Information Interfaces and Presentation]: Hypertext/Hypermedia

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

HI-KDD’12, August 12–16, 2012, Beijing, China. Copyright 2012 ACM 978-1-4503-1548-7/12/08 $15.00.

General Terms
Algorithms; Experimentation; Human Factors; Measurement

Keywords
Adverse Drug Reaction, Online Health Community, Association Rule Mining.

1. INTRODUCTION
During the pharmaceutical product development, the pre-marketing review process focus on identifying the risk associated with drugs and the risks must be established and clearly communicated to pre-scribers and consumers. However, the pre-marketing review process cannot possibly identify all potential adverse effects. Therefore, post-marketing surveillance programs must be conducted to identify new adverse effects and evaluate their potential impact on the public health. It is found that adverse effects contribute to 5% of all hospital admission and represent the fifth most common cause of death in hospital. In this work, we focus on harnessing social media for signal detection of adverse drug reactions. These detected signals are not meant to be proven adverse effects but need to be further validated by signal analysis. The signal analysis includes determination of causality, evaluation of frequency, evaluation of biological gradient, and determination of health consequences through appropriate medical and epidemiological evaluation that exclude biases and confounding variables. Such signal analysis requires sophisticated clinical and laboratory evaluations, which are not intended to be part of this work. This work only focuses on signal detection. When a signal is confirmed by further analysis, subsequent appropriate actions must be taken to inform the prescribers and consumers.

Adverse Drug Reactions (ADRs) is defined as an appreciably harmful or unpleasant reaction, resulting from an intervention related to the use of a medicinal product, which predicts hazard from future administration and warrants prevention or specific treatment, or alteration of the dosage regimen, or withdrawal of the product [6]. ADRs represent a serious problem all over the world. They may complicate a patient’s medical conditions and increase the morbidity, even mortality. Previous study showed
that in year 2000, there were about 100,000 deaths in the U.S. due to medical errors, of which about 7,000 were attributed to drug reactions [14].

Currently, ADRs detection relies heavily on post-marketing surveillance, because clinical trials, which are required before the drugs are approved for marketing, are too constrained in scale and time span to identify many potential ADRs. In United States, FDA (Food and Drug Administration) conducts most of the post-marketing surveillance. Pharmaceutical companies, hospitals, consumers etc. spontaneously report to the FDA’s Adverse Event Reporting System (AERS). This system is limited by its passive nature, which largely depends on volunteer reports. It was estimated that the reporting rate of AERS is lower than 10% [24].

With the rapid development of Internet, there are many online health communities booming and many patients go to these websites to seek or offer healthcare information. Previous study revealed that previously unreported ADRs can be identified from patients’ reports through centralized reporting systems and that their quality is similar to those that health professional reports [2]. Therefore, these online health communities provide great platforms for the patients to discuss about the drugs they are taking, which provide enormous valuable information for detecting potential ADRs. If we can make good use of this information, we may detect ADRs much more timely and efficiently than existing reporting system. As a result, an effective adverse drug reaction signal detection system is desired to crawl, analyze, and identify signals from the health social media sites such as PatientsLikeMe and MedHelp or popular social media sites such as Facebook and Twitter, in supporting the postmarketing surveillance.

2. RELATED WORK
In recent years, many studies are focused on this area. Edwards and Aronson defined ADRs as an appreciably harmful or unpleasant reaction, resulting from an intervention related to the use of a medicinal product, which predicts hazard from future administration and warrants prevention or specific treatment, or alteration of the dosage regimen, or withdrawal of the product [6]. Edwards and Aronson also defined signal as possible causal relation between an adverse event and a drug. In addition, they provided clear definitions to Unexpected Adverse Reaction, Serious Adverse Effect and Adverse Event/ Adverse Experience [6]. Unexpected Adverse Reaction is defined as an adverse reaction, the nature or severity of which is not consistent with domestic labeling or market authorization, or expected from characteristics of the drug. Serious Adverse Effect indicates any untoward medical occurrence that at any dose results in death, requires hospital admission or prolongation of existing hospital stay, results in persistent or significant disability/incapacity, or is life threatening. The term “severe” is often used to describe the intensity (severity) of a medical event, as in the grading “mild”, “moderate”, and “severe”; thus a severe skin reaction need not be serious, while cancers and congenital anomalies or birth defects should be regarded as serious, and medical events that would be regarded as serious if they had not responded to acute treatment should also be considered serious. Adverse event/adverse experience means any untoward occurrence that may present during treatment with a pharmaceutical product but which does not necessarily have a causal relation to the treatment. In this work, we focus on detecting signals of adverse drug reactions in general health consumer contributed content in social media sites.

Some researchers studied the reasons why people report ADRs. Birrell and Edwards summarized 14 categories of reasons for reporting ADRs by conducting a research through 34 national drug monitoring centers and the top six reasons are (1) motivation to contribute to medical knowledge, (2) reaction previously unknown to reporter, (3) reaction to new drug, (4) all significant reactions reported, (5) known association between drug and reaction, and (6) severity of reaction [3].

Some previous studies employed observational methods to conduct ADRs assessment and detection, such as medical record review, solicited surveillance, patient survey, administrative data and laboratory and clinical values [8]. However, by reviewing these methods, Hakobyan et al. [8] showed that they were inadequate for identifying all possible ADRs and would not provide sufficient information about ADRs to clinicians and patients. Although Wu and Makuch incorporated external data such as established databases or pre-NDA (New Drug Application) data into observational cohort study and provided direct evidence for a reduction in sample size with these data [22], it didn’t change the fact that observational studies are time-consuming and costly.

In recent year, instead of concentrating on observational studies, many researchers and health professionals use database-related quantitative methods to detect and predict ADRs. In the United States, current postmarketing methods primarily rely on FDA’s spontaneous reporting system MedWatch (http://www.fda.gov/Safety/MedWatch/default.htm). There are also spontaneous reporting centers in other countries such as England and Japan [7, 15], and based on these reporting data various data mining methods have been practically implemented. For example, the FDA currently adopts an algorithm called Multi-item Gamma Poisson Shrinker (MGPS) for detecting potential signals from its MedWatch data [20]; UK Medicines Control Agency employs Proportional Reporting Ratios (PRR) and chi-squared to recognize adverse reactions, events related to the underlying disease and signals requiring further evaluation by comparing the proportion of all reactions to a drug of interest to the same proportion of all other drugs in UK Yellow Card database [7]; the Uppsala Monitoring Center uses Bayesian Confidence Propagation Neural Network as its signal detection strategy with World Health Organization database [18]; the Netherlands Pharmacovigilance Foundation Lareb utilizes the method using the 95% confidence interval for the Reporting Odds Ratio [21]. The performance of these methods were compared by Kubota et al. with a Japanese spontaneous reporting database and the results showed that the ability of detecting a signal varies among these methods [15]. A number of other data mining methods such as empirical Bayes model [5, 9, 16] and pharmacovigilance map method [1] have also been used with spontaneous reporting dataset.

Although the methods mentioned above performed more efficiently than traditional techniques, their performance is likely to be highly situation dependent because of the weakness and potential biases such as latency and inconsistency inherent in spontaneous reporting systems [9]. For example, FDA’s MedWatch is a passive system that depends on voluntary, spontaneous reports of suspected ADRs filed by healthcare professionals, drug manufactures, and/or consumers and it was estimated that less than 10% of all ADR cases were reported to MedWatch. In addition, early generation of a new signal can be very difficult because a large number of interesting cases cannot be timely collected due to the underreporting nature of the current
reporting system [24]. In order to solve this problem, different databases are used as alternatives. For example, using electronic medical record which is more accessible in various healthcare organization, Ji et al. developed a fuzzy logic-based computational recognition-primed decision (PRD) model to calculate the extent of causality between a drug and some of its adverse effects [12]. Based on this model, Ji et al. proposed a novel intelligent agent software system (called ADRMonitor) approach for proactively monitoring and detecting potential ADRs of interest using electronic patient records [11]. Also, on the basis of PRD model, another important signal-detection strategy is known as causal association mining algorithm in which a new interestingness measure, causal-leverage, is used to predict potential ADRs from electronic health databases [23, 24]. In addition to PRD model-based approaches, Polliot et al. generated logistic regression models that correlate postmarketing ADRs with screening data from the PubChem BioAssay database [19]. Jin et al. brought up a new interestingness measure, residual-leverage to mine unexpected temporal associations for generating ADRs signals from real-world healthcare administrative databases [10]. Nehemiah and Kannan proposed a diagnostic decision support system for adverse drug reaction using temporal reasoning. In the study, the analysis is carried out based on Modified Association Classification algorithm, which is a modified version of Apriori algorithm and uses Interestingness and Local Support measures to calculate the risk ratio and the odds ratio [13].

Data mining techniques based on electronic health data could generate earlier ADR signals than spontaneous reporting data because of its much lower underreporting rate. However, this kind of data is not available for every researcher who is interested in this area but those who are cooperating with hospitals, clinics or any other health organizations and communities. Most researchers may only have one dataset of electronic health records depending on the affiliating or collaborating health unit. The integration of electronic health records from multiple resources is still a technical and policy challenge. A single dataset of electronic health records may have limitations on the patient records that it may cover. Therefore, the availability of large scale electronic patient data from multiple sources is a limitation for its application on ADRs research in spite of its usefulness. In addition, the electronic health records are submitted by health professionals. That means the data is collected only when the health consumers visit the health professional and the adverse drug effect is recorded by the health professional. Nowadays, with the development of Internet and Web 2.0, more and more online healthcare community are emerging and booming such as MedHelp (http://www.medhelp.org) and PatientsLikeMe (http://patientslikeme.com). Everyday tens of hundreds of users post topics or comment on other users’ posts talking about their health conditions, treatment experience, drugs taken as well as ADRs of the drugs through these social media platforms. This cyber-based technique empowers patients and healthy individuals to play a substantial role in their own health and treatment and these social media data is available and accessible to public. If this data could be used effectively and efficiently, ADRs can be detected more accurately and earlier than using either spontaneous reporting data or electronic health data. In our knowledge, very few studies have employed social media to predict ADRs. For example, Chee et al. used machine learning method to classify drugs into FDA’s watchlist and non-watchlist based on messages extracted from an online health forum - Health & Wellness Yahoo! Groups but it required a training dataset to train the ensemble classifiers [4]. Practically, it takes a tremendous amount of human effort to prepare a training data for detecting the ADR signals and it may not be feasible if we have a large number of drugs and adverse drug reactions. Leaman et al. [17] used the DailyStrength health-related social network as the source of user comments. They extracted the adverse reactions by matching the terms in user comments with a lexicon that combined concepts and terms from four resources and compared the extracted adverse reactions with the annotated results generated by two annotators.

In this work, in order to explore the potential of detecting ADRs using online healthcare communities, we proposed to employ association rule mining to extract interesting associations of drugs and adverse reactions. When people talk about the ADRs of a specific drug, the co-occurrence of the drug and its ADR in the posts or comments of an online healthcare social media could be regarded as an association rule, and its interestingness and impressiveness can be measured by investigating such metrics as support, confidence and leverage. Association rule mining was first utilized in the field of data mining. Also, in the area of ADRs detection, this method was employed by several researchers to identify potential casual relationships between drugs and adverse reactions from electronic health data [10, 23, 24]. This study is trying to initially test the effectiveness of using association rule mining to extract accurate adverse reactions caused by certain drugs from online healthcare communities.

3. METHODOLOGY

3.1 Association Rule Mining

Let \( X = \{x_1, x_2, x_3, ..., x_m\} \) be a set of items. An association rule is an implication of the form \( A \Rightarrow B \), where \( A \subset X, B \subset X \), and \( A \cap B = \Phi \). Both \( A \) and \( B \) is a set of items which is referred to as an itemset. An itemset which contains \( k \) items is a \( k \)-itemset. For example, the set \{Lansoprazole\} is a 1-itemset whose element is a drug; the set \{Lansoprazole, diarrhea\} is a 2-itemset that contains two elements – a drug and one of its adverse reactions. The occurrence frequency of an itemset is the number of records that contain the itemset, denoting the support count of the itemset in the whole dataset. Thus, there are two common measures used in association rule mining: support and confidence. Support is defined as

\[
\text{support} (A \Rightarrow B) = \frac{\text{count}(A \cup B)}{\text{total count}}
\]

where \(\text{count}(A \cup B)\) is the number of occurrence that \(A\) and \(B\) co-occur and \(\text{total count}\) is the total number of records in the whole dataset. This measure calculates the proportion of records in which \(A\) and \(B\) co-occur at least once among all the records in the dataset. The importance of support lies in the fact that if \(A \cup B\) has a very low support value, it means that this association rule \((A \Rightarrow B)\) may occur simply by chance and is not interesting for us. Confidence is defined as

\[
\text{confidence} (A \Rightarrow B) = \frac{\text{support} (A \cup B)}{\text{support} (A)} = \frac{\text{count of}\ (A \cup B)}{\text{count of}\ (A)}
\]

where \(\text{support} (A)\) stands for the proportion of records in which \(A\) occurs at least once. This measure determines the extent to which the appearance of \(A\) implies the appearance of \(B\). Based on these two measures, an association rule could be identified if both of its support and confidence values exceed a pre-determined threshold.

When applying association rule mining to ADRs detection, for example, in this study, the dataset is all the threads talking about a specific drug and its ADR in the posts or comments of an online healthcare social media could be regarded as an association rule, and its interestingness and impressiveness can be measured by investigating such metrics as support, confidence and leverage. Association rule mining is first utilized in the field of data mining. Also, in the area of ADRs detection, this method was employed by several researchers to identify potential casual relationships between drugs and adverse reactions from electronic health data [10, 23, 24]. This study is trying to initially test the effectiveness of using association rule mining to extract accurate adverse reactions caused by certain drugs from online healthcare communities.
a record. Since the problem is defined as identifying the association rules in the form of drug ⇒ adverse reaction, there are two kinds of itemsets we are interested in: one is a 1-itemset whose element is a certain drug; another is a 2-itemset whose elements are a certain drug and one of its adverse reactions.

So given the total number of threads in the dataset, we can calculate the count of (drug) (number of threads in which the drug name is mentioned) and the count of (drug U adverse reaction)(number of threads in which both the drug and its adverse reaction are mentioned). Then we can get support (drug ⇒ adverse reaction) and confidence (drug ⇒ adverse reaction).

However, one of the shortcomings of using support and confidence is that these two measures work well when the frequency of ADRs is high. As a matter of fact, in online healthcare communities most of which have special sections for people to talk about a certain drug, users could discuss various aspects of the drug, such as drug dose, drug prescription and syndromes, and so on. It is possible most opinions are focused on dosage rather than ADRs especially for those rare ADRs. If support and confidence are employed in this situation, the threshold must be set relatively low to identify the rare ADRs. In such case, many false drug ⇒ adverse reaction associations would be detected. In order to handle this problem, in this study, another interestingness measure called leverage is used which is defined as [10]

\[
\text{leverage}(A \Rightarrow B) = \frac{\text{support}(A \Rightarrow B)}{\text{support}(A) \times \text{support}(B)}
\]

where support (B) is the proportion of records in which B occurs at least once (in this study, B represents ADR). This measure indicates the proportion of records in which both A and B occur by excluding that if A and B are independent of each other.

### 3.2 Consumer Health Vocabulary (CHV)

In order to match adverse reactions with the discussion threads of each drug to compute the occurrence of each ADR, we cannot simply apply the typical ADR terms used by health professionals, available at resources such as UMLS. The online healthcare communities are characterized by openness and casualness, people may use diverse expressions on the same adverse reactions communities are characterized by openness and casualness, people may use diverse expressions on the same adverse reactions.

Consumer Health Vocabulary (CHV) is a computerized collection of health expressions derived from actual consumer utterances (authored by consumers), linked to professional concepts, and reviewed and validated by professionals and consumers [25]. CHV would reflect the different ways consumers express and think about health topics, helping to bridge this vocabulary gap. In this study, we used the CHV to expand the lexicon. We searched for all the FDA adverse reactions terms in CHV wiki¹ to get the expression that are most used by consumers. For example, we found 11 different expressions of diarrhea in CHV (Table 11). We used this method to expand every adverse reaction term reported by FDA, and the final lexicon was used to match with the dataset.

### 3.3 ADR Mining Algorithm

In this study, in the drug section of our selected online healthcare community, each thread including an original post and all the following comments represents an analysis unit. Punctuations and stopwords² were removed from each thread which was then tokenized by splitting at whitespace. In order to keep the users’ original expression, we didn’t implement stemming for each word. After pre-processing, a thread is represented as a list of tokens. Since we are only interested in drug’s ADRs, most of which are nouns and noun phrases, we only count the frequencies of terms that are related to ADRs. We created a lexicon of ADRs by drawing on an external resource (CHV) which provides us with synonyms and other expressions of a specific ADR. Adverse reactions would be identified in threads by comparing a sliding window of tokens from the threads with each term in ADRs lexicon. The sliding window functions as an n-gram terms generator. For example, if the size of window is set to 3, a thread will be represented as a list of tri-gram terms. At the beginning, we need to set the value of maximum size of sliding window as n. Secondly, for each sliding window j with size smaller than n, we represent each thread i of a drug as a list of corresponding j-gram terms which would be compared with each term in ADRs lexicon. If a matching is detected, we add one to the count of co-occurrence of this drug and the matching adverse reaction. With the number of threads containing each pair of drug-adverse reaction k, we can easily determine support, confidence and leverage of the association.

Below is the pseudo code for association rule mining algorithm.

<table>
<thead>
<tr>
<th>Table 1 Expanding Terms of Diarrhea</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Diarrhea</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

¹ http://consumerhealthvocabulary.chp.utah.edu/CHVwiki/
² http://norm.al/2009/04/14/list-of-english-stop-words/
13: else continue;
14: end for
15: end for
16: for each drug-adverse reaction association k
17: \[ \text{support}(k) = \frac{\text{number of threads containing } k}{\text{total number of threads for all drugs}} \]
18: \[ \text{confidence}(k) = \frac{\text{number of threads containing } k}{\text{number of threads for this drug}} \]
19: \[ \text{leverage}(k) = \text{support}(k) - \text{support (drug in } k) \times \text{support (ADR in } k) \]
20:

4. EXPERIMENT

4.1 Dataset
In this study, we collect the dataset from MedHelp, which is one of the largest online healthcare communities worldwide. Since founded in 1994, MedHelp provides a platform for people to share needs for better medical information and support. Figure 1 presents the front page of Medelp Web site. The Drugs section is one of the sub-forums in MedHelp, and there are tens of thousands kinds of drugs included. For each drug in this section, there is a brief introduction, and MedHelp users can start a thread of this drug with a post, on which all users can comment. There are up to thousands of threads under each drug. Figure 2 presents an example of the drug information of Biaxin. To run our algorithm on the dataset, the drug should have active discussion in MedHelp. So we selected the drugs with more than 500 threads of discussion, and collected all the original posts and comments of these drugs. We selected 5 drugs that have more than 500 threads in MedHelp for our experiment.

4.2 Gold Standard
Currently in United States, FDA is responsible for the postmarketing drug safety, and information about ADRs is reported to the FDA’s Adverse Event Reporting System (AERS). We selected the drugs that have been alerted by FDA and used the safety alert information released by FDA as the ground truth to evaluate our experiment results.

<table>
<thead>
<tr>
<th>Drug Name</th>
<th>Adverse Reactions</th>
<th>Number of Threads</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biaxin</td>
<td>Heart Disease</td>
<td>686</td>
</tr>
<tr>
<td>Lansoprazole</td>
<td></td>
<td>592</td>
</tr>
<tr>
<td>Luvox</td>
<td>Heart Condition; Suicidal</td>
<td>570</td>
</tr>
<tr>
<td>Prozac</td>
<td>Suicidal; Depression</td>
<td>718</td>
</tr>
<tr>
<td>Tacrolimus</td>
<td>Cancer</td>
<td>583</td>
</tr>
</tbody>
</table>

Figure 1. MedHelp Social Media site

Figure 2. Drug information and discussions in MedHelp

To gather all the posts and comments of the 5 drugs from MedHelp efficiently, we implemented a highly parallelized automatic web crawler. All data was obtained from the raw HTML using PHP codes since there is no open API provided by MedHelp. For each thread, we extracted the subject, username, timestamp and content. However, in this study, we only use the content to mine the association rules, but the other information can be analyzed in our future study.

4.3 Table 2 Adverse Reactions Reported by FDA

Table 2 Adverse Reactions Reported by FDA

3 http://www.medhelp.org/health_topics/drugs_list

Table 2 presents the five selected drugs, the selected adverse reactions of the drugs, and the number of threads available in MedHelp.

4.3 Results and Discussion
In this study, we used our dataset to match the ADRs lexicon to find the adverse reactions of the 5 drugs that are alerted by FDA. We set the value of maximum size of sliding window as three because the longest term of ADR we obtained in CHV consisted of three words after pre-processing. We computed the support, confidence and leverage of each pair of drug ⇒ adverse reaction.

Table 1 shows the support, confidence and leverage of each pair of drug ⇒ adverse reaction. The pairs that have been alerted by FDA are highlighted. For each ADR except cancer, it can be seen that the drugs that are reported to cause the ADR have higher values than any other drugs. We can get the same results from the column charts in Figure 3 and Figure 4, which are easier to visualize the result than Table 1.

Table 1 Support, Confidence and Leverage

<table>
<thead>
<tr>
<th>Drug Name</th>
<th>Diarrhea</th>
<th>Heart Disease</th>
<th>Depression</th>
<th>Suicidal</th>
<th>Cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Support</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biaxin</td>
<td>1.75×10⁻³</td>
<td>6.99×10⁻⁴</td>
<td>2.92×10⁻¹</td>
<td>9.53×10⁻¹</td>
<td>3.56×10⁻¹</td>
</tr>
<tr>
<td>Lansoprazole</td>
<td>2.41×10⁻¹</td>
<td>6.35×10⁻³</td>
<td>2.54×10⁻²</td>
<td>9.53×10⁻⁴</td>
<td>2.83×10⁻⁰</td>
</tr>
<tr>
<td>Luvox</td>
<td>9.84×10⁻¹</td>
<td>6.67×10⁻⁴</td>
<td>1.08×10⁻¹</td>
<td>1.81×10⁻⁰</td>
<td>1.81×10⁻⁰</td>
</tr>
<tr>
<td>Prozac</td>
<td>6.35×10⁻¹</td>
<td>2.54×10⁻¹</td>
<td>1.27×10⁻⁰</td>
<td>8.89×10⁻¹</td>
<td>8.26×10⁻⁰</td>
</tr>
<tr>
<td>Tacrolimus</td>
<td>3.81×10⁻¹</td>
<td>1.27×10⁻¹</td>
<td>1.27×10⁻⁰</td>
<td>9.53×10⁻⁴</td>
<td>1.46×10⁻⁰</td>
</tr>
</tbody>
</table>

| **Confidence** |          |               |            |          |        |
| Biaxin       | 8.02×10⁻⁵ | 3.21×10⁻⁵     | 1.34×10⁻⁸ | 4.37×10⁻⁵ | 1.63×10⁻⁶ |
| Lansoprazole | 1.28×10⁻⁵ | 3.38×10⁻⁵     | 1.35×10⁻³ | 5.07×10⁻⁵ | 1.50×10⁻⁵ |
| Luvox        | 5.44×10⁻⁵ | 1.68×10⁻⁵     | 5.85×10⁻⁴ | 1.00×10⁻⁰ | 1.00×10⁻⁰ |
| Prozac       | 2.79×10⁻⁵ | 1.11×10⁻⁵     | 5.79×10⁻⁴ | 3.90×10⁻⁵ | 3.62×10⁻⁵ |
| Tacrolimus   | 2.06×10⁻⁵ | 6.86×10⁻⁵     | 6.86×10⁻⁵ | 5.15×10⁻⁵ | 7.59×10⁻⁵ |

| **Leverage** |          |               |            |          |        |
| Biaxin       | 4.04×10⁻⁵ | 1.80×10⁻⁵     | -3.63×10⁻² | -5.55×10⁻⁵ | 1.27×10⁻³ |
| Lansoprazole | 1.26×10⁻⁵ | 1.87×10⁻⁵     | -3.11×10⁻² | -4.66×10⁻³ | 8.56×10⁻⁴ |
| Luvox        | -1.31×10⁻⁵ | 2.36×10⁻⁵     | 5.19×10⁻⁸ | 1.27×10⁻¹ | -8.68×10⁻⁸ |
| Prozac       | -7.70×10⁻⁵ | -2.89×10⁻⁵    | 5.85×10⁻⁵ | 2.09×10⁻⁵ | -1.56×10⁻⁵ |
| Tacrolimus   | -7.60×10⁻⁵ | -3.14×10⁻⁵    | -4.30×10⁻² | -4.57×10⁻⁷ | -4.79×10⁻⁷ |

As shown in Figure 3 and Figure 4, for diarrhea, it is obvious that the pair Lansoprazole ⇒ Diarrhea has the highest support, confidence and leverage, which has been alerted by FDA. For heart disease, the pairs Biaxin ⇒ Heart Disease and Luvox ⇒ Heart Disease have respectively the first and second highest support as well as respectively the third and first highest value in confidence and leverage. These two drugs have been reported by FDA to be related to adverse reaction of heart diseases. As to depression, although the Prozac ⇒ Depression association has the highest support and leverage, it has the second highest confidence comparing to all the other Drug ⇒ Depression associations. The reason for this may be ascribed to the fact that the leverage formula is able to eliminate the portion of independent relationship between a drug and an adverse reaction. For suicidal, the drugs Luvox and Prozac, which have been reported to cause suicidal thoughts or suicidal actions, have the first two highest value in support, confidence and leverage. As to the last association rule which we are supposed to identify, however, Tacrolimus ⇒ Cancer appeared to be unimpressive. This result may due to the characteristics of cancer itself that a number of cancers cannot be diagnosed in time. Therefore, many cancer patients would not be aware of their situation at early stages. Thus it is less possible for them to discuss their cancers symptoms caused by the drugs they are taking in online healthcare communities. On the other hand, health consumers can assess their symptoms in diarrhea, heart disease, depression, and suicidal thoughts easily without consulting health professionals. Therefore, discussions on these ADRs can be easily identified.

It is found that we cannot simply apply one threshold on support, confidence, or leverage for all drugs and ADRs to detect the drug-adverse reaction association. The support, confidence, and leverage values vary substantially across different drugs and across different ADRs. By applying a simple threshold, we can easily miss the true drug-adverse reaction associations or identify many false drug-adverse reaction associations. This can be reflected by the diverse discussions on drugs and the variation of vocabulary usage in describing ADRs.

We can see from the results that for each adverse reaction except cancer in this experiment, the drug ⇒ adverse reaction pairs of interest, which are alerted by FDA, ranked highly among all the drug ⇒ adverse reaction pairs according to Confidence and Leverage. This indicates that our algorithm is effective in terms of detecting FDA alerted drug ⇒ adverse reaction pairs.

Moreover, in this experiment, we also find several other impressive association rules that have not been reported by FDA,
such as Lansoprazole ⇒ Heart Disease, Lavox ⇒ Depression, Biaxin ⇒ Cancer and Lansoprazole ⇒ Cancer. These drug ⇒ adverse reaction pairs all have high value in Confidence and leverage, which means there are many MedHelp users are discussing about them. These may be potential ADR signals deserving our attention and further investigation. But current study is not addressing this problem, and it can be a part of our future study.

5. CONCLUSION
Nowadays, with the booming of online healthcare communities, more and more patients find it convenient to discuss their health conditions, treatment experience, drug they are taking and adverse reactions of them through these online social media platforms. Since these data are available and accessible to public, if we can make good use of them, ADRs might be detected much earlier and more accurately than using either spontaneous FDA reports or electronic health data. However, very few related studies have focused on social media to identify ADRs, so there is a huge potential in this research area. This study collected posts and comments data of 5 drugs from an online healthcare community – MedHelp, used as grounded truth 5 FDA alerted adverse reactions of these drugs, and employ association rule mining to detect drug ⇒ adverse reaction of interest. In the experiment, we calculated the values of support, confidence and leverage for each pair and the results show that our method is able to effectively detect FDA alerted adverse reactions. We also believe that our approach is promising in discovering other potential ADRs.

However, as an initial step of our research, this study has some limitations which need improvements in the future. First of all, we only experimented on 5 drugs. In the future, we need to extend the dataset and apply our method to other kinds of drugs. Secondly, currently, we only discussed 5 adverse reactions which need to be increased. Also, when using external sources to expand ADRs terms, we solely relied on CHV which could miss many other expressions by online social media users. Their frequent misspellings could also decrease the matching performance. Our next work will be concentrated on systematically expanding the set of ADRs terms.

6. REFERENCES


