Development of an Adverse Drug Reaction Corpus from Consumer Health Posts for Psychiatric Medications

Maryam Zolnoori, MS, PhD Candidate[1], Timothy B. Patrick, MS, PhD[1], Kin Wah Fung, MD, MS, MA[2], Paul Fontelo, MD, MPH[2], Anthony Faiola, PhD[3], Yi Shuan Shirley Wu, PharmD Candidate[4], Kelly Xu, PharmD Candidate[4], Jiaxi Zhu, MS[5], Christina E. Eldredge, MD, PhD Candidate[6]

[1]Department of Health Sciences, University of Wisconsin Milwaukee, Milwaukee, WI; [2]National Library of Medicine, Bethesda, MD; [3]Department of Biomedical and Health Information Sciences, University of Illinois at Chicago, Chicago, IL; [4]School of Pharmacy, University of Pittsburgh, Pittsburgh, PA; [5]Emmes Corporation, Rockville, MD; [6]School of Information, University of South Florida, Tampa, FL

Abstract

UWM-Adverse Drug Events Corpus (UWM-ADEC) is an annotated corpus that has been developed from consumer drug review posts in social media. In this corpus, we identified four types of Adverse Drug Reactions (ADRs) including physiological, psychological, cognitive, and functional problems. Additionally, we mapped the ADRs to corresponding concepts in Unified medical language Systems (UMLS). The quality of the corpus was measured using well-defined guidelines, double coding, high inter-annotator agreement, and final reviews by pharmacists and clinical terminologists. This corpus is a valuable source for research in the area of text mining and machine learning for ADRs identifications from consumer health posts, specifically for psychiatric medications.

Introduction

Clinical trials and post-marketing surveillance systems established by regulatory agencies, such as the Adverse Event Reporting System (AERS) of the Food and Drug Administration (FDA), are not sensitive enough to detect the potential risks of drugs before marketing and moreover the occurrence of potential adverse drug reactions (ADRs) after wider use in patients. It is estimated that current surveillance systems capture less than 10% of the ADRs occurrence, due to voluntary nature of data collection and perhaps, patients' negative perceptions of the reporting systems (Yang, Yang, Jiang, & Zhang, 2012). These limitations have led to major concerns in public health because of recent reports thousands of incidents of hospitalizations and deaths (Karimi, Metke-Jimenez, Kemp, & Wang, 2015).

Recent studies have shown the potential significance of using consumer health posts from social media as a supplementary health data source to improve identifying ADRs. Therefore, regulatory agencies such as the FDA's Sentinel Initiative, have considered this source for actively monitoring for ADRs. However, there are challenges to automatic extraction of ADRs from social media, such as colloquial expressions of ADRs and deviation of sentence/phrase structure from formal sentence/phrase structure. These deviations can significantly reduce recall and precision of the automatic extraction of ADRs from consumer health posts.

A human annotated corpus can significantly improve performance of computerized systems aimed to identify health entities from unstructured consumer health posts. Development of such corpus is a very costly process. In line with the needs of improving performance of text mining algorithms in the area of pharmacovigilance, we developed a corpus of ADRs from a healthcare forum called "askapatient.com", which collects drug reviews from patients. We extracted ADRs from the review posts in this forum and mapped them to their corresponding terms in Unified medical language Systems (UMLS). To our knowledge, this corpus is the first corpus that covers a wide-range of ADRs associated with psychiatric medications, including physiological, psychological, cognitive, and functional adverse reactions.

Background

The lexicon-based approach for name entity recognition in the area of pharmacovigilance currently dominates other methods of health entity extraction in consumer health posts. The lexicons have been mostly developed by combining standard medical vocabularies including COSTART (that was developed by the FDA for coding post-

marketing ADR reports and was later replaced by MedDRA), the FDA Adverse Event Reporting System (FAERS), MedEffect (Canadian Adverse Reaction and Medical Device Problem Reporting database), SIDER (which has been developed based on resources published by public sources, mainly the FDA such as structured product labeling (SPL)), the Drug Bank Database, and the European agency for the Evaluation of Medical Product (EMEA). The lexicons were mainly built on clinical trial findings and clinicians' reports, which often have low coverage of colloquial expressions available in consumer health posts. To address this problem, pharmacovigilance studies have used a few approaches mostly focused on augmentation of the standard medical lexicons by embedding Consumer Health Vocabularies (CHV). CHV was developed mainly with the purpose of covering colloquial expression of health professional vocabularies (Zeng & Tse, 2006). Here, we explain three studies which have adopted lexicon-based approaches for identifying ADRs from consumer health posts.

Leaman et al. (2010) constructed a lexicon of SIDER, MedEffect, and COSTART, which was augmented with CHV and a small set of ADR colloquial expression to identify adverse drug reactions in consumer drug reviews in the "Daily Strength" forum. This study had 78.3% precision and 69.9% recall. Benton et al. (2011) complied a lexicon of dietary supplements, pharmaceutical terms mentioned in the Cerner Multum's Drug Lexicon, list of signs and symptoms in the Medicinenet database, FAERS, and CHV to identify ADRs of hormonal drugs used for breast cancer treatment in breast cancer healthcare forums. The reported precision was 77% and recall 35.1% . Liu and Chen (2013) constructed AZD Drug Minor on UMLS, which provided 56.5% recall and 82% precision for ADRs identification in a healthcare forum.

Pharmacovigilance studies that focused on identifying ADRs from consumer healthcare forums mostly attributed systems errors to misspelling, colloquial expression of ADRs, use of non-standard terms, and high variability of semantic representations of a specific ADR in health posts. In addition, augmentation of the standard lexicons with CHV did not improve the system's recall significantly, indicating that the CHV is not rich in colloquial expressions of ADRs. Therefore, there is a need for an annotated corpus that not only clarifies the text segments of health posts for the presence of specific information, such as ADRs, but also fills the gap between patient and clinician terminologies by mapping colloquial expressions to standard medical terminologies.

In line with this need, Ginn et al. (2014) developed an open source Twitter corpus, which was built on 10,822 instances of randomly selected tweets (each instance of tweet is a maximum of 140 characters) for drugs prescribed for chronic illness. The tweets were double coded by two annotators for presence of ADRs, spans of ADRs, drug indications, and beneficial effects. For this data set, the Inter Annotator Agreement (IAA) calculated using Cohen's Kappa was 71%. The authors normalized the identified medical terms by mapping layperson expressions to the UMLS standard terminology. Karimi et al. (2015) developed CADEC corpus, which was built on drug review posts in online message board "askapatients.com". The corpus consists of 1,231 comments for two sets of drugs, Diclofenac and Lipitor. The drug reviews were annotated for span of ADRs (6,318) where mapped to both SNOMED-CT and MEDRA terminologies. The pair-wise agreement between annotators was 60.4 % , when span and annotation settings were both strict.

The UWM-ADEC corpus is specifically developed for identifying ADRs associated with psychiatric medications. Although these medications have shown substantial evidence of effectiveness in treatment of mental illness such as depression and anxiety, they are associated with significant numbers of physiological, psychological, and cognitive ADRs unique to these types of medications. We built UWM-ADEC on drug reviews from "askapatient.com" for two classes of psychiatric medications including SSRI (Selective Serotonin Reuptake Inhibitor) and SNRI (Serotonin–norepinephrine reuptake inhibitor). In addition, we identified functional problems associated with drugs' adverse effects, such as limitations in daily functioning and social activities from the drug reviews. Identifying drug-induced functional problems was not previously identified in CADEC and the Twitter corpus. Functional problems can result in patient non-adherence behavior, and therefore may lead to an increased risk of illness relapse, emergency room visits and hospitalizations.

UWM-ADEC can be used for text mining systems and machine learning systems, specifically for psychiatric medication pharmocovigilance and hypothesis testing related to the impact of the ADRs on attitude, discontinuation, and other medical entities.

Methodology

Dataset Information

We examined data from an Online Message Board (OMB) "askapatient.com" that compiles uncensored user comments on the effects of taking different types of medication from people with a range of clinical diagnoses. In this OMB, patients can record their experience with a medication by filling out a form for a medication brand name. This form is composed of eight fields including rating, reason for prescription, side-effects, comments, gender, age, duration/dosage, and date of posting the review. Patients can rate their satisfaction with drugs ranging from 1 to 5, where 1 presents the least satisfaction and 5 presents the highest satisfaction. Patients are instructed to report drug ADRs in the side-effect field and the details of their experience in the comment field. However, patients were noted to report various aspects of their experiences, such as drug effectiveness or perceived distress due to ADRs, in both fields. **Table 1** shows an example of posts for Cymbalta in "askapatient.com".

Table 1. An example of a post for Cymbalta in "askpatient.com".

Rating	Reason	Side-effect	Comment	Gender	Age	Duration	Date
3	fibromyalgia/depression	Nausea, diarrhea, upset stomach, dry mouth, sleepiness	I have only been on 30mg for 4 days and have the extreme runs. Upset stomach and no appetite. Pain in minimal though and I feel less anxious and depressed.	F	38	4 days	2009-10-05

Drug Source

We used drug review posts in "askapatients.com" to collect information for four psychiatric medications: Sertraline (brand name: Zoloft) and Escitalopram (brand name: Lexapro) from Selective Serotonin Reuptake Inhibitor (SSRI) Class and venlafaxine (brand name: Effexor XR) and duloxetine (brand name: Cymbalta) from Serotonin Norepinephrine Reuptake Inhibitor (SNRI) Class. These four drugs have been primarily prescribed for depression and mood disorders. According to a dataset from Symphony Health Solutions, these medications had the highest prescription rates in 2012.

Data Collection

Because this healthcare forum does not have application program interface (API), we designed a web-crawler to collect information from the OMB. Since there is an option for filtering drug reviews for a specific drug, we could collect the data without requiring further effort. All the data in askapatient.com is anonymous and publicly available. Therefore, we did not seek any IRB approval for the data collection phase.

Dataset Statistics

We randomly selected 892 posts from a healthcare forum called "askapatient.com". **Table 2** shows demographic information of the whole sample. The gender proportion in the sample for female is significantly higher than male for both classes of drugs. Age range of the reviewer is 14-83 years old with the average of 37, and median of 35; implying that patients less than 40 are more likely to report their experience with drugs. Duration of drug usage ranged from 1 day to 20 years with an average of 18 months and median of 5 months, indicating that the duration of usage is highly skewed due to the effect of outliers. Posting reviews as soon as 1 day of treatment may indicate patients' high concern for potential ADRs.

Table 2. Corpus statistics

Dataset statistics	Dataset
Sample Size	892
No. of reviews with text	887
Time span	Feb 2001 Sep 2016
Rating	3.16
Gender	F 669 (76%) M 212 (24%) Missing value (11)
Age	Avg. 37 Med. 35 Missing values (12)
Age range	14-83 Missing values (3)
Duration of usage	Avg. 18 months Med. 5 month
Duration of usage (range)	1 day - 20 years

Annotation

We created the corpus in two main phases: (1) ADR identification and (2) terminology association, also known as normalization, in which we linked the identified entities to

controlled vocabularies. In the next sections, we explain the annotation guidelines and the annotation process.

Developing Guidelines for ADRs Identification

Guidelines for ADRs identification includes the ADR definitions and rules for proper identification of entities. Table 3 includes the entity definitions and the associated rules for identification with examples. The identification rules are related to patient certainty in linking ADRs with the drug, identifying patient subjective complaints and functional problems as ADRs, as well as excluding unnecessary context such as "similes" and "metaphors" from ADRs. Identifying patient subjective complaints are important because they may reflect subtle physiological, psychological, or cognitive ADRs associated with drugs. For example, "felt like I couldn't stop moving" reflects patient restlessness, which is a sign of akathisia. Identifying functional problems in drug review posts is also significant, not only for understanding how ADRs influence the normal daily activities of patients and their interpersonal relationships, but also for estimating the indirect cost associated with the ADRs. Collecting this information also enhances clinicians' abilities to predict the impact of ADRs on patient functionality, such as limitations of daily activities, social participation, and work performance. We further categorized identified ADRs as physiological (Phys), Psychological (Psycho), Cognitive (Cogn), and functional problem (FP).

Table 3. Guidelines for ADRs identification with examples

Entity	Definitions	Example	Rules for identification	Example
ADRs	Any sign or symptom that patient explicitly associated it with drug consumption, except the phase of dosage reduction and discontinuation.	My doctor increased my dose and I felt severe dizziness (ADR).	1. Certainty: If patient is not confident about the association between ADRs and drug, the ADR was not extracted. 2. Subjective complaints: If ADR is expressed as subjective complaint, it should be extracted with the entire necessary context. 3. Functional problems: if patient mentions their experiences with drugs as functional problems, such as problem with daily functioning and social activities, it should be extracted and labeled as an ADR. 4. Excluding simile and metaphor: If patient used a simile or metaphor to provide information about his/her feelings towards ADRs, that simile or metaphor should not be extracted. 5. Duplicates: Duplicate ADRs in a sentence should be independently extracted, that is, all the occurrences of the entities are identified. 6. Qualifiers: If an ADR were associated with qualifiers presenting severity or persistency of it, it needs to be identified.	1. It caused hair loss and stomach bloating (ADR), however I am not sure that hair loss (not ADR) is because of the drug. 2. "It certainly erased the anxiety, but I hardly feel human anymore (ADR). 3. I would just stay around and do nothing all day (ADR). 4. Very hard to take a deep breath (ADR) like someone is squeezing my lungs. (Smile –non necessary) 5. The anxiety (ADR) was debilitating. I also had severe headache (ADR), but the anxiety (ADR) was worse. 6. Anxiety is now though the roof (severity) • Constant (persistency) bad (severity) headaches.

Annotation Process

Four annotators participated in the process of identification and extraction of the three entities explained in Table 3. In the second step, the documents were divided into three sets and each set was reviewed by an annotator for entity identification. In order to calculate inter-annotator agreement, the entire dataset was reviewed by the fourth annotator. We did not extract general mentions of entities, such as "side-effects" in the sentences. For example, in these sentences, "*I really suffered from side-effects*," side effects and was not extracted.

Calculating Inter-Annotator Agreement

To calculate inter-annotator agreement, we used pair-wise agreement between the annotators using the following formula (Metke-Jimenez & Karimi, 2015):

$$Agreement\ (A_i, A_j) = \frac{match\ (A_{i,}\ A_j\ , \alpha, \beta)}{max\ (n_{A_i}\ ,\ n_{A_j}\ )}$$

Where A_i represents the set of data annotated by annotator *i*; A*j* represents the set of data annotated by annotator *j*; n_{A_i} is the size of identified entities in A_i and n_{A_j} is the size of identified entities in A_j; Max (n_{A_i}, n_{A_j}) is the maximum number of identified entities; α parameter presents span strictness of identified entities and β parameter represents tag strictness of identified entities. The computed pairwise agreement for strict match for ADRs identification was 0.86.

Terminology Association

While sentence classification and entity identification in drug review posts have significant implications for automatic systems that focus on information retrieval, translating these entities to the language of health professionals fills the gap between layperson and professional expressions of medical entities, such as ADRs. This translation may benefit the generation and testing of medical hypotheses by providing unambiguous and standard information for statistical data collection and analysis.

This translation process (terminology mapping) typically involves identifying terms used by healthcare consumers and mapping them to their equivalent concepts available in medical standard vocabularies. This process is also referred to as normalization in other research (Karimi et al., 2015). To normalize the entities in our corpus, we mapped the identified entities to their corresponding concepts in Unified Medical Language System (UMLS). The UMLS Metathesaurus is a compendium of many standard medical vocabularies that provides a mapping structure among vocabularies, allowing one to translate among various terminology systems. The Metathesaurus is organized by concepts. Each concept is assigned one Concept Unique Identifier (CUI) and one or more semantic type (categories). Mapping ADRs to UMLS, in addition to normalization benefit, often reveals a list of consumer health vocabularies that has not been covered by current medical terminologies.

Guidelines for Terminology Association

Due to the different ways in which symptoms, feelings, concerns etc. are described by lay persons and medical professionals, simple matching of words will sometimes fail to capture the synonymy in meaning. For example, the consumer term "feeling sick in my stomach" is equivalent to the medical term "nausea" but no words are shared. Therefore, proper mapping of consumer terms to the concepts in the UMLS must take into account both lexical and semantic matching. Since this process sometimes involves subjective judgment, to ensure consistency in mapping, we have drawn up mapping guidelines, which were iteratively updated. These guidelines were based on insights we gained by reviewing publications including clinical trial studies targeting ADRs of the drugs specified in this study and qualitative studies investigating the themes of patient experiences with the drugs. In these publications, ADRs are often grouped into three broad areas: physiological, psychological, and cognitive, an approach which we have also adopted in our study.

In some cases, the symptom mentioned by the patient is more fine-grained than the meaning of a UMLS concept, whose meaning is more general and broader in scope. In such cases, we label the map as a "specific-to-general" map. One example is the UMLS concept "executive dysfunction". According to our research, executive dysfunction as a cognitive ADR is associated with inability to initiate and follow processes of completing a task, such as problems with initiating a task, problems with organizing a task, or problems with switching between tasks. So for a patient complaint of "cannot follow through on simple tasks", we map it to "executive dysfunction" as a more general concept.

Mapping Process

Three annotators with diverse backgrounds (pharmacist, physician, and health scientist) mapped the ADRs to proper UMLS concepts based on the guidelines of mapping that we developed for this study. Annotators used the UMLS Terminology Services, UTS browser (2017) for finding proper UMLS and SNOMD-CT concepts. Example of mapping the concepts to UMLS is shown in Table 4.

Table 4. Examples of mapping ADRs to UMLS Concepts

Drug_ID	Sen_ID	Original Term	UMLS (1)
cymbalta.124	1	**Felt sick**	C0857027 / Feeling Sick /Sign or Symptom
lexapro.12	3	**"Zombie" like**	C0857486/ Felt like a zombie/ Finding
cymbalta.12	2	**Constipation**	C0009806/ Constipation/ Sign or Symptom
cymbalta.131	1	**Excessive sleepiness**	C0013144/ Drowsiness/ Finding
Effexor.78	1	**Minor muscle spasms**	C0037763 / Spasm/ Sign or Symptom
effexor.97	2	**Sweating like crazy all the time**	C0700590 / Increased sweating / Sign or Symptom
effexor.111	7	**Brain zap**	No concept

Corpus Statistics

Table 5 lists frequency of identified ADRs for the corpus, as well as type of ADRs separately. Overall, we identified 4776 ADRs where 31% were duplicates, with the lowest number of duplicates for functional problems, followed by psychological problems. The findings indicate the level of subjectivity of functional and psychological ADRs that leads to creating different phrases by patients to describe their feelings and problems. Functional problems only made up 2% of the total ADRs, indicating that patients prefer to discuss physical and psychological effects of the drugs in review posts rather than their impacts on their quality of life. For the purpose of designing more effective medication adherence interventions, it would be useful if healthcare forums also asked patients to report the impact of drugs on their daily functioning and social activities.

Table 5. Frequency of identified ADRs for the total corpus

	Total		Physiological		Psychological		Cognitive		Functional	
	ADRs	Unique	All	Unique	All	Unique	All	Unique	All	Unique
ADRs in Corpus	4776	69% (All) 3285	3522	64% (All) 2274	900	81% (All) (716)	272	80% (All) (217)	82	95% (All) (78)

Statistics for annotation from the normalization stage were also shown in Table 6. The final normalization set contains 695 concepts from UMLS concepts, from which 61% belong to the physiological category and only 5% of the concepts are related to the functional category. We also report the two most frequently mapped concepts with their frequencies across the corpus for each category of ADRs.

Table 6. Statistics for annotation from normalization stage

	Total	Physiological	Psychological	Cognitive	Functional
No. Unique concepts	695	425 (61% total)	196 (28% total)	42 (6% total)	31 (5%)
No. Unique concepts	695	425 (61% total)	196 (28% total)	42 (6% total)	31 (5%)
1st most freq. concept	Sleeplessness (171)	Sleeplessness (171)	Anxiety (94)	Foggy feeling in head (47)	Difficulty in daily functioning (10)
2st most freq. concept	Nausea (169)	Nausea (169)	Detailed recall of dream (62)	Unable to concentrate (30)	Emergency room admission (9)

Normalization Challenge

While normalization of consumer health posts has significant implications for understanding pharmacological aspects of medications, it is a subjective process. We attempted to address this by developing guidelines that

include underlying concepts for both patient and professional expressions of entities. But, some expressions strongly related to the life context of patients. For example, we did not map “hardly feel human anymore” to any concepts due to uncertainty of the underlying concepts associated with it. It is not clear what the patient meant with this expression: is it about the patient feeling emotionally detached, having a problem in performing daily activities, or is it about feeling detached from his mind and his body (de-realization)?

There were also some cases that, while the expression of an ADR is clear and can be translated to an equivalent medical concept, there are no UMLS concepts available for it. For example, brain zap, which is known as the professional term “brain shivers”, does not have any concept in UMLS.

Limitations

Sample size

The size of sample is limited to 892 posts for four psychiatric medications. While this sample size is a good representative of the four most common psychiatric medications, it may not be a good representative of other consumer posts in this forum or other healthcare forums. It is also possible that a specific group of patients tend to report their experiences with drugs in this forum leading to reporting bias.

Limitation for coverage of drug types

Our corpus covers sentence labeling and entities identifications for two classes of psychiatric medications, SSRI and SNRI. While limiting the dataset to a specific set of drugs enabled us to have a better understanding of conceptual models associated with layperson and professional expressions of medical entities, it may not include rare ADRs related to other classes of psychiatric medications and medications for other diseases or disorders.

Lack of information on drug-drug interactions, drug-herb interaction, and drug overdose

The focus of patients in review posts is mostly on the selected drug. Hence, it is not clear, whether the reported adverse effects are merely caused by the drug or it is the result of interaction of the drug with other potential drugs or herbal treatment that administered by patients. Moreover, some of the ADRs for psychiatric medications, such suicidal ideation or emergency visits can happen because of patient’s overdose, this information is not available in the review posts because of the nature of these reports.

Uncertainty of data in social media

Although patient self-reported experiences is a reliable source for evaluating pharmacological effects of medications, there is still the risk of inaccurate and false information. In addition, we only identified and extracted ADRs that patients directly associated with their medications, however, there is the possibility that patients misinterpreted the symptoms of their psychiatric condition as an ADR of their psychiatric medication.

Possibility of human errors in data analysis

Although the entire data set is double coded, there is still the possibility that annotators did not interpret a sentence correctly and therefore assign a wrong label to it. In addition, the span of the identified entities may include less or more information than necessary. These problems affect the performance of any machine learning system trained on this corpus to identify drug effectiveness, ADRs, and drug indications in consumer health posts.

Conclusions and Future Work

We have created a corpus of ADRs with the purpose of improving recall and precision of automatic systems designed for identifying ADRs from social media. The source of this corpus is patient reviews of psychiatric drugs in a medical forum called askpatient.com. Sentences in review posts were annotated for the presence of ADRs and span of ADRs. This corpus can benefit researchers in several areas including 1) developing and evaluating systems that automatically identify ADRs from consumer health posts, 2) developing systems that automatically map free text to UMLS, 3) Creating a structured vocabulary of layperson expressions of adverse effects and indications which can be used in electric health records (EHR) for facilitating seamless information between patients and clinicians. This can be achieved by mapping information in personal health records (PHR) to EHR systems.

We are in the process of annotating withdrawal symptoms, drug indications, and effectiveness from the consumer health reports. We are also mapping the entities to corresponding terms in SNOMED-CT terminology.

References

Benton, A., Ungar, L., Hill, S., Hennessy, S., Mao, J., Chung, A., . . . Holmes, J. H. (2011). Identifying potential adverse effects using the web: A new approach to medical hypothesis generation. *Journal of Biomedical Informatics, 44*(6), 989-996.

Chapman, W. W., Bridewell, W., Hanbury, P., Cooper, G. F., & Buchanan, B. G. (2001). A simple algorithm for identifying negated findings and diseases in discharge summaries. *Journal of Biomedical Informatics, 34*(5), 301-310.

Ginn, R., Pimpalkhute, P., Nikfarjam, A., Patki, A., O'Connor, K., Sarker, A., . . . Gonzalez, G. (2014). *Mining Twitter for adverse drug reaction mentions: a corpus and classification benchmark.* Paper presented at the Proceedings of the fourth workshop on building and evaluating resources for health and biomedical text processing.

Gurulingappa, H., Mateen-Rajput, A., & Toldo, L. (2012). Extraction of potential adverse drug events from medical case reports. *J Biomed Semantics, 3*(1), 15.

Huynh, T., He, Y., Willis, A., & Rüger, S. (2016). *Adverse drug reaction classification with deep neural networks*.

Karimi, S., Metke-Jimenez, A., Kemp, M., & Wang, C. (2015). CADEC: A corpus of adverse drug event annotations. *Journal of Biomedical Informatics, 55*, 73-81.

Leaman, R., Wojtulewicz, L., Sullivan, R., Skariah, A., Yang, J., & Gonzalez, G. (2010). *Towards internet-age pharmacovigilance: extracting adverse drug reactions from user posts to health-related social networks.* Paper presented at the Proceedings of the 2010 workshop on biomedical natural language processing.

Liu, X., & Chen, H. (2013). AZDrugMiner: an information extraction system for mining patient-reported adverse drug events in online patient forums *Smart Health* (pp. 134-150): Springer.

Metke-Jimenez, A., & Karimi, S. (2015). Concept Extraction to Identify Adverse Drug Reactions in Medical Forums: A Comparison of Algorithms. *arXiv preprint arXiv:1504.06936*.

Metke-Jimenez, A., Karimi, S., & Paris, C. (2014). *Evaluation of text-processing algorithms for adverse drug event extraction from social media.* Paper presented at the Proceedings of the first international workshop on Social media retrieval and analysis.

Nguyen, T., Larsen, M. E., O'Dea, B., Phung, D., Venkatesh, S., & Christensen, H. (2017). Estimation of the prevalence of adverse drug reactions from social media. *Int J Med Inform, 102*, 130-137.

Nikfarjam, A., Sarker, A., O'Connor, K., Ginn, R., & Gonzalez, G. (2015). Pharmacovigilance from social media: mining adverse drug reaction mentions using sequence labeling with word embedding cluster features. *Journal of the American Medical Informatics Association*, ocu041.

Sarker, A., & Gonzalez, G. (2015). Portable automatic text classification for adverse drug reaction detection via multi-corpus training. *Journal of Biomedical Informatics, 53*, 196-207.

Yang, C. C., Yang, H., Jiang, L., & Zhang, M. (2012). *Social media mining for drug safety signal detection.* Paper presented at the Proceedings of the 2012 international workshop on Smart health and wellbeing.

Zeng, Q. T., & Tse, T. (2006). Exploring and developing consumer health vocabularies. *Journal of the American Medical Informatics Association, 13*(1), 24-29.