


December 2017

Utilizing Consumer Health Posts for Pharmacovigilance: Identifying Underlying Factors Associated with Patients' Attitudes Towards Antidepressants

Maryam Zolnoori
University of Wisconsin-Milwaukee

Follow this and additional works at: <https://dc.uwm.edu/etd>

 Part of the [Databases and Information Systems Commons](#), [Medicine and Health Sciences Commons](#), and the [Psychology Commons](#)

Recommended Citation

Zolnoori, Maryam, "Utilizing Consumer Health Posts for Pharmacovigilance: Identifying Underlying Factors Associated with Patients' Attitudes Towards Antidepressants" (2017). *Theses and Dissertations*. 1733.
<https://dc.uwm.edu/etd/1733>

This Dissertation is brought to you for free and open access by UWM Digital Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of UWM Digital Commons. For more information, please contact open-access@uwm.edu.

**UTILIZING CONSUMER HEALTH POSTS FOR PHARMACOVIGILANCE:
IDENTIFYING UNDERLYING FACTORS ASSOCIATED WITH PATIENTS'
ATTITUDES TOWARDS ANTIDEPRESSANTS**

by

Maryam Zolnoori

A Dissertation Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in Health Sciences

at

The University of Wisconsin-Milwaukee

December 2017

ABSTRACT

UTILIZING CONSUMER HEALTH POSTS FOR PHARMACOVIGILANCE: IDENTIFYING UNDERLYING FACTORS ASSOCIATED WITH PATIENTS' ATTITUDES TOWARDS ANTIDEPRESSANTS

by

Maryam Zolnoori

The University of Wisconsin-Milwaukee, 2017
Under the Supervision of Professor Timothy Patrick

Non-adherence to antidepressants is a major obstacle to antidepressants therapeutic benefits, resulting in increased risk of relapse, emergency visits, and significant burden on individuals and the healthcare system. Several studies showed that non-adherence is weakly associated with personal and clinical variables, but strongly associated with patients' beliefs and attitudes towards medications. The traditional methods for identifying the key dimensions of patients' attitudes towards antidepressants are associated with some methodological limitations, such as concern about confidentiality of personal information. In this study, attempts have been made to address the limitations by utilizing patients' self report experiences in online healthcare forums to identify underlying factors affecting patients attitudes towards antidepressants. The data source of the study was a healthcare forum called "askapatients.com". 892 patients' reviews were randomly collected from the forum for the four most commonly prescribed antidepressants including Sertraline (Zoloft) and Escitalopram (Lexapro) from SSRI class, and Venlafaxine (Effexor) and duloxetine (Cymbalta) from SNRI class. Methodology of this study is composed of two main phases: I) generating structured data from unstructured patients' drug reviews and

testing hypotheses concerning attitude, II) identification and normalization of Adverse Drug Reactions (ADRs), Withdrawal Symptoms (WDs) and Drug Indications (DIs) from the posts, and mapping them to both The UMLS and SNOMED CT concepts. Phase II also includes testing the association between ADRs and attitude. The result of the first phase of this study showed that “experience of adverse drug reactions”, “perceived distress received from ADRs”, “lack of knowledge about medication’s mechanism”, “withdrawal experience”, “duration of usage”, and “drug effectiveness” are strongly associated with patients attitudes. However, demographic variables including “age” and “gender” are not associated with attitude. Analysis of the data in second phase of the study showed that from 6,534 identified entities, 73% are ADRs, 12% are WDs, and 15 % are drug indications. In addition, psychological and cognitive expressions have higher variability than physiological expressions. All three types of entities were mapped to 811 UMLS and SNOMED CT concepts. Testing the association between ADRs and attitude showed that from twenty-one physiological ADRs specified in the ASEC questionnaire, “dry mouth”, “increased appetite”, “disorientation”, “yawning”, “weight gain”, and “problem with sexual dysfunction” are associated with attitude. A set of psychological and cognitive ADRs, such as “emotional indifference” and “memory problem” were also tested that showed significance association between these types of ADRs and attitude. The findings of this study have important implications for designing clinical interventions aiming to improve patients' adherence towards antidepressants. In addition, the dataset generated in this study has significant implications for improving performance of text-mining algorithms aiming to identify health related information from consumer health posts. Moreover, the dataset can be used for generating and testing hypotheses related to ADRs associated with psychiatric mediations, and identifying factors associated with discontinuation of antidepressants. The dataset and guidelines of this study are available at

<https://sites.google.com/view/pharmacovigilanceinpsychiatry/home>

© Copyright by Maryam Zolnoori, 2017
All Rights Reserved

To

My parents for their encouragement

My Sisters for their support

My advisors for their invaluable guidance and advice

Everyone who seeks enlightenment through knowledge

TABLE OF CONTENTS

LIST OF FIGURES	ix
LIST OF TABLES	xi
LIST OF ABBREVIATIONS	xiii
ACKNOWLEDGEMENTS	xv
1. Introduction.....	1
1.1 Statement of the Problem.....	2
1.2 Significance of the Study	5
1.3 Statement of Aims.....	7
1.4 Data Source and Methodology of the Study	8
1.5 Contributions	9
2. Background	11
2.1 Treatment of Depression	12
2.2 Psychopharmacology Treatment VS. Psychological Treatment	13
2.3 Types of Antidepressants	13
2.3.1 Selective serotonin reuptake inhibitors (SSRIs)	13
2.3.2 Serotonin–Norepinephrine Reuptake Inhibitors (SNRIs)	15
2.4 Experience of Patients With Antidepressants	15
2.5 Duration of Antidepressants Usage.....	16
2.6 Adherence to Antidepressant Treatment	16
2.6.1 Importance of adherence to antidepressant treatment.....	17
2.6.2 Factors Affecting on non-Adherence behavior	17
2.7 Potential Predictive Factors Associated With Attitude Towards Antidepressants	22
2.7.1 Pharmacological treatment factors.....	23
2.7.2 Healthcare system factors	27
2.7.3 Psycho-social factors	29
2.7.4 Patient related factors.....	30
2.7.5 Depression disorder characteristics.....	32
2.8 Self-Report Scales Used in Antidepressants Studies for Measuring Attitude.....	33
2.8.1 Methodological Limitations of self-report scales	33
2.9 Social Media and its Application in Healthcare.....	36
2.10 Challenges in Health Information Extraction from Consumer Health Posts	37
2.10.1 Automatic extractions of ADRs from consumer health posts	39
2.10.2 Corpora Developed Using Consumer Health posts	44
2.1.3 Corpora developed using biomedical literature.	45
3. Methodology	48

Phase 1	52
3.1 Data Source Information	52
3.1.1 Data sources	52
3.1.2 Drug Source	52
Phase 2	53
3.2 Data collection	53
3.2.1 Website policy	53
3.2.2 Designing application program interface (API)	54
3.2.3 Calculating the sample size	54
Phase 3	55
3.3 Developing the Analytical Framework for Generating Structured Data	55
3.3.1 Developing the initial analytical framework using deductive approach	56
3.3.2 Coding (annotating) the drug reviews using the initial analytical framework	56
3.3.3 Generating codes (themes) using inductive approach (open coding)	60
3.3.4 Developing the final analytical framework by refining the themes	61
Phase 4	67
3.4 Applying the Final Analytical Framework to the Drug Reviews	67
3.4.1 Data Pre-processing	68
3.4.2 Annotating sentences using analytical framework (sentence labeling)	68
Phase 5	76
3.5 Entity Identification	76
3.5.1 Guidelines for entity identification	77
3.5.2 Entity identification process	78
Phase 6	81
3.6 Terminology Association (Entities Normalization)	81
3.6.1 Guidelines for Terminology Association	83
3.6.2 Creating a dictionary	85
3.6.3 Procedure of mapping	86
Phase 7	89
3.7 Usability of the Dataset (Corpus)	89
3.7.1 Identifying the underlying factors affecting patient attitude towards drugs	90
3.7.2 Testing Association between ADRs and Attitude	95
3.7.3 Statistical Methods	96
4. Results	97
4.1 Dataset Statistics	97
4.2 Statistics on Developing and Applying the Analytical Framework	98
4.3 Statistics on Summarized Dataset	102
4.4 Testing Hypotheses	104
4.4.1 Measuring association between variables and levels of attitude	104
4.4.2 Developing a predictive model	113
4.5 Statistics on Entities Identification	115
4.6 Statistics on Terminology Association	119
4.7 Testing association between specific type of ADRs and levels of Attitude	121
5. Discussion	137

5.1 Significance of Social Media for Providing Insight about Attitude and Adherence Behavior	138
5.2 Significance of Social Media in Identifying Underlying Factors Associated with Attitude	138
5.3 Limitations of Social Media in Identifying Underlying Factors Associated with Attitude	140
5.4 Significance of Social Media in Identifying ADRs and WDs Associated with Antidepressants	141
5.5 Implications of this Study.....	142
5.5.1 Developing tools in the area of text mining algorithms and machine learning for extraction of health related information from consumer health posts.....	142
5.5.2 Testing the association between physiological, psychological, cognitive, and functional problems with attitude	144
5.5.3 Developing self-report scales.....	144
5.5.4 Implications of the analytical framework	144
5.5.5 Implication of the findings related to the factors associated with attitude.....	145
5.6 Limitations of this Study	145
Conclusion	149
Future Work.....	150
REFERENCES.....	152
Appendix A: Self-report questionnaire designed/used for measuring ADRs associated with antidepressants.....	161
Appendix B: Self-Report Scales Used in Antidepressants Studies for Measuring Attitude	165
Curriculum Vitae	172

LIST OF FIGURES

Figure 3-1 Summary of the research methodology.....	51
Figure 4-1 Distribution of gender for each level of patients attitude.....	106
Figure 4-2 Frequency of gender for each level of attitude.....	106
Figure 4-3 Distribution of age for each level of attitude	107
Figure 4-4 Distribution of duration of usage for each level of attitude	107
Figure 4-5 Distribution of presence and absence of ADRs for attitude levels	108
Figure 4-6 Frequency of ADR for each level of attitude	108
Figure 4-7. Distribution of perceived distress from ADRs (high, low, medium).....	109
figure 4-8 Frequency of perceived distress from ADRs	109
Figure 4-9 Distribution of drug effectiveness.....	110
Figure 4-10 Frequency of drug effectiveness	110
Figure 4-11 Distribution of withdrawal experience for levels of attitude	111
Figure 4-12 Frequency of withdrawal experience for each level of attitude	111
Figure 4-13 Distribution of complaints for lack of knowledge	112
Figure 4-14 Frequency of complaints for lack of knowledge for each level of attitude.....	112
Figure 4-15 Frequency of headache for each level of attitude.....	124
Figure 4-16 Frequency of insomnia for each level of attitude.....	124
Figure 4-17 Frequency of drowsiness for each level of attitude.....	125
Figure 4-18 Frequency of blurred vision for each level of attitude	125
Figure 4-19 Frequency of constipation for each level of attitude.....	126
Figure 4-20 Frequency of diarrhea for each level of attitude	126
Figure 4-21 Frequency of decreased appetite for each level of attitude	127

Figure 4-22 Frequency of increased appetite for each level of attitude.....	127
Figure 4-23 Frequency of sweating for each level of attitude	128
Figure 4-24 Frequency of disorientation for each level of attitude	128
Figure 4-25 Frequency of yawning for each level of attitude.....	129
Figure 4-26 Frequency of weight gain for each level of attitude.....	129
Figure 4-27 Frequency of problem with urination for each level of attitude.....	130
Figure 4-28 Frequency of sexual dysfunction for each level of attitude	130
Figure 4-29 Frequency of nausea or vomiting for each level of attitude.....	131
Figure 4-30 Frequency of nausea or vertigo for each level of attitude.....	131
Figure 4-31 Frequency of emotional indifference for each level of attitude	132
Figure 4-32 Frequency of apathy for each level of attitude.....	132
Figure 4-33 Frequency of emotionally detached for each level of attitude	133
Figure 4-34 Frequency of mood swings for each level of attitude	133
Figure 4-35 Frequency of restricted emotion (emotional indifference + apathy + emotionally detached + restricted motivation) for each level of attitude.	134
Figure 4-36 Frequency of anxiety for each level of attitude.....	135
Figure 4-37 Frequency of motivation for each level of attitude	135
Figure 4-38 Frequency of difficulty in concentration for each level of attitude.....	136
Figure 4-39 Frequency of memory problem for each level of attitude.....	136

LIST OF TABLES

Table 2-1 Potential Predictive Factors Affecting Non-adherence towards Antidepressants	19
Table 2-2 Factors Affecting Patients Attitude Towards Antidepressants (Suggested by Literature)	23
Table 2-3 Specifications of Studies Used Lexicon-based Approach and Machine Learning Methods for Automatic Extraction of Health-Related Information from Consumer Health Posts.	43
Table 2-4 Specifications of Existing Corpus Relevant to the Corpus Specified in this Study	46
Table 3-1 An Example of a Post for Cymbalta in "askapatient.com"	52
Table 3-2 Total Number of Review Posts and the Calculated Sample Size	55
Table 3-3 Description of Predefined Themes Developed Deductively from Literature Review .	57
Table 3-4 Predefined Themes Developed Inductively with Examples from Drug Reviews	60
Table 3-5 New Themes Generated Through the Process of Inductive Approach (Open Coding)	62
Table 3-6 Themes (items) in the Fianl Analytical Framework with Description and Example from Patients Comments.....	64
Table 3-7 Examples of Regular Expression codes for Reducing Noise Patterns in Review Posts	69
Table 3-8 A Snapshot of the Coding (Annotation) Environment	70
Table 3-9 Computed IAA for each Themes in the Sentence Classification Component with Examples of Disagreement	75
Table 3-10 Entities' Definitions and Identification Rules with Examples.....	79
Table 3-11 Requirements for Selecting Proper UMLS and SNOMED CT Concepts	87
Table 3-12 Examples of Mapping Entities to both UMLS and SNOMED-CT Concepts	89
Table 3-13 Lists of the Variables, Possible Values, and Limitations Associated with each of them for Measuring Patient Attitude toward Antidepressants.....	92
Table 4-1 Dataset Statistics.....	99
Table 4-2 Statistics on Frequency of Posts, Sentences, and Tokens	101
Table 4-3 Frequency of Themes after Resolving Disagreements for the Whole Dataset.....	101

Table 4-4 Statistics on Variables Used for Testing Hypotheses.....	103
Table 4-5 List of Variables, Hypothesis, Statistics Test, and computed P-values.....	104
Table 4-6 Values of statistical model.....	113
Table 4-7 Confidence Interval for the Outcome Variables in the Model	114
Table 4-8 Statistics on Entity Identification Component for Each Class of Drugs and Type of Entities Separately	117
Table 4-9 The Top Five Identified Entities for Each Class of Drug s and Type of Entities Separately.....	118
Table 4-10 Statistics on Entities Normalized with the UMLS and SNOMED-CT Concepts	120
Table 4-11 The Most Frequent UMLS Concepts in the Corpus without SNOMED Expressions	121
Table 4-12 Frequency of Identified Qualifiers Associated with ADRs and WDs.....	121
Table 4-13 Testing the Association between ADRs and Attitude	122
Table Appendix B.1. Drug Attitude Inventory (DAI): summary of the components' items.....	165
Table Appendix B.2. Antidepressant Compliance Questionnaire (ADCQ): Summary of the components' items	167
Table Appendix B.3. Beliefs about Medications Questionnaire (BMQ): a summary of the components' items	169
Table Appendix B.4. Rating of Medication Influence Scale (ROMI): Summary of the Components' items	170

LIST OF ABBREVIATIONS

ADCQ	Antidepressants Compliance Questionnaire
ADR	Adverse Drug Reaction
ADR-PD	Perceived distress from ADRs
AERS	Adverse Event Reporting System
ASEC	Antidepressant Side-Effect Checklist
ATT	General attitude toward medication
BMQ	Beliefs about Medicines Questionnaire
CDC	Centers for Disease Control and Prevention
CHV	Consumer Health Vocabularies
CNN	Convolutional Neural Network
CNNA	Convolutional Neural Network with Attention
Cogn	Cognitive
CRNN	Convolutional Recurrent Neural Network
DAI	Drug Attitude Inventory
DI	Drug Indication
DXD-Dec	Intentional decision for Withdrawal
DXD-S	Intentional Withdrawal -Stopping
DXD-W	Intentional Withdrawal -weaning off
EF	Drug Effectiveness
EHRs	Electronic Health Records
EMA	European agency for the Evaluation of Medical Product
Experience of WD	Experience of Withdrawal
FDA	Food and Drug Administration
FP	Functional Problem
FS	Lack of Financial Support (FS)
HBM	Health Belief Model
IAA	Inter Annotator Agreement
ICF	International Classification of Functioning, Disability and Health
INF	Ineffectiveness
KN	Lack of Knowledge
KNN	K-nearest Neighborhood
LUNSERS	Liverpool University Neuroleptic Side Effect Rating Scale
MedEffect	Canadian Adverse Reaction and Medical Device Problem Reporting Database
N	Negative
NB	Naïve Bayes
NLTK	Open-source Natural Language Toolkit
OMB	Online Messaging Board
P	Positive
PHR	Personal Health Record
Phys	Physiological
PPI	Patient –physician interaction

Psycho	Psychological
QP	Qualifiers representing the Persistency
QS	Qualifiers representing the severity
RCNN	Recurrent Convolutional Neural Network
ROMI	Rating of Medication Influence Scale
SAFTEE	Systematic Assessment for Treatment Emergent Side Effects
SAS	Simpson Angus Scale
SNOMED CT	Systematized Nomenclature of Medicine Clinical Terms
SNRI	Serotonin and Norepinephrine Reuptake Inhibitors
SPL	Structured Product Labeling
SS (~SS)	Lack of Social Support
SSRI	Selective Serotonin Reuptake Inhibitor
SUG	Patient suggestion to readers about drug
SVM	Support Vector Machine
TCAs	Tricyclic antidepressants
TPB	Theory of Planned Behavior
UMLS	Unified Medical Language System
DXD-F	Unintentional Withdrawal
WD	Withdrawal symptoms
WD-PD	WD-perceived distress
WHOQOL-BREF	WHO Quality of Life Assessment-Brief version

ACKNOWLEDGEMENTS

I would never have been able to accomplish my dissertation without the support I received from my advisors, friends, and family. I would like to express my deepest gratitude to my advisor Dr. Timothy Patrick for his strong support, caring, and patience during the past three years I was a student at UWM. He is a prime example of what a scientist should be. I would like to express my sincere appreciations for Dr. Kin Wah Fung and Dr. Paul Fontelo from the National Library of Medicine for their continued support and the countless hours they have spent to provide me with their guidance and advice during my dissertation process. I would like to thank my doctoral committee Dr. Anthony Faiola, Dr. Jake Luo, Dr. Virginia Stoffel, Dr. Hadi Kharrazi, and Dr. Rohit Kate for their support, reviews, and helpful comments. I have many thanks for my friends Yi Shuan Shirley Wu, Kelly Xu, Jiaxi Zhu, Christina Eldredge, Soo Kyung Park, Margaret Downs, and Lillian Folk for their significant contributions to preparing the dataset of this study. My research could not be possible without them. Finally, I would like to express my heartfelt appreciations for my parents and sisters for their encouragement and confidence in my abilities.

Chapter 1: Introduction

1.1 Statement of the Problem

Depression is a major public health concern due to the high prevalence and it is a substantial burden on families, society, and healthcare system (De las Cuevas, Peñate, & Sanz, 2014). It is estimated that 150 million people suffer from depression at a certain point in their life (World Health Organization, 2003) and by 2020, depression will be the most common cause of disability in the world (Murray & Lopez, 1997), mainly because of cognitive and emotional impairment, such as memory recall and low motivation.

In the United States, according to The Centers for Disease Control and Prevention (CDC), about 9 percent of Americans show symptoms of desperation, regret, and sadness that result in depression (CDC 2014) and approximately six percent of adults annually are affected by depression, which is the major cause of suicide (NCCMH, 2010). The financial burden of depression in the United States increased from \$173.2 billion to \$210.5 billion (21.5 percent) from 2005 to 2010 (Greenberg, Fournier, Sisitsky, Pike, & Kessler, 2015). Forty-five percent of this cost was related to depression treatment, fifty percent of the cost was indirect cost related to patients' dysfunction and low productivity in the workplace, and five percent of this cost was related to suicide-related behavior.

Antidepressants from SSRI (Selective Serotonin Reuptake Inhibitor) and SNRI (Serotonin and Norepinephrine Reuptake Inhibitors) classes have proliferated pharmacological interventions for depression in both psychiatric and primary care settings. This level of popularity may be related to efficacy of these drugs and their relatively mild adverse effects (Sansone & Sansone, 2010). Therapeutic benefits of antidepressant treatment depend on maintaining an appropriate drug regimen for a certain period of time and a proper method of discontinuation. Many patients recognize the necessity of the treatment (Murata, Kanbayashi,

Shimizu, & Miura, 2012), but non-adherence to the treatment is a major obstacle to its clinical effectiveness (López-Torres, Párraga, Del Campo, & Villena, 2013). Non-adherence to antidepressants is associated with the increased rate of emergency visits, longer hospital admission, low quality of life, and significant cost at individual, family, and health care system levels (Vives et al., 2014).

Research showed that across chronic medical illness in general, non-adherence is loosely associated with demographic information and clinical variables, such as severity of the symptoms or types of drug adverse effects. However, it is more strongly associated with patients' beliefs and attitudes toward medication, low perceived necessity, and high perceived medication risks (Aikens & Klinkman, 2012), (Acosta, Rodríguez, & Cabrera, 2013; Aikens, Nease, & Klinkman, 2008; Brown et al., 2005; Chakraborty, Avasthi, Kumar, & Grover, 2009; De las Cuevas et al., 2014; López-Torres et al., 2013; Richardson, McCabe, & Priebe, 2013; Verdoux et al., 2000) . It is more likely that patients with positive attitudes show better adherence behavior than patients with negative or indifferent attitudes(Aikens, Nease, Nau, Klinkman, & Schwenk, 2005; De las Cuevas et al., 2014; Richardson et al., 2013). Ultimately, It is patient perception of beneficial aspects of the medication versus the harmful aspects that will determine the acceptance or rejection of the medication (Acosta et al., 2013; Christensen, 2004).

Identifying the key dimensions of patients' attitudes towards antidepressants is challenging work that is often neglected (Demyttenaere et al., 2004). This partly due to the lack of reliable and comprehensive methods to capture underlying factors affecting patients' attitudes. It has been evident that identifying potential predictive factors for attitude may have significant implications for understanding the complexity of this phenomenon and for designing effective

interventions to improve drug therapy adherence (Kessing, Hansen, Demyttenaere, & Bech, 2005).

Some studies have attempted to develop structured self-report measures to operationalize the concept of attitude towards psychiatric drugs in various scales. The Drug Attitude Inventory (DAI) (Hogan, Awad, & Eastwood, 1983) and the Antidepressants Compliance Questionnaire (ADCQ) (Demyttenaere, Adelin, Patrick, Walthère, & Michèle, 2008) are examples of self-report scales employed by studies focused on measuring patient attitudes toward antidepressants. Although, the self-report scales are well validated, they are associated with some limitations. Firstly, the ADCQ is the only self-report scale that is primarily designed for measuring attitude to antidepressants. However, other scales, such as the DAI, which are primary designed for patients with psychotic disorders, have been widely used in the studies addressing attitudes to antidepressants. Secondly, the scales do not include a comprehensive set of potential factors influencing attitudes, such as antidepressants' adverse effects or patients' knowledge. Therefore, studies interested in determining association of the factors with attitude need to employ relevant self-report scales such as the Self-report Antidepressant Side-Effect (ASEC), which may increase complexity and cost of a study. Third, to reduce complexity of data analysis, most of the studies represented outputs of the self-report scales as dichotomous variables (positive vs. negative attitude) that may not capture meaningful variance in data. Finally, stigma associated with mental disorders (Griffiths, Calear, & Banfield, 2009), sample bias and inherent limitations with such data collection, concern about confidentiality of personal information, and patients' reluctance to reveal personal information may reduce the reliability of the studies' findings.

Detecting Adverse Drug Reactions (ADRs) for antidepressants is also associated with several challenges. First of all, clinical trial studies, as the pre-marketing phase of

pharmacovigilance systems, suffer from limitations, such as specific inclusion and exclusion criteria for recruiting patients and limited duration and size of the trials. These limitations may hinder the discovery of long-term effects of antidepressants' usage, such as antidepressants-induced weight gain or rare adverse effects, such as "eye pain", which may occur in less than 5000 patients (Ferguson, 2001). In addition, the post-marketing phase of pharmacovigilance systems, such as Food and Drug Administration (FDA)'s Adverse Event Reporting System (FAERS), are voluntary in nature for clinicians and the general public, leading to detection that may not be timely and is incomplete (Sarker et al., 2015). To facilitate the process of assessing adverse effects of medications, some studies designed instruments to measure the adverse effects in clinicians' offices, including Systematic Assessment for Treatment Emergent Side Effects (SAFTEE) (Levine & Schooler, 1986) and the UKU Side Effect Rating (Lingjaerde, Ahlfors, Bech, Dencker, & Elgen, 1987). However, these instruments are not widely in use because of their length, complexity, and demand on clinicians' time.

1.2 Significance of the Study

A fundamental assumption of this project is that social media technologies may be used as a much needed and important source of data for determining attitudes leading to non-adherence with antidepressant medications. Social media sites show promise as a much needed and important source of data for determining attitudes leading to non-adherence with antidepressant medications. Novel social media technologies have provided patients with a unique platform to freely report their experiences and express their attitudes about healthcare services and treatments. Social media applications with a focus on healthcare topics have been constantly growing in recent years (Metke-Jimenez & Karimi, 2015). The findings of a public opinion survey conducted in 2009 by Pew Research Center's Global Attitudes Project showed that 61

percent of Americans looked online for general health information, 41 percent read others experience and 30 percent were actively participating in creating new knowledge (Fox & Jones, 2012).

Facebook, Twitter, YouTube, and patient forums such as “patientslikeme.com”, and “Online Messaging Boards” (OMB) such as “webmd.com” are popular online platforms used by patients or caregivers for creating online healthcare communities. Patients and caregivers discuss various health concerns and treatment experiences in the communities (Metke-Jimenez & Karimi, 2015). Such communication can happen in various forms, such as sharing videos or images, questions/answers, posts/comments, and expressing emotional reactions using emoticons and “like” buttons (Mao et al., 2013).

Currently, patients’ self-reports about their experience with pharmacological therapy in online communities has received growing attention in the area of psychiatric disorders, indicating that individual’s cognitive representations may be associated with therapeutic outcomes in this group of patients (Richardson et al., 2013). The International Society of Drug Bulletins emphasized in 2005 that “patient reporting systems should periodically sample the scattered drug experiences that patients reported on the internet” (Leaman et al., 2010). The self-reports captured in OMB may provide practical information useful for identifying patient concerns and the root determinants of patients’ specific attitudes or behaviors that might not be traditional self-report tools, such as questionnaire or interview, and in a physician’s office (Benton, Ungar, et al., 2011; Harpaz et al., 2014; Sarker et al., 2015). This is particularly the case with antidepressants.

Regarding the growing emphasis on individuals’ cognitive representations about treatment in the format of self-report evaluation, the ability to measure individuals’ attitudes

toward medications directly from their reviews in social media may increase early detection of factors that contribute to negative outcomes by highlighting aspects of treatment that prompt negative attitudes towards pharmacological treatment (Kane, Kishimoto, & Correll, 2013). In general, signals detected from social media, and in particular from OMBs could be used by pharmaceutical companies and the healthcare system to supplement existing pharmacovigilance systems. Patients in drugs' review may disclose adverse effects associated with a particular drug, such as antidepressants that may not be captured by pharmacovigilance systems (Benton, Ungar, et al., 2011; Harpaz et al., 2014; Sarker et al., 2015).

The premise of this study is that patients' self-reports on OMBs of their experiences with antidepressants therapy may constitute a reliable source to uncover various dimensions of attitude towards these medications. To date, there are no reports in the medical or social sciences literature that focused on identifying underlying factors influencing attitude towards antidepressant treatment as reported by patients in social media, particularly in OMBs.

1.3 Statement of Aims

This study is composed of two main phases. The aim of the first phase is to determine usability of drug reviews in social media in providing insight into medication tolerability, adherence, and perception (attitude) towards antidepressants treatment. We aim to

Aim (1): Explore reliability of drug reviews in social media to identify underlying factors for patients' attitudes and adherence towards antidepressants.

Aim (2): Evaluate usability of drug reviews in social media compare to self-report scales in measuring patients' attitudes towards antidepressants.

The aim of the second phase of this study is to determine usability of drug reviews in social media in detecting ADRs associated antidepressants. We aimed to:

Aim (3): Evaluate reliability of drug review posts in social media for detecting pharmacological aspects of antidepressants, including adverse drug reactions (ADRs), withdrawal symptoms (WD), and drug effectiveness (EF).

Aim (4): Explore the usability of Drug reviews in social media in addressing limitations of self-report scales, such as the Antidepressants Side-Effect Checklist (ASEC)

1.4 Data Source and Methodology of the Study

“Askapatient.com”, a particular OMB, constitutes the source of data for this study. It enables respondents to post reviews about prescribed medications and rate the medications on a scale from 1 to 5 where 1 represents the least satisfaction with the medication and 5 represents the highest satisfaction. In addition, for each review, respondents may disclose clinical and demographic attributes including age, gender, duration of use, and dosages of medication usage. We considered the rating number associated with each review as the respondent’s overall attitude towards the medication. Although, the rating originally indicates respondents’ satisfaction with the medications in “Askapatient.com”, patient’s satisfaction with a medical therapy in several studies was characterized by the patient’s attitude and belief (Taylor & Cronin Jr, 1994; Williams, 1994). Satisfied patients have positive attitude to the therapy, while dissatisfied patients have negative attitude. Therefore, we considered the respondents’ rate for a medication as their overall attitudes, which is equivalent to the ordinal discrete values (Likert scale) that used to represent output of structured self-report scales, such as the the Antidepressant Compliance Questionnaire (ADCQ) and the the Drug Attitude Inventory (DAI).

892 patients’ reviews collected randomly from “askapatient.com” for the four most commonly prescribed antidepressants including Sertraline (Zoloft) and Escitalopram (Lexapro) from SSRI class, and Venlafaxine (Effexor) and duloxetine (Cymbalta) from SNRI class.

A mixed-method approach was used to provide structured data from the unstructured drug reviews and to test the hypotheses. In qualitative analysis, the framework method with hybrid approach of inductive and deductive analysis was used to provide the analytical framework for data analysis. The drug reviews at level of sentences were coded against the analytical framework. In quantitative analysis, the identified adverse drug reactions (ADRs), withdrawal symptoms (WD), and Drug Indications (DIs) were identified and normalized by mapping to both UMLS and SNOMED CT concepts. In addition, statistical methods have been used for developing a predictive model and testing the association between ADRs and patients attitude toward antidepressants.

1.5 Contributions

Major contributions of this study are as follows:

- ① Developing an analytical framework for analyzing patients' self-report experiences of pharmacological treatment in the social media.
- ② Generating structured data from unstructured patients' self-report experiences of psychiatric medications using the Framework Method.
- ③ Using a systematic approach to develop a corpus consisted of three main components: (1) Sentence classification, (2) Entity identification (ADRs, WD, and DIs), and (3) entities normalization: mapping the entities to equivalent medical concepts in both UMLS and SNOMED CT.
- ④ Filling the gap between layperson and professional terminologies of psychiatric medications by identifying semantic links among the expressions of medical terms.
- ⑤ Identifying ADRs and WDs associated with two classes of antidepressants (SSRI and SNRI) using real-world patients.

- ⑥ Identifying underlying factors associated with patients' attitudes towards antidepressants using consumer health posts.
- ⑦ Developing a predictive model of factors affecting patients' attitudes toward antidepressants using data provided by real-world patients.
- ⑧ Measuring association between physiological, psychological, and cognitive ADRs and levels of attitude.

The dataset and guidelines of this study are available at:

<https://sites.google.com/view/pharmacovigilanceinpsychiatry/home>

Chapter 2: Background

2.1 Treatment of Depression

Treatment of depression disorder in patients with depression may include psychotherapy treatment or psychopharmacology treatment. Psychological treatment is characterized by three essential factors: (1) it is designed with the purpose of reducing anxiety and depression symptoms, 2) it is based on psychological theory, such as learning theory, and 3) it involves a structured interaction between a facilitator and a patient (Orgeta, Qazi, Spector, & Orrell, 2015). Eligible interventions for psychological treatment include cognitive behavioral therapies, relaxation training therapies, psychodynamic therapies, interpersonal therapies, and supportive or counseling therapies. McHugh, Whitton, Peckham, Welge, and Otto (2013) using a meta-analytic review showed that patient preferences across diverse settings yielded a significant three-fold preference for psychological treatment relative to pharmacology treatment. Therefore, improving access to psychotherapy treatment is needed to connect more patients to their preferred treatment. However, the focus of this study is on pharmacological treatment.

Over the years, antidepressants as psychopharmacology treatment of depression have become the major source for treatment of depression, particularly for moderate to severe depression. Chisholm (2015) has introduced antidepressants as a cost-effective choice for depression treatment that is quite affordable and feasible to be prescribed in primary care settings. From 1999 to 2012, the percentage of Americans used antidepressants increased from 6.8 percent to 13 percent (Karter, 2015). In 2010, antidepressants were the second most commonly prescribed medication after cholesterol medications, and about 254 million prescriptions were written, resulting in nearly \$ ten billion (B. L. Smith, 2012).

2.2 Psychopharmacology Treatment VS. Psychological Treatment

There is a wide range of pharmacological and psychological treatment that healthcare professionals use to treat patients with depression. The lay public's preference is psychotherapy, and psychopharmacology treatment is rejected by most participants in studies conducted in the United States (US), United Kingdom (UK), Germany, and Australia. Although the preference of patients with depression is psychological treatment, financial factors such as low reimbursement by insurance companies and high out of pocket costs have led to increasing rate of pharmacological treatment for depression (Olfson & Marcus, 2009).

Because this study aimed to identify the underlying factors affecting attitudes towards antidepressants, the focus of the rest of this study is on psychopharmacology treatment of depression.

2.3 Types of Antidepressants

Antidepressants for forty years were the core pharmacological treatment for depression. Tricyclic antidepressants (TCAs) were the first antidepressants introduced to market in the 1950s (Ferguson, 2001). The efficacy of this drug was linked to mood-elevating properties. Although TCAs were in practice for several years, anticholinergic and cardiac side-effects, the risk of morbidity and mortality in overdose cases, and the lack of specific treatments led to development of new classes of antidepressants called Selective Serotonin Reuptake Inhibitors (SSRIs) and Serotonin–Norepinephrine Reuptake Inhibitors (SNRIs).

2.3.1 Selective serotonin reuptake inhibitors (SSRIs)

SSRI class is currently the most common antidepressants recommended to be prescribed for patients suffering from different phases of depression (Sayyah, Eslami, AlaiShehni, & Kouti, 2016). This class encompasses five agents: sertraline (brand name: Zoloft), Escitalopram (brand

name: Lexapro), paroxetine (brand name: Paxil), citalopram (brand name: Celexa), and fluoxetine (brand name: Prozac). SSRIs in comparison with TCAs use the same property (serotonin reuptake inhibitors) for mood elevation. However, the mechanism in SSRIs is selective, which causes fewer adverse reactions and therefore enhances the safety of the drugs. Since SSRIs treatment has reduced the risk of pharmacotherapy, antidepressants now can be prescribed for patients with mild or moderate phases of depression, possibly even subsyndromal symptomatic depression (SSD) (Santarsieri & Schwartz, 2015). Some SSRI agents, in addition to approval for treating major depression, received approval for treating other mental disorders such as obsessive-compulsive disorder and social anxiety disorder. Sertraline and paroxetine have the highest approval by the FDA to treat a wide range of mental disorders.

SSRI agents have subtle yet notable differences in adverse effects and effectiveness. For example, Fluoxetine's efficacious aspect takes longer to emerge in comparison with other SSRIs, and Escitalopram has better effectiveness in comparison with Citalopram (Gorman, Korotzer, & Su, 2002). All SSRIs may cause physiological and psychiatric adverse effects, such as gastrointestinal side effects, insomnia, weight gain, and fatigue. Sertraline may lead to more gastrointestinal disturbance than others, while Citalopram and Paroxetine are accompanied with more sedative side effects. In terms of withdrawal symptoms, Sertraline and Paroxetine, in the case of sudden discontinuation, can cause more severe effects. SSRIs plus the use of other medications may also lead to different drug-drug interactions. Lower adverse effects and the effectiveness of SSRIs, as well as diagnostic factors such as phase of depression and patient conditions, cause healthcare providers to prescribe a SSRI.

2.3.2 Serotonin–Norepinephrine Reuptake Inhibitors (SNRIs)

The mechanism of Serotonin–Norepinephrine Reuptake Inhibitors (SNRIs) is by inhibiting reuptake of serotonin and norepinephrine, which is similar to the mechanism of TCA class, but like the SSRI class, SNRIs have less adverse effects. Therefore, SNRIs are prescribed for a wide range of psychotic disorders. SNRIs include mainly Venlafaxine (brand name: Effexor), Desvenlafaxine (brand name: Pristiq), Duloxetine (brand name: Cymbalta) and Levomilnacipran (brand name: Fetzima). Like the SSRI class, SNRIs may cause physiological and psychiatric adverse effects, but they tend to induce more nausea, insomnia, dry mouth, and, in rare cases, elevated blood pressure (Santarsieri & Schwartz, 2015).

In terms of efficacy, SNRI agents may show higher effectiveness than SSRI agents. For example, results of three studies have shown that a greater number of patients treated with Venlafaxine showed depression symptom remission compared with patients on SSRIs and on placebos (Nemeroff et al., 2003; D. Smith, Dempster, Glanville, Freemantle, & Anderson, 2002; Thase, Entsuah, & Rudolph, 2001). On the other hand, one study did not find any significant statistical difference between the efficacy of Duloxetine and SSRI agents (Goldstein et al., 2004).

According to the dataset from Symphony Health Solutions with one year of prescription-filling data (from July 1, 2011 to June 30, 2012) for patients from Washington DC, Maryland, Virginia and West Virginia, Sertraline (Zoloft), Escitalopram (Lexapro), Venlafaxine (Effexor), and duloxetine (Cymbalta) are the four most common antidepressants prescribed for depression.

2.4 Experience of Patients With Antidepressants

The experiences of people using antidepressants vary considerably. It has been accepted that 10%-30% of people with mental disorders do not respond to medications (Al-Harbi, 2012). Some patients found the medications very helpful in treating symptoms of the disorders and they

could tolerate troublesome adverse effects. Others found the drugs ineffective and believed the risks associated with adverse effects outweighed the beneficial aspects of antidepressants and therefore had a low level of acceptance of the drugs (De las Cuevas & Sanz, 2007).

2.5 Duration of Antidepressants Usage

According to NCCMH (2010), antidepressant treatment should be continued at least for six months to reduce risk of relapse, even if the therapy goal was achieved and the full remission of depression symptoms was observed. More importantly, the relapse rate would be lower if the acute treatment dosage was maintained rather than reducing the dosage over the six months (NCCMH, 2010).

2.6 Adherence to Antidepressant Treatment

Despite the clinical trials' proof of antidepressants efficacy, non-adherence to treatment is a major barrier to their effectiveness in clinical practice (López-Torres et al., 2013). Overall, therapeutic benefits of antidepressant treatment depend on maintaining an appropriate drug regimen for a certain period time and a proper method of discontinuation. Many patients recognize the necessity of treatment due to warnings by their healthcare providers about the importance of adherence (Murata et al., 2012). Unfortunately, a significant number of individuals diagnosed with depression do not follow the prescription and they may discontinue antidepressant use abruptly, which may lead to severe withdrawal symptoms, increasing emergency room visits, and even hospitalization (Grenard et al., 2011). Research showed that about 50 percent of patients with depression discontinue their medication during the first month of treatment, and nearly 68 percent of patients, discontinue their medication in the first 3 months of treatment (De las Cuevas et al., 2014). Results of other studies indicate that between 30

percent and 68 percent of patients with depression did not complete their prescribed course of antidepressant treatment (Aikens et al., 2005).

2.6.1 Importance of adherence to antidepressant treatment

Adherence to antidepressant regimes is important to achieve expected clinical outcomes. According to one meta-analysis, efficacy of antidepressants cannot be achieved with poor adherence in patients with depression (Grenard et al., 2011). That underlies the need for innovations to assist patients in following their prescription. Non-adherence can also increase the risk of relapse (Alekhya et al., 2015; NCCMH, 2010). The ratio of mental illness relapse is about five times greater in non-adherent patients compared to adherent patients (Masand, Roca, Turner, & Kane, 2009). Relapse of depression, in turn, may affect long-term prognosis, severity of depression recurrence, and difficulty of treatment. Moreover, non-adherence to antidepressants is associated with increased rate of emergency visits, longer hospital admission, low quality of life, and significant cost at individual, family, and health care system levels (Vives et al., 2014).

2.6.2 Factors Affecting on non-Adherence behavior

In response to the substantial negative impact of non-adherence behavior, identifying high risk factors associated with non-adherent patients is important from the view of both clinical practice and interventions targeting these patients (Rivero-Santana, Perestelo-Perez, Pérez-Ramos, Serrano-Aguilar, & De las Cuevas, 2013). Several studies identified predictive factors associated with non-adherence behavior in pharmacological treatment. Jin, Sklar, Oh, and Li (2008) in a literature review categorized potential predictive factors for non-adherence in five categories:

- Patient-centered factors, such as demographic factors and patient-physician relationship;
- Therapy-related factors, such as treatment complexity and medication adverse effects;

- Healthcare system factors, such as lack of accessibility;
- Social and economic factors, such as social support and treatment affordability; and
- Disease factors, such as severity and duration of a disease.

Research focused on antidepressant non-adherence found a broad range of factors associated with premature discontinuation of these drugs. Likewise, in the categorical system provided by Jin et al. (2008), the potential predictive factors for antidepressants non-adherence can also be summarized in the same five groups (**Table 2-1**). Examples of factors that can be included in the categories of patient-centered factors are patient forgetfulness (Bulloch & Patten, 2010), patient's specific personality type or issue such as being extroverted or having a personality disorder (Akerblad, Bengtsson, Holgersson, von Knorring, & Ekselius, 2008; Holma, Holma, Melartin, & Isometsä, 2010; Woolley, Fredman, Goethe, Lincoln, & Heeren, 2010), substance abuse (Holma et al., 2010), and low motivation to continue treatment (Masand, 2003). The group of therapy-related factors can include antidepressants' adverse effects (Fortney et al., 2011) such as sexual dysfunction, (Cohen, Kühn, Sträter, Scherbaum, & Weig, 2010), concern about adverse effects, such as possibility of addiction (Brown et al., 2005), impact on personality (Chakraborty et al., 2009), and being doubtful about drug effectiveness (Hoencamp, Stevens, & Haffmans, 2002). The group of socioeconomic factors contains issues such as cost of antidepressants and lack of coverage by insurance companies. This may lead to the erroneous views in patients that antidepressants are not necessary (Kennedy, Tuleu, & Mackay, 2008). The group of healthcare system factors includes poor instruction of intake (Woolley et al., 2010), and lack of follow-up care by clinicians (Masand, 2003). Finally, the group of disease factors include patient's perceived severity of depression and duration of symptoms (Demyttenaere et al., 2008).

Table 0-1 Potential Predictive Factors Affecting Non-adherence towards Antidepressants

Category	Example of potential Factors in each category
Patient-centered factors	Patient forgetfulness, specific personality type
Therapy-related factors	Antidepressants' adverse effects, drug ineffectiveness
Socioeconomic factors	Cost of antidepressants
Healthcare system factors	Lack of follow-up care by clinicians, instruction of intake
Disease factors	Perceived severity of depression and duration of symptoms

There are several inconsistencies in the literature regarding the significant factors affecting drugs adherence in patients with depression. For example, from ten studies that tested the impact of depression severity on non-adherence, only two of them concluded that depression severity was a significant predictive factor (Rivero-Santana et al., 2013). Among studies that evaluated the relationship between basic demographic factors and non-adherence, there is no consistency on the findings related to gender, educational level, living situation, income, marital status, or employment status (Rivero-Santana et al., 2013).

One of the major reasons for this inconsistency is the complexity of non-adherence and the underlying factors. For example, education level might not directly be related to non-adherence behavior (Rivero-Santana et al., 2013). In fact, other confounding variables such as attitude towards medication may drive any observed association between education and non-adherence.

2.6.2.1 Attitude as a significant predicative factor for non-adherence to antidepressants

Patients' self-reports about medication effectiveness have received growing attention in the area of psychopharmacology. Cognitive representations, that is, the person's thoughts about their medication, may be associated with therapeutic outcomes (Richardson et al., 2013). This perspective is supported by cognitive representation models such as the Health Belief Model (HBM) (Green & Murphy, 2014) and the Population and Theory of Planned Behavior (TPB) (Ajzen, 1985). According to these models, the intention and motivation of patients to present a

specific health behavior, such as taking medication, is a function of three factors: 1) individual attitude towards the behavior, 2) perception of social norms regarding the behavior, and 3) intention to perform the behavior.

Attitude is a tendency or disposition of an individual to a particular act (Bergman, 1998), representing the overall evaluation of that act. In other words, attitude is a function of an individual's belief, which is the base of comparing and contrasting potential harms and benefits of a specific behavior (Richardson et al., 2013). For example, a patient who believes that taking antidepressants causes more harm than good has a negative attitude towards taking antidepressants.

There is well established that understanding patients views about their problems and their treatments can significantly improve process of treatment (Moncrieff, Cohen, & Mason, 2009). Aikens et al. (2005) emphasized the importance of association between patients' perceptions about medication and the clinical outcome, such as adherence. Aikens et al. (2005) showed that outcomes of the interventions targeting non-adherence behaviors (such as intensified follow up plan) ultimately depend on the patients' willingness to take the drugs.

Research showed that across chronic medical illness in general, non-adherence is loosely associated with demographic information and clinical variables, such as severity of the symptoms or type of drug adverse effects, but more strongly it is associated with patients' beliefs and attitudes, perceived necessity, and perceived medication risks (Aikens & Klinkman, 2012), (Acosta et al., 2013; Aikens et al., 2008; Brown et al., 2005; Chakraborty et al., 2009; De las Cuevas et al., 2014; López-Torres et al., 2013; Richardson et al., 2013; Verdoux et al., 2000). It is more likely that patients with positive attitudes show better adherence behavior than patients with negative or indifferent attitudes (Aikens et al., 2005; De las Cuevas et al., 2014; Richardson

et al., 2013). It is ultimately patient perception of beneficial aspects of the medication versus the harmful aspects that will determine the acceptance or rejection of a medication (Christensen, 2004) (Acosta et al., 2013). For example, recent studies found that a majority of patients on antidepressants have erroneous views towards these drugs, which in turn influence adherence behavior (Chakraborty et al., 2009) (Jacob, Ab Fatah Ab Rahman, & Hassali, 2015). A study with a sample of 573 primary care patients indicated that the only identifiable baseline predictive factor of early discontinuation was belief about the appropriateness of taking antidepressants (Aikens & Klinkman, 2012). Moreover, a significant negative correlation has been found between patients' attitudes and beliefs toward antidepressants and the percentage of days of missed dosages (Jacob et al., 2015). On the other hand, Hung, YIaCHAO, and JAUa (2014) found that the main reason for patients continued acceptance of depression treatment was a positive view toward antidepressants effectiveness.

“Patients' perspective towards antidepressants treatment is complex, neglected and ambivalent” (Demyttenaere et al., 2004). It has been evident that identifying potential predictive factors for attitude may have significant implications for understanding the complexity of this phenomenon and therefore, designing effective interventions to improve drug therapy adherence (Kessing et al., 2005). Patients with depression may have some specific beliefs and perceptions that may lead to failure to follow antidepressant instructions. For example, concerns about addictive possibilities and perceived distress received from adverse effects are significant factors leading to negative attitude toward antidepressants and consequently non-adherence behavior (Fawzi et al., 2012).

According to Horne's theoretical model (Lin et al., 1995), beliefs about medication in general seem to be most relevant for adherence at the start of treatment (Vergouwen, Burger,

Verheij, & Koerselman, 2009). For example, believing that antidepressants are associated with harmful aspects increase the patients' sensitivity and decrease the patients' tolerability for antidepressants.

Aikens and Klinkman (2012) suggested that there are different factors affecting medication adherence in the acute and maintenance phase of antidepressant treatment. In the acute phase of treatment, a change in functional health predicts change in medication beliefs. Patients' perceptions become more positive as medication becomes effective. On the other hand, occurrence of side effects reinforces the belief that there is harmful aspect of the drug. In the maintenance phase, patients' beliefs about antidepressant necessity versus harmfulness were the only identifiable predictor of adherence. Perceived necessity for antidepressant treatment is directly connected to the antidepressants perceived effectiveness. The perceptions of potential harm is connected to adverse effects and withdrawal symptoms in the case of missing a dosage or dose reduction (Aikens et al., 2005).

2.7 Potential Predictive Factors Associated With Attitude Towards Antidepressants

Attitudes of patients and caregivers toward chronic psychiatric disorders appeared to be shaped by pharmacological treatment factors, healthcare system factors, social-cognitive and psychological factors, patient-related factors, and depression factors.

Table 0-2 Factors Affecting Patients Attitude Towards Antidepressants (Suggested by Literature)

Category	Factors in each category
Pharmacological treatment factors	<ul style="list-style-type: none"> • Perceived effectiveness • Perceived necessity • Perceived concerns • Adverse drug reaction • Perceived distress from adverse effects
Healthcare system factors	<ul style="list-style-type: none"> • Patient-provider relationship • Healthcare settings • Affordability
Social-cognitive and psychological factors	<ul style="list-style-type: none"> • Stigma and cultural related factors • Partner supports
Patient-related factors	<ul style="list-style-type: none"> • General concern and necessity • Knowledge about pharmacological aspects of medication • Socio-demographic factors • Educational level
Depression factors	<ul style="list-style-type: none"> • Depression severity, type, and duration • Patient insight about depression

Table 2-2 shows the categories and factors in each category affecting patient attitudes towards antidepressants. The following sections explain the factors and their association with attitudes.

2.7.1 Pharmacological treatment factors

Identified factors in the pharmacological treatment category include medication perceived effectiveness, adverse effect reaction (side-effects), perceived distress from side effects, perceived necessity, and perceived concern of treatment.

2.7.1.1 Perceived effectiveness

Drug effectiveness is a strong predictive factor for shaping patients' attitudes in both the acute and the maintenance phase of treatment. Patients with depression generally express the effectiveness of antidepressants as subjective experiences that mostly indicate the impact on their quality of life. The patients commonly suffer from emotional problems and cognitive impairments, such as "stress" and "slow information processing speed" (Culang-Reinlieb, Sneed,

Keilp, & Roose, 2012) that restrict their capabilities in daily functioning and social participations. Therefore, from patients' viewpoints, antidepressant effectiveness is the extent to which they help patients reduce emotional problems (Demyttenaere et al., 2004; Ferguson, 2001), improve cognitive functionality, such as psychomotor speed and ability to plan and initiate a behavior (Demyttenaere et al., 2004; Reilly JL, 2011), and enhance their coping mechanisms, such as dealing with stress or difficult relationships (Demyttenaere et al., 2004; Prins, Verhaak, Bensing, & van der Meer, 2008).

2.7.1.2 Perceived necessity

From the layperson's viewpoint, the necessity of antidepressants as the main source of treatment is much lower than other methods of treatment, such as psychotherapy. People usually view counseling as the best solution for depression treatment (Demyttenaere et al., 2004). Therefore, it is not surprising that the perspective of patients with mental disorders regarding the perceived necessity of pharmacological treatment is in line with this viewpoint. From 49% to 84% of patients with mental disorders who are looking for treatment feel that pharmacological treatment is not a desirable solution (Prins et al., 2008). A plausible implication of these findings is that patients with depression may discontinue taking medication in the acute phase of treatment if they find perceived effectiveness less than expectations. Surprisingly, they may also discontinue the treatment prematurely if they feel better, assuming they no longer need medication to feel better.

2.7.1.3 Perceived concerns

Patients' perceived concerns about antidepressants mechanism and its long-term effects on patients' quality of life are strongly associated with patients' negative attitudes towards antidepressants (Aikens et al., 2005; Benton, Hill, et al., 2011; Demyttenaere et al., 2004; Horne,

Weinman, & Hankins, 1999; Hunot, Horne, Leese, & Churchill, 2007). The main concerns of patients on antidepressants are related to addictive possibilities (Chakraborty et al., 2009; Jacob et al., 2015; Prins et al., 2008) and immunity to antidepressants (Kessing et al., 2005).

Demyttenaere et al. (2004) in a self-report scale, the Antidepressant Compliance Questionnaire (ADCQ), designed for assessing patients attitudes toward antidepressants introduced four dimensions to measure patients' concerns including addictive possibility of antidepressants, control over feelings and thoughts, altering personality, and immunity to antidepressants. In another self-report scale, the Beliefs about Medicine Questionnaire (BMQ) (Horne et al. (1999), perceived concern is measured based on three dimensions: 1) long-term effects and its impact on quality of life; 2) concern about the mechanism and addictive possibilities of antidepressants; and 3) concern about losing autonomy because of the effects of the medication.

2.7.1.4 Adverse drug reaction

Adverse effects of antidepressant treatment correlate with the perceived harmfulness of antidepressants (Aikens et al., 2005). This provokes negative attitudes towards the drugs (Dougherty, Klein, Olino, Dyson, & Rose, 2009; Murata et al., 2012; Ng, Smith, King, Ong, & Schweitzer, 2012), and in turn leads to patients' non-adherence (De las Cuevas et al., 2014; Dougherty et al., 2009). Cash and Brown (2000) reported that related adverse-effects of weight gain influence patient attitude toward antidepressants and therefore lowers patients' acceptability of the drugs. Bradley, McGrath, Brannen, and Bagnell (2010) also found that among adolescents, the most adverse effects that made adherence to antidepressant treatment difficult were the increase of weight among girls and the sexual side effects among boys.

Patients may express adverse effects of antidepressants as physiological symptoms that are directly correlated with pharmacological aspects of the medications. But there is still a large

number of patients who express adverse effects in terms of subjective complaints (Hogan et al., 1983) such as “not feeling like themselves”, “ unable to sit”, or “ alcohol craving”. These adverse effects are recognized as behavior toxicity of pharmacological treatment and may reflect extrapyramidal or autonomic side-effects, or subtle and frequently syndromes of akathisia (Hogan et al., 1983). Further, patients may report side effects as emotional problems (Kikuchi, Uchida, Suzuki, Watanabe, & Kashima, 2011; Price, Cole, & Goodwin, 2009; Sansone & Sansone, 2010) or cognitive dysfunction (Sayyah et al., 2016) such as “aggressiveness” or “brain fog”. In addition, impacts of adverse effects on daily functioning and social participation (Giannangelo, Bowman, Dougherty, & Fenton, 2005) and in overall impact on quality of life, such as job loss are determinant factors that influence patients’ attitudes towards antidepressants (Hofer et al., 2004). Self report scales for measuring ADRs associated with antidepressants are reported in Appendix A.

2.7.1.5 Perceived distress from adverse effects

“There is preliminary evidence that pharmacogenetic variations may affect the efficacy and tolerability of antidepressant drugs” (NCCMH, 2010). It is also conceivable that patients self-attention to internal bodily sensation can vary from patient to patient and this can lead to their expressing a range of intensity for identical physiological and emotional pain (Hogan et al., 1983). The perceived pain (distress) from medication can influence patient attitude toward medication and in turn can determine patient drug intake behavior. For example, patient may frequently forget to follow a prescription. This information implies that in addition to a specific ADR and the level severity associated with it, perceived stress received from adverse reaction is another treatment factor affecting on patient attitude toward medication. For example, a young female patient experiencing a small weight gain may receive high perceived distress that may

result in antidepressant discontinuation. While for another patient who experienced the same amount of weight gain, the level of perceived distress from this adverse effect may be low.

Overall, patient distress from antidepressants adverse effects is not only a function of patient tolerability, but also a function of perceived stress received from depression symptoms and also concern about the long term effect of adverse effects on mental and physical functioning.

Regarding the importance of perceived distress in shaping attitudes toward antidepressants, educational strategies designed merely to inform patients the type of side-effects may be relatively ineffective in patients whose interceptive cues provide less than favorable information about the medication in their body system.

2.7.2 Healthcare system factors

Patient-provider relationship and healthcare setting are factors that are associated with medication attitude. Few studies discussed medications affordability as factors that may influence attitudes towards medications.

2.7.2.1 Patient-provider relationship

Several studies emphasized patient positive experience as a strong predictive factor of patients' positive attitudes toward medications and adherence behavior (Chakraborty et al., 2009; Day et al., 2005; Demyttenaere et al., 2004; Hunot et al., 2007; Prins et al., 2008; Vives et al., 2014). The importance of this interaction is not limited to the prescriber-patient relationship. Day et al. (2005) suggested that patients' admission experiences and relationship with healthcare staffs also influence patients' attitudes toward medication. Studies have found that in a patient-centered approach where providers actively seek to understand patients' perspectives to treatment and respect patients' preferences by involving them in the process of treatment

decision making and providing supportive continued management, providers are most likely to enhance and maintain patients positive attitude toward therapy (Day et al., 2005; Hunot et al., 2007).

The Antidepressants Compliance Questionnaire (ADCQ) (Demyttenaere et al., 2004), incorporated patient-physician interaction as one of the main components for assessing patient attitude toward antidepressants. Patients' perceptions of physicians' knowledge, patients' perceptions of the sufficiency of knowledge provided by physicians about disorder and treatment process, and patients' perceptions of communication effectiveness such as the perception of clinicians interest to a patient's problem and the level of support received from clinicians are three dimensions that ADCQ addressed in assessing patient physician interaction Demyttenaere et al. (2004).

2.7.2.2 Healthcare settings

Few studies evaluated the relationship between patients' attitudes toward psychotic medications and healthcare settings, i.e. primary care sites vs. psychiatric sites. Aikens et al. (2005) suggested that patients with depression treated in primary care settings are more likely to discontinue antidepressants at initial stages. A plausible explanation is that patients have more uncertainty about diagnosis results and prescribed treatment in primary care settings compared with psychiatric settings. This uncertainty can lower patients' perceived necessity for antidepressants, which may result in a negative attitude toward medications.

2.7.2.3 Affordability

Affordability is a significant factor for predicting patient adherence behavior (Jin et al., 2008). Nevertheless, studies related to attitude did not consider this factor as a potential predictive factor for attitude toward antidepressants, but Weiden et al. (1994) suggested lack of

insurance coverage and unaffordability of a drug may reduce the necessity of medication from a patient's view point.

2.7.3 Psycho-social factors

Psycho-social factors that are significantly associated with attitude towards medication are stigma and cultural related factors, and perceived support from partners as well as family and friends.

2.7.3.1 Stigma and cultural related factors

Stigma against depression and antidepressants are significant factors for negative attitude and premature discontinuation in patients with depression (Acosta et al., 2013; Jacob et al., 2015; Malpass et al., 2009). Stigma about the use of antidepressants is similar to the stigma against depression, such as perceived emotional and mental weakness and inability to cope with daily problems. However, lack of belief in the therapeutic efficacy of antidepressants and showing non-adherence behavior are indicators of avoiding to being labeled as mentally ill (Acosta et al., 2013; Castaldelli-Maia et al., 2011).

2.7.3.2 Partner support

Patients perceived support from partners, family, and friends is associated with patients positive attitude towards antidepressants (Demyttenaere et al., 2004). This factor is incorporated as one of the components in the Antidepressants Compliance Questionnaire (ADCQ). Items in this component measure patient's perceived support from partners regarding their views about being diagnosed with depression and treated with antidepressants (Demyttenaere et al., 2004).

2.7.4 Patient related factors

Patients related factors that may influence patients' attitudes about antidepressants are general concern and necessity of medications, knowledge about antidepressants' mechanisms and adverse effects, socio-demographic information, and patient's education.

2.7.4.1 General concern and necessity

Horne et al. (1999) in the Concern-Necessity Framework discussed that patients' general beliefs about the necessity and harmfulness of medications are significant factors affecting patients' attitudes toward prescribed medications and adherence behavior. Patients who have concerns about pharmacological treatment and believe that natural remedies or changing their life style will have better healthcare outcomes than pharmacological treatment are more likely to reject a prescribed medication or show non-adherence behavior, particularly at initial phase of treatment.

2.7.4.2 Knowledge about medication

Clarifying the purpose of a medication, possible physiological adverse effects, the nature of regimen, and the proper time for discontinuation are a set of pharmacology knowledge that healthcare providers need to share with patients with depression.

A qualitative study (Haslam, Atkinson, Brown, & Haslam, 2005) showed that many patients taking antidepressants suffer from side-effects that they felt they were not sufficiently informed about them and mechanisms of their management. Gabriel and Violato (2010) discussed that patients with more knowledge about their illness and their treatment are likely to be have positive perception about treatment and, in tern, to be more adherent.

2.7.4.3 Socio-demographic factors

Gender, age, ethnicity, and education level are important predisposing characteristics influencing attitude and perceived needs in people with depression.

Women are two times more likely to seek treatment for depression as they believe less in the addictiveness of antidepressants (Prins et al., 2008). Young, senior, and minority-group patients are more vulnerable having a low perceived need and a negative attitude towards treatment. Young people prefer to cope with depression symptoms without seeking for help, because they view depression as a sign of weakness and feel embarrassed to express their feelings. On the other hand in the elderly, diagnosis is usually missed, particularly in patients with a stressful life and retirement (Prins et al., 2008). Minority groups tend to have less access to healthcare services, and they also have a different perceived need and attitude toward pharmacological treatment (Scheppers, Van Dongen, Dekker, Geertzen, & Dekker, 2006). Contrary with Prins et al. (2008) findings pertinent to age and gender, some studies found that age and gender are not significant predictive factors for attitude towards antidepressant (Jacob et al., 2015; Murata et al., 2012; Ng et al., 2012).

2.7.4.4 Educational level

Education level is loosely associated with attitude toward medication and belief toward antidepressants (Murata et al., 2012; Ng et al., 2012). Intuitively, it may be expected that individuals with a higher education are more motivated to follow prescriptions. However, patients with a higher education may be more skeptical about the mechanism of antidepressants or the accuracy of diagnosis. In addition, they may have more access to alternative sources of treatment, thereby they will show less positive attitudes toward antidepressants compared to patients with low education. On the other hand, patients with a low education level may have more trust in a physicians' diagnosis and treatment plan (Jin et al., 2008).

2.7.5 Depression disorder characteristics

Depression disorder characteristics including type, severity, and duration and patient insight about depression characteristics may influence patients attitude toward antidepressants.

2.7.5.1 Depression severity, type, and duration

“Patients with major depressive disorder appeared to have a more negative view of antidepressants than those with bipolar disorder” (Acosta et al., 2013). Murata et al. (2012) also found that there is no correlation between types of depression including melancholic, nonmelancholic, and bipolar depression and attitude toward antidepressants. In terms of duration, Ng et al. (2012)) did not find a significant correlation between the duration or severity of depression and the patients attitude was found.

2.7.5.2 Patient insight about depression

A majority of patients believe that depression is a psychological problem rather than a biological problem (Hansson, Chotai, & Bodlund, 2010; Jacob et al., 2015; Jorm, Christensen, & Griffiths, 2005). This erroneous view causes patients to reject biological interventions, i.e., antidepressants in moderate- to-severe cases, as they feel they can heal themselves by simply changing their behavior or personality(Prins et al., 2008)

The factors (explained in this section) identified by literature as important factors affecting attitude will be used for developing initial themes for data analysis. The the detail of data analysis procedure is explained in the methodology section.

2.8 Self-Report Scales Used in Antidepressants Studies for Measuring Attitude

Identifying dimensions of patients attitude towards psychotic medications, particularly antidepressants is a challenging work, partly due to the lack of reliable and comprehensive methods to capture predictive factors affecting patients' attitudes. Few studies have attempted to develop structured methods, such as self-report measures, to operationalize the concept of attitude towards psychiatric drugs in various scales (Demyttenaere et al., 2004; Hogan et al., 1983; Horne et al., 1999; López-Torres et al., 2013). The well-validated self-report scales for assessment of attitude among patients with psychotic disorders include the Drug Attitude Inventory (DAI) (Hogan et al., 1983), The Rating of Medication Influence Scale (ROMI) (Weiden et al., 1994), the Antidepressants Compliance Questionnaire (ADCQ) (Demyttenaere et al., 2008), and the Beliefs about Medicines Questionnaire (BMQ) (Hogan et al., 1983). The detail of the self-report scales is presented in **Appendix B**.

2.8.1 Methodological Limitations of self-report scales

Self report scales, including the ADI (Drug Attitude Inventory), the ADCQ (Antidepressant Compliance Questionnaire), the BMQ (Beliefs about Medications Questionnaire), and the ROMI (Rating of Medication Influence Scale), are non-intrusive and efficient to capture information about patients' attitudes and beliefs toward psychiatric medications (Nunkoosing, 2005). However, they are associated with some limitations.

First, these tools are different in dimensions used for measuring attitude and the impact of medications on patients' quality of life (Richardson et al., 2013) Wolters, Kneegtering, van den Bosch, and Wiersma (2009). While ADI, ADCQ, and BMQ assigned some items for assessing patient's perceived necessity and concern directly or indirectly, ROMI did not include factors for measuring patient concern about medications. In terms of healthcare settings, ADCQ and ROMI

included items expressing quality of patient-physician relationship and also psychosocial support, but ADI and BMQ did not. BMQ is unique in terms of assessing patient belief in general towards medications. This factor has a significant role in adherence behavior at initial stage of treatment. In addition, ROMI is unique in suggesting items assessing patient perception of illness.

Second, the focus of the questionnaires is on extrapyramidal side effects and patients' subjective experiences with medications, whereas physiological side effects such as muscle spasm may have a strong impact on forming patients' attitudes toward medications. Studies interested in finding association between side effects and attitude, need to administer other questionnaires, such as self-report antidepressant Side-Effect Checklist (ASEC) or Liverpool University Neuroleptic Side Effect Rating Scale (LUNSERS) Ng et al. (2012).

Third, the DAI, the ROMI, the BMQ, and the ANT questionnaire are primarily developed to assess attitudes toward antipsychotics. They have thus been mainly used among patients with psychosis or schizophrenia, and few studies used them in patients with depression (Grover, Chakrabarti, Sharma, & Tyagi, 2014).

Fourth, studies employing the scales represented the scales' output score in a continuous range, in a set of ordinal discrete values (such as a Likert scale) or in a binary construct (positive attitude vs. negative attitude). Although, representing data in an ordinal discrete values may capture meaningful variance in data to some extent, a binary construct discards the variance. To simplify process of data analysis, most of the studies represented output in a binary construct.

Fifth, self-report questionnaires depend on individuals' ability to understand items and their willingness to reveal personal information. Patients tends to report minimally with a

reluctance to be critical (Nunkoosing, 2005). Participant answers may be biased because of “social desirability despite its anonymity and confidentiality and distance from the researchers or those responsible for clinical care” De las Cuevas and Sanz (2007). Furthermore, reliability of questionnaires may be affected by patients’ memory recall, specifically for questions that need long term recall period or are not specified by time period. These limitations can affect response accuracy and consequently the validity and reliability of the questionnaire.

Sixth, there is always this risk that patients with most the positive or most negative attitude may have a high tendency to participate in studies focused on measuring attitude or adherence, while patients with indifferent attitudes may be less willing to participate in these studies. This problem can lead to sample bias and thereby limit generalizability of studies’ findings to the population.

The premise of this study is that patients’ self-reports experience in online healthcare forums may constitute a reliable source to uncover various dimensions of attitude towards the medications, and in turn, addresses the limitations of self-report scales (questionnaires) in measuring patients attitudes toward antidepressants.

2.9 Social Media and its Application in Healthcare

Novel social media technologies have provided patients with a unique platform to freely report their experiences and express their attitudes about healthcare services and treatments. The number of social media applications with a focus on healthcare topics has been constantly growing in recent years (Metke-Jimenez & Karimi, 2015). The findings of a public opinion survey conducted in 2009 by Pew Research Center's Global Attitudes Project showed that 61 percent of Americans looked online for general health information, 41 percent read others experience and 30 percent were actively participating in creating new knowledge (Fox & Jones, 2012).

Patients with mental disorders usually prefer to share health experiences and concerns with each other rather than in clinical research studies or with their healthcare providers (Blenkinsopp, Wilkie, Wang, & Routledge, 2007; Leaman et al., 2010). This may be stigma associated with mental disorders and help seeking for the conditions (Griffiths et al., 2009).

Currently, many patients' self-reports about their experience with pharmacological therapy in online communities are specifically in the area of adverse drug reactions. The International Society of Drug Bulletins emphasized in 2005 that "patient reporting systems should periodically sample the scattered drug experiences that patients reported on the internet" (Leaman et al., 2010).

Previous studies showed that clinical trials and post-marketing surveillance systems established by regulatory agencies, such as the Adverse Event Reporting System (AERS) of Food and Drug Administration (FDA), are not able to detect potential risks of drugs pre and post marketing. It is estimated that such surveillance systems capture less than 10 % of the Adverse Drug Reaction (ADR) occurrence, due to the voluntary nature of the systems in data collection

and perhaps patients' negative perceptions of the systems (Yang, Yang, Jiang, & Zhang, 2012). These limitations have led to major concerns in public health; thousands of incidents of hospitalizations and deaths have occurred due to undetected and uncontrolled risks (Karimi, Metke-Jimenez, Kemp, & Wang, 2015).

According to recent studies showed, patients'-self reports of ADRs and drug efficacy to surveillance systems have the same quality as reports by healthcare professionals (Uher et al., 2009). As such, patients' self-reports are used as a reliable source for risk discovery through programs such as FDA's MedWatch program or the UK MHRA's Yellow Card Scheme. However, many patients do not report to these systems, perhaps due to negative attitude towards providers, ignorance of the availability of these systems, or their severity of illness (Yang et al., 2012). Instead, they prefer to report and discuss their detailed experience with prescribed medications in various social media platforms, such as online support groups and message boards. Accordingly, these social media platforms have turned out to be reliable sources for discovering various aspects of medications benefits and risks, such as drugs' adverse effects, drug effectiveness, and drug impacts on patients' quality of life.

2.10 Challenges in Health Information Extraction from Consumer Health Posts

Social media health-related content is typically found in the form of unstructured, natural language text. Regarding the relatively large size of this type of data in social media, methods for automatic extraction and analysis of the data received considerable attention (Sarker et al., 2015). However, performance of the methods is affected by the inherent complexity of the posts. This complexity is the result of:

- 1) Colloquial phrase and sentence structure,

2) Deviation of lay person language from professional medical language in expressing the pharmaceutical effects of drugs, and

3) Ambiguity in presentation of a specific term; a given term may be used variously as a side-effect, a withdrawal symptom, or a drug indication.

Due to the first challenge, the conventional methods for named entity recognition for detecting boundaries of terms and phrases are subject to bias inherent in the data (Liu & Chen, 2013). Therefore, a dictionary-based approach to named entity recognition has been used as an alternative. Nevertheless, because of the second and third challenges, the systems face low recall, leading to high frequency of undetected desirable terms, such as side-effects and withdrawal symptoms.

The lexicon-based approach for name entity recognition in the area of pharmacovigilance currently dominates other methods of data 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) (Edwards et al., 2013), MedEffect (Canadian Adverse Reaction and Medical Device Problem Reporting database) ("Medeffect Canada. MedEffect Canada [Internet],"), SIDER (which has been developed based on resources published by public sources, mainly the FDA such as structured product labeling (SPL)) (Kuhn, Letunic, Jensen, & Bork, 2015), the Drug Bank Database (Wishart et al., 2006), and the European agency for the Evaluation of Medical Product (EMA) (Gardner, 1996). 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 used few

approaches that mostly focused on augmenting 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).

2.10.1 Automatic extractions of ADRs from consumer health posts

Leaman et al. (2010) constructed a lexicon of SIDER, MedEffect, and COSTART, which were augmented with CHV and a small set of colloquial expression of ADRs to identify adverse drug reaction in consumer drug reviews in the "Daily Strength" forum. In addition, the search algorithm was set on window-based search with size five to capture syntactic variation in ADR expressions composed of two or more parts, such as "gained huge weight" for "weight gain". To address the ambiguity and distinguish ADRs from drug indications and beneficial aspects of drugs, they used a rule-based approach built on the closest verb to the ADR expressions. The study had 78.3% precision and 69.9% recall, and F-measure 73.9%. (see **Table 1** for a summary of these studies.)

Benton, Ungar, et al. (2011) compiled 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. Since the authors could not filter reviews for the desirable drugs, they used co-occurrence techniques to determine the association between mentions of the drugs and ADRs. The authors did not provide any information about employing specific methods to address colloquial phrases or ambiguity in the health posts. The reported recall was 35.1%, precision 77%, and F-measure 73.9%.

Liu and Chen (2013) designed AZDDrugMinor, a general application for detecting ADRs in health post forums. MetaMap (A. R. Aronson, 2006), a tool that mainly developed for

identifying in unstructured texts medical concepts expressed by standard medical vocabularies, was configured to recognize all terms belonging to ‘Chemicals and Drugs’ and ‘Disorder’ semantic groups. Then the identified terms were compared against the list of ADRs in FAERS. To identify negated side-effects, NegEX (Chapman, Bridewell, Hanbury, Cooper, & Buchanan, 2001) was used. To distinguish ADRs from drugs indications, they used the drugs indications list specified in FAERS database. The application was tested by 200 sentences collected from a forum and showed 56.5% recall, 82% precision, and F-measure 66.9%.

Nikfarjam, Sarker, O’Connor, Ginn, and Gonzalez (2015) designed the ADRMine, a feature-based machine learning system, to identify adverse effects from health posts' sentences in Daily Strength and Twitter. First, to address the challenge of colloquial expression of ADRs in the posts, they trained a supervised sequence labeling CRF classifier on a set of corpora from Daily Strength (4720 reviews) and Tweepers (1340). In these two corpora, the drug reviews were annotated for the presence of ADRs, the span (boundary) of the ADRs, and the beneficial effects of drugs. To address the ambiguity (distinguishing ADRs from instances of other semantic types), they used a list of features, including a binary feature indicating presence/absence of ADRs that construed, based on the same ADRs lexicon developed by Leaman et al. (2010), part of speech of the token, context features, and negation. Moreover, to extend the semantic variability of the ADRs annotated in the corpus (Daily Strength and Tweets), they used the word embedding technique (word2vec) (Ganguly, Roy, Mitra, & Jones, 2015) trained on unseen drug review posts. The result for daily strength was 78% recall, 86% precision, and 82% F-measure, and for Twitter 76% recall, 68% precision, and F-

measure 72%. The authors attributed the errors to the colloquial expressions of ADRs, spelling errors, and informal sentence structure.

(Sarker & Gonzalez, 2015) designed a feature-based machine learning system to classify text segments from drug review posts into the ADR or non-ADR categories. First, they trained three supervised classification approaches, including Naïve Bayes (NB), Support Vector Machines (SVM) and Maximum Entropy (ME) trained on the two corpuses of daily strength and Twitter introduced by (Nikfarjam et al., 2015). To improve performance of the system, they created a binary feature using the same ADRs lexicon developed by Leaman et al. (2010). The binary feature indicating presence/absence of ADRs mentions. They also created a numeric feature, which was calculated by counting the number of ADR mentions divided by the number of words in the text segment. The authors also used other features for representing semantic properties, including n-grams, UMLS semantic types and CUIs, synonyms extracted from WordNet, and sentiword scores. Among the tree feature-based classifiers, SVM significantly outperformed NB and ME in both Daily Strength and Twitter corpora with F-Score 67% and 54% respectively. Authors attributed the relatively low F-scores to misspelling, short posts, ambiguous statements, and colloquial expression of ADRs, such as descriptive phrases of adverse effects, non-standard terms and high variability of semantic representation of specific ADRs in health posts.

Huynh, He, Willis, and Rüger (2016) trained and tested four different structures of Neural Network for classifying ADRs assertive text segments on a social media source (the same Twitter corpus introduced by (Nikfarjam et al., 2015)) and a non-social media source, the Adverse Drug Effects (ADEs) (Gurulingappa, Mateen-Rajput, & Toldo, 2012), which was constructed by sampling from MEDLINE case reports. Among Convolutional Neural

Network (CNN), Recurrent Convolutional Neural Network (RCNN), Convolutional Recurrent Neural Network (CRNN), and Convolutional Neural Network with Attention (CNNA), CNN and CRNN had 51% F-measure on Twitter data and 87% F-measure on the ADEs corpus. The low performance of Neural Network classifiers on the Twitter corpus compared with the non-social media corpus occurred due to high structural variability in text segments including ADRs and high syntactic and semantic variability of ADR expressions in social media posts.

(Nguyen et al., 2017) attempted to calculate the frequency of the top 10 ADRs of the top 10 psychiatric medications in LiveJournal, Reddit posts, Reddit comments, and Twitter. The data for top 10 ADRs was collected from the SIDER database. To improve performance of their system in capturing semantic and syntactic variability of the 10 ADRs (mentioned by SIDER), they employed the word embedding technique (word2vec framework). For example, for instances of “diarrhea”, the word2vec included the alternative spelling or misspelling in the final tally. The authors did not provide any information about the accuracy of their system in finding semantic and syntactic variations of the ADR terms, but they concluded that the frequency of ADRs calculated in social media by extending the list of ADRs has better agreement with the frequency of the ADRs in the SIDER database.

As the low recall of the ADRs extraction systems and low F-measure of ADRs assertive text segment classifiers show, informal sentence structure, low coverage of standard medical lexicons for colloquial expression of ADRs, and ambiguity in semantic types of medical terms are three main obstacles in enhancing performance of the systems. Even the use of advanced machine learning methods, such as deep learning and the word2vec framework did not upgrade performance of the systems significantly. Moreover, augmenting the standard lexicons with CHV did not improve results of the systems significantly, indicating that the CHV is not rich in

colloquial expression of ADRs. Therefore, there is clearly a need for an annotated corpus that not only clarifies the text segments of health posts for presence of specific information, such as drug effectiveness and ADRs and detecting desirable medical terms, but also filling the gap between patients' and clinicians' terminologies by mapping the colloquial expressions to standard medical terminologies.

Table 0-3 Specifications of Studies Used Lexicon-based Approach and Machine Learning Methods for Automatic Extraction of Health-Related Information from Consumer Health Posts.

Study	Purpose	Professional lexicon used	Methods for extending standard medical lexicon	Identifying semantic type of expressions	performance
Leaman, Wojtulewicz et al. (2010)	Identifying ADRs from Daily Strength	COSTART, SIDER, MedEffect,	CHV, a small set of colloquial language mapped to UMLS, Window based search	Rule-based approach that uses cures from nearby terms	78.3% precision, 69.9% recall, 73.9% F-measure
Benton, Ungar et al. (2011)	ADRs of hormonal breast cancer drugs from breast cancer forums	Dietary supplements, Pharmaceuticals terms mentioned in Cerner Multum's Drug Lexicon, sign and symptoms in Medicine website, FAERS database	CHV	---	77 precision, 35.1 % recall, 73.9% F-measure
Liu and Chen (2013)	A general application for detecting ADRs in health post forums	UMLS, MetaMap the findings compared with by FAERS database.	CHV	NexEX (for negation detection), FAERS Database to distinguish drug indication from ADRs	82% precision, 56.5% recall, 66.9% F-measure.
Nikfarjam, Sarker et al. (2015)	Identifying ADRs in Daily Strength and Twitter	The lexicon generated by Leaman, Wojtulewicz et al. (2010) used for constructing binary feature indicating presence/absence of ADRs	Training a supervised CRF classifier on corpuses from Daily Strength and Tweepers, Word2Vec to extend the semantic variability of the ADRs annotated in the corpus	Applying Features including art of speech of the token, context features, and Negation detection.	86% precision, 78% recall, 82% F-measure

Sarker and Gonzalez (2015)	Binary classification of ADR assertive text	lexicon developed by Leaman, Wojtulewicz et al. (2010) used for generating a binary and numeric features	Training Classifier Methods on two corpuses: twitter and Daily Strength	Generating a set of features, representing semantic properties of a text (e.g., n-grams, sentiment, polarity, topic, synonyms extracted from WordNet)	Twitter Corpus: 53% F-measure Daily Strength Corpus: 67% F-measure
Huynh, He et al. (2016)	Binary classification of ADRs assertive text	No lexicon	Four different structure of Neural Network based Classifiers were trained on a Twitter corpus	—	Convolutional Recurrent Neural Network: 51% F-measure
Nguyen, Larsen et al. (2017)	Frequency estimation of top 10 ADRs of the top 10 psychiatric medications. Sources: LiveJournal, Reddit posts, Reddit comments, and Twitter.	10 top side effects of psychiatric medications in the SIDER.	The word2vec framework for generating the alternative spelling or misspelling in the final tally	—	No computed recall/specificity/F-measure. Conclusion: The word2vec increase the estimation frequency to the point of agreement with SIDER.

2.10.2 Corpora Developed Using Consumer Health posts

Currently, there are two open-source corpora based on health-related social media: (1) a Twitter corpus that was built on 10,822 instances of randomly selected tweets (each instance of tweet is a maximum of 140 characters) for 71 drugs prescribed for chronic illness, and (2) the CADEC corpus, constructed based on drug review posts on the online message board “askapatient.com”. These corpora are summarized in **Table 2**.

The tweets in the Twitter corpus were double coded by two annotators for presence of ADRs, spans of ADR indications, and beneficial effects. For this data set, the Inter Annotator

Agreement (IAA) calculated using Cohen's Kappa was 71%. The authors normalize the identified medical terms by mapping layperson expressions to the UMLS (Ginn et al., 2014)

The CADEC 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), symptoms (275), and disease (283), and drug names (1,800). The pair-wise agreement between annotators was calculated using Metke-Jimenez, Karimi, and Paris (2014) method, which was 86.5% , when span and annotation settings were both relaxed and 60.4 % , when span and annotation settings were both strict. All the entities other than drug names were mapped to SNOMED CT. All the drugs were mapped to AMT (the Australian Medicines Terminology). In addition, all the ADRs also mapped to MedDRA (Karimi et al., 2015).

2.1.3 Corpora developed using biomedical literature.

There are also other corpora for identifying ADRs that developed based on biomedical literature, mainly from Medline case reports and abstracts. These are summarized in **Table 2**.

Gurulingappa, Klinger, Hofmann-Apitius, and Fluck (2010) developed a corpus of ADRs and diseases using randomly selected MEDLINE abstracts. It consists of 813 mentions of adverse effects and 1,428 mentions of disease. Further, Gurulingappa, Mateen-Rajput, et al. (2012) developed a corpus using 2,972 medical case reports randomly selected from MEDLINE. The corpus was annotated for mentions of drugs (5,063), adverse effects (5,776), dosage (231), as well as the relationships between drug-adverse effect (6,821) and drug-dosage (279). In addition, Gurulingappa, Rajput, et al. (2012) created Adverse Drug Effects (ADEs), a corpus of annotated sentences, indicating the presence/absence of ADRs, which were obtained from 2,094 MEDLINE medical case reports. The corpus contains annotations of 5,063 drugs, 5,776

conditions (e.g. diseases, signs, symptoms), and 6,821 relations between drugs and conditions representing clear adverse events.

Table 0-4 Specifications of Existing Corpus Relevant to the Corpus Specified in this Study

Study	Source	Type	Entity and Size	Normalized
Gurulingappa, Klinger et al. (2010)	Medicine abstract	Biomedical literature	813 ADRs 1,428 mentions of disease	...
Gurulingappa, Rajput et al. (2012)	2,972 medical case reports randomly selected from MEDLINE	Biomedical literature	Drugs (5,063), Adverse effects (5,776), dosage (231) Relationships between drug-adverse effect (6,821) and drug-dosage (279)
Gurulingappa, Mateen-Rajput et al. (2012)	2,094 MEDLINE medical case reports.	Biomedical literature	5,063 drugs, 5,776 conditions (e.g. diseases, signs, symptoms), 6,821 relations between drugs and conditions	...
Ginn, Pimpalkhute et al. (2014).	10,822 instances of randomly selected tweets	Social Media-Twitter	Presence of ADRs, spans of ADR indications, and beneficial effects	Mapped to UMLS concepts
(Karimi, Metke-Jimenez et al. 2015)	1,231 comments for two sets of drugs	Social Media (askapatient.com)	ADRs (6,318), symptoms (275), and disease (283), and drug names (1,800)	Identified entities mapped to SNOMED-CT, AMT, and MEDRA.
Our corpus	896 comments for two classes of psychiatric medications: SSRI and SNRI	Social Media (askapatient.com)	Sentence classification: presence of Side-effects, Withdrawal symptoms, Drug indicators, Effectiveness, and Ineffectiveness. Entity identification: side-effects, withdrawal symptoms, and drug indicators.	Identified entities mapped to both UMLS and SNOMED-CT

Although the corpora annotated for mentions of ADRs using biomedical literature and clinical notes in Electronic Health Records (EHRs) have important implications for automatic extractions of ADRs from these resources, they may not provide significant performance improvement for ADRs identification in consumer health posts in social media. As Sarker and Gonzalez (2015) showed, incorporation of ADEs corpus (Gurulingappa, Mateen-Rajput, et al., 2012) with corpus obtained from social media data does not provide significant improvement in the accuracy of ADRs of an assertive sentences-classifier system, because the corpus structure is

not compatible with the social media corpora. In addition, corpora developed with biomedical literature follow grammatical rules and are not rich in colloquial expression of medical entities. Therefore, they do not address the challenges of ADR medical entity extractions from consumer health posts in social media.

In response to these challenges, we developed a corpus in this study that addresses pharmacological effects of psychiatric medications including ADRs, drugs indications, and drug effectiveness. We followed a systematic approach to develop this corpus consisted of two main components: entity identification, and entity normalization. The identified entities were mapped to the equivalent medical concepts in both the UMLS and SNOMED CT. ADRs and WDs were further classified as physiological, psychological, cognitive, and functional problems (e.g., limitation in daily functioning, social activities or inter-personal relationships) that did not receive any attention in previous studies. The methodology for developing this corpus is explained in the methodology section of this study.

Chapter 3: Methodology

This study involves a mixed-method approach for providing structured data from unstructured drug reviews for testing hypotheses concerning attitudes and adverse drug reactions associated with antidepressants. Methodology for developing the structured data and testing the hypotheses consisted of seven major phases:

- Phase 1: identification of data source and drug source
 - Drug sources of this study are Sertraline (brand name: Zoloft), Escitalopram (brand name: Lexapro), venlafaxine (brand name: Effexor XR), duloxetine (brand name: Cymbalta).
 - Data source of this study is a healthcare forum called “askapatient.com”.
- Phase 2: Data collection

This phase includes the following steps:

 - Developing an Application Program Interface (API) for data collection
 - Calculating the sample size: 892 reviews were collected for four drugs specified in this study.
- Phase 3: Developing an Analytical Framework for generating structured data from unstructured text

This phase includes the following steps:

 - Developing the initial analytical Framework using deductive approach (reviewing the literature)
 - Annotating the drug reviews using the initial analytical framework
 - Generating themes using inductive approach (open coding)
 - Developing the final analytical framework by refining themes obtained from inductive and deductive approaches.
- Phase 4: Applying the final analytical framework to the drug reviews

This phase includes the following steps:

 - Data preprocessing: This step includes using regular expression codes to remove personal information and noisy pattern from sentences structure.
 - Splitting drug reviews to sentences.
 - Annotating sentence using analytical framework: This step includes developing guideline for annotators, developing annotation environment, calculating inter-annotator agreement and resolving disagreement between annotators.
- Phase 5: Entity identification

This phase includes the following steps:

 - Developing guidelines for identifying Adverse Drug Reactions (ADRs), Withdrawal Symptoms (WDs), and Drug Indications (DIs)
 - Annotation process
 - Calculating inter-annotator agreement

- Phase 6: Terminology association (entities normalization)
 - Developing guidelines for mapping the identified entities (ADRs, WDS, and DIs) to UMLS and SNOMED-CT. The guidelines include requirements for selective proper/preferred concepts, procedure of mapping, and instruction for selecting proper standard concept for colloquial expressions of ADRs/WDs/DIs.
 - Reviewing the mapping for consistency
- Phase 7: Usability of the data set

The structured dataset provided through phase 1 to phase 6 used for generating and testing hypotheses related to association between variables and ADRs identified in this study. This phase includes the following steps:

 - Summarizing the dataset
 - Using imputation methods for handling missing values
 - Testing hypotheses
 - Developing a predictive model

Figure 1-3 shows the summary of the research methodology and the outcome of each phase, which is specified in the result section.

Summary of the Research Methodology

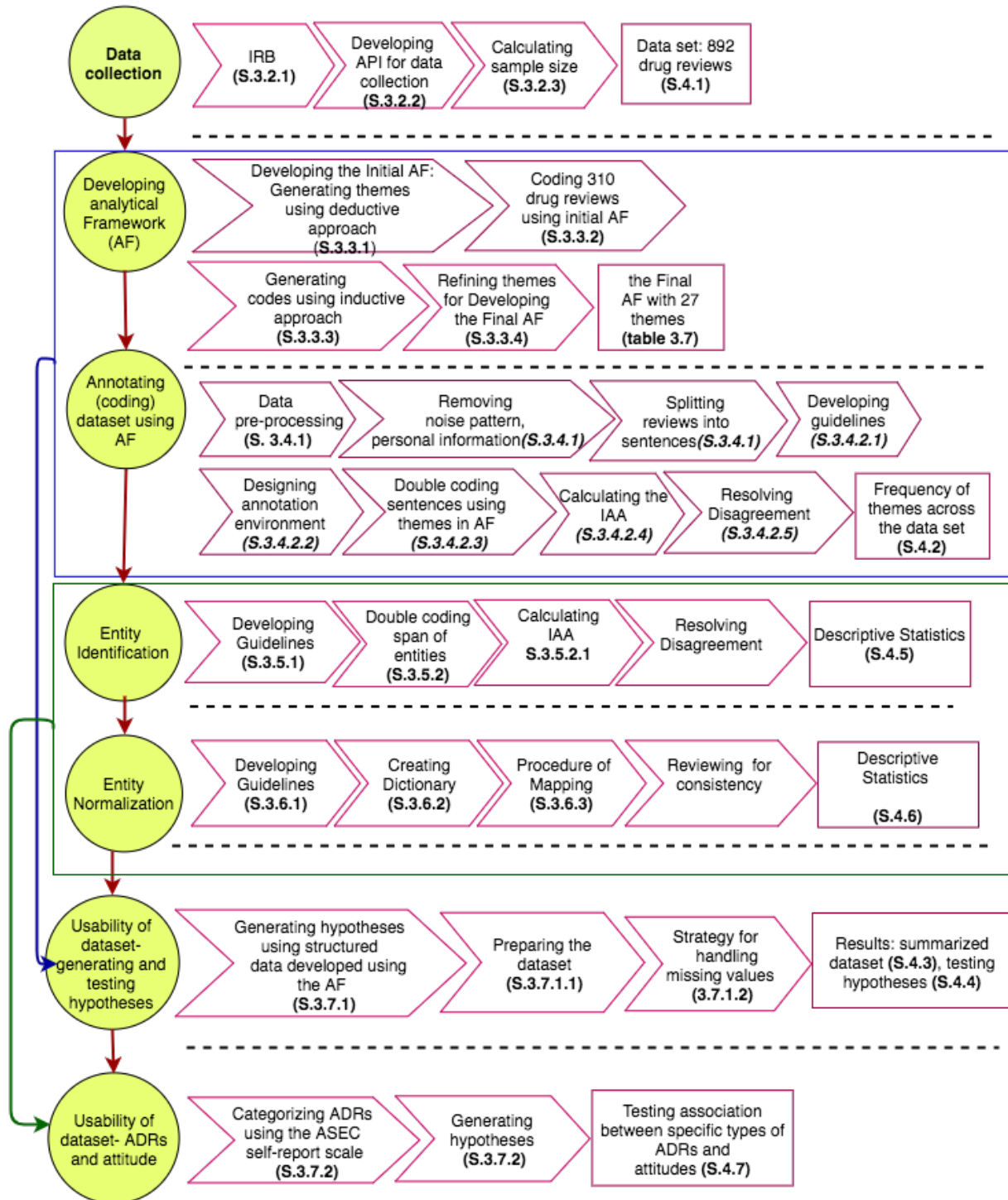


Figure 0-1 Summary of the research methodology

Phase 1

3.1 Data Source Information

3.1.1 Data sources

We examined data from a healthcare forum "askapatient.com" that compiles uncensored user comments on the effects of taking different sorts of medication from people with a range of diagnoses. In this forum, 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 supposed to report drugs' ADRs in the side-effects field and the detail of their experience in the comment field. However, patients reported various aspects of their experiences, such as drugs' effectiveness or perceived distress received from ADRs, in both fields. **Table 3-1** shows an example of posts for Cymbalta in "askapatient.com".

Table 0-1 An Example of a Post for Cymbalta in "askapatient.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

3.1.2 Drug Source

We used drug review posts in "askapatient.com" to collect information for four psychiatric medications, including 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, and according to a dataset from Symphony Health Solutions*, they had the highest prescription rate in 2012. Although the medications have shown substantial evidence of effectiveness in promoting mood elevating properties, they are associated with a significant number of ADRs causing debilitating impacts on patients' daily activities and social participation.

Phase 2

3.2 Data collection

3.2.1 Website policy

All the data in the healthcare forum “askapatient.com” are anonymous and publicly available, however some respondents may disclose their email address or their phone number in the field of comment. We sought IRB approval through University of Wisconsin-Milwaukee (UWM) and we received the following response:

“Based on the activities described in the Determination of UWM IRB Submission form, your study does NOT involve research with human subjects and IRB review and approval is NOT required for your project. By analyzing publicly available posts, you will not be obtaining data through intervention or interaction with the individual and you will not be obtaining private information.”

We also attempted to contact the website administrators, but we did not receive any replies despite repeated requests. Regarding the anonymous data recording process, website privacy policy, and response of UWM to our IRB request, we decided it is ethically acceptable to conduct a passive data collection from the healthcare forum. Thus we did not seek IRB approval for data collection phase. To further protect respondents' identities, we removed all the emails and phone numbers from dataset.

3.2.2 Designing application program interface (API)

The website does not have API (application program interface) access. We developed a web-crawler system to collect all the respondents comments available in this website for the four antidepressant drugs. Because data in this website are classified based on the type of drugs, it is possible to easily filter other drugs that are not purpose of this study. For each respondent's comment, type of disorder, side effects, comments, dosage, duration of the drug intake, age, and gender were collected, if they were available. The information was collected in August 2016.

3.2.3 Calculating the sample size

Since the focus of this study is on patients' attitudes towards antidepressants, we first filtered review posts in which patients reported depression with or without comorbid with other mental/physical illness as the reason for drug prescription. **Table 3.2** provides information about the total number of reviews available in the forum for the four drugs focused on in this study and the number of reviews after filtering the dataset for depression.

In order to select a subset of data that sufficiently represents the whole dataset in the healthcare forum, we used the following sample size formula (Charan & Biswas, 2013):

$$Sample\ Size = \frac{\frac{z^2 \times p(1-p)}{e^2}}{1 + (\frac{z^2 \times p(1-p)}{e^2 N})}$$

Where z is the confidence interval, e is the margin of error, p is the standard deviation, and N is the population size. For this study, z is 90%, e is 5%, p is 0.5 to ensure the sample is large enough, and N is the number of post reviews for each drug in the healthcare forum. The calculated sample for Zoloft, Lexapro, Cymbalta, and Effexor XR were 213, 219, 231, 229 respectively. We used Python Random Module (Python, 2017) to select the calculated sample from the healthcare forum randomly.

Table 0-2 Total Number of Review Posts and the Calculated Sample Size

Drugs	Number of reviews available in the forum	Number of reviews include depression as the reason for drug prescription	Sample Size
Zoloft	1528	980	213
Lexapro	1913	1145	219
Cymbalta	2481	1573	231
Effexor XR	2129	1447	229

Phase 3

3.3 Developing the Analytical Framework for Generating Structured Data

Management and summarization of unstructured data in qualitative research is a critical aspect in data analysis. The Framework method is a flexible tool that provides researchers with the approaches of developing themes to generate highly structured data from qualitative data.

The Framework method was developed at the qualitative research unit at the National Center for Social Research in the United Kingdom for large-scale policy research. However, it is becoming an increasingly popular approach in medical and health research data (Gale, Heath, Cameron, Rashid, & Redwood, 2013). Like other qualitative approaches, the Framework method helps to identify similarity and discrepancy in qualitative data before determining the relationship between different parts of the data. This identification involves developing a descriptive summary of the data clustered around themes.

Depending on the research question, the Framework method uses an inductive, a deductive, or a combination of inductive and deductive approaches on the inductive-deductive continuum. The difference between these three approaches is the way in which the themes for developing structured data are generated. In the deductive approach, themes are developed based on literature, previous theories, or specifics of a research question. In the inductive approach, generates themes from data by open coding and refining themes. In a combined approach, a

study starts with predefined themes to explore, but also it leaves space to discover the unexpected aspect of respondents experiences with the subject of the study.

To provide structured data from the review posts, we adopted the Framework method by focusing on a combined approach (inductive-deductive). A deductive approach was used that involved reviewing the literature the set of important variables affecting attitudes towards antidepressants, which construct the initial that provided a framework for data analysis. Conducting the inductive approach generated the list of themes for data analysis that were not discussed by the literature. The following steps were followed for constructing the final set of codes for qualitative analysis of reviews.

3.3.1 Developing the initial analytical framework using deductive approach

We conducted a review of literature to identify important clinical and personal factors affecting patients attitude to antidepressants. **Table 3.3** provides the identified factors (codes) from literature with their description. The factors were categorized into five categories, including pharmaceutical treatment factors, healthcare system factors, psycho-social factors, patient related factors, and disorder related factors. The identified factors were used as the initial codes for constructing the initial Analytical Framework. Each identified code may include sub-codes indicating the spectrum of the levels of patient experience for that variable.

3.3.2 Coding (annotating) the drug reviews using the initial analytical framework

The purpose of this step is to understand to what extent the pre-defined themes could be applied to the comments. In addition, it helped us to identify passages of comments that could not be mapped to the predefined codes.

For the purpose of initial analysis, 310 reviews from the combined drug samples were randomly selected. Because, reviewers entered their experiences with antidepressants in two

fields, including side-effects and comments, we first combined the two fields and generated one segment of text for each review that was identified by *drug id*.

Table 0-3 Description of Predefined Themes Developed Deductively from Literature Review

Categories	Factors (predefined codes)	Description
Pharmacological treatment factors	Perceived effectiveness	Patient's subjective assessment of antidepressant helpfulness in reduction of depression symptoms, enhancing emotional and cognitive functionalities, and in overall, enhancing life quality.
	Perceived necessity	Patient subjective assessment of antidepressant necessity in improving and maintaining current and future health conditions. For example, patient may find an antidepressant vital in reducing the risk of relapse.
	Perceived concern	Patient subjective assessment of antidepressant harmful aspects in long-term. Patient may view antidepressant as an agent, which is addictive, take control over feelings and thought, and altering personality in long-term.
	Side-effects	Any adverse reactions that patient report as side effects of antidepressants intake. Antidepressants' adverse reaction may include physiological side-effects, emotional syndromes, cognitive impairment, and limitations on daily functioning, and in overall reducing quality of life
	Perceived distress from side-effects	Patient's perceived distress from antidepressants side effects depends on patient's self-attention to internal bodily sensation that may have an influence on patient tolerability of side effects. Patient may show the distress by using adjective showing severity of a side effect, negative impact on work or daily activities, or visiting emergency department. A severe side-effect including self-harm and suicidal ideation or attempt reflects distress received from antidepressants.
Healthcare system factors	Patient-provider relationship	Patient expresses their satisfaction or dissatisfaction with healthcare providers from various aspects, such as perceived support from providers or perception of healthcare providers knowledge about illness and treatment.
	Healthcare settings	Patients may demonstrate a higher level of trust and satisfaction towards diagnosis and treatment offered in a psychiatric setting compared with a primary care setting.
	Affordability	Patient may complain about the high cost of antidepressants and lack of insurance coverage.
Psycho-social factors	Stigma and cultural related factors	Patient may express their concern about stigma and cultural-related factors that may make prolonged pharmacological treatment notoriously difficult for patients. For example, the public may view antidepressants intake as a sign of weakness or incapacity to deal with daily emotional distress that may influence on patient acceptance of antidepressants.
	Partner support	Patients may express their perceived support from partners (family and friends) about having depression as a proper diagnosis and having an antidepressant as a proper treatment.
Patient-related factors	General concern and necessity	Patient may express their general view toward medications. For example, they may view all the medications harmful and believe natural remedies and changing life style will have a better healthcare outcome than pharmacological treatment.
	Knowledge about medication	Patient may complain about lack of knowledge of side effects or complexity of treatment, proper time of discontinuation, and withdrawal adverse effects.
	Education level	Patient may disclose their level of education in the comments or in the section of demographic information.
Disorder related factors	Patient insight about source of depression	Patient may express their insight towards the source of depression, and severity of symptoms. For example, patient may attribute source of depression to psychological problem rather than biological factors.
	Severity of depression /symptoms	Patient with major depression disorder may have a higher tolerability of adverse effects. In addition, patient's perceived effectiveness of antidepressant may be higher in the patients.
	Type of depression	Type of depression may be an important factor in shaping patient's attitude toward antidepressants.

Selected reviews with *drug ids* were arranged in rows and predefined codes were arranged in columns, developing the initial framework of analysis. Microsoft Excel was used for developing the initial framework of analysis.

Figure 3-1 provides an excerpt of the initial framework of analysis that was generated using pre-defined codes. As illustrated in the figure, we highlighted each meaningful passage of text and attached it to a cell corresponding to the appropriate code in the initial framework. For passages of text that did not fit into the existing themes in the framework, we determined and documented the theme of the passage and recorded it in a new column named “**Not-Matched**”.

We also created a column named “Not-applicable” that contains passages without any meaningful information about respondents’ experiences with antidepressants. For example:

- *“Feel free to email me, I have been through it all with Effexor.”*
- *“Thank God I was pregnant or I may NEVER have forced myself through to the other side!!!”*

Patients in the review posts did not provide any information about the pre-defined themes of “type of healthcare setting” where they received treatment for depression, “stigma and cultural related factors”, or “educational level”. In addition, the initial analysis showed that there is little information available in the reviews for some of the predefined themes, including “affordability”, “general concern and necessity”, “partner support”, and “patient insight about depression”. **Table 3.4** provides the identified codes and related sub-codes with examples from patients’ reviews for each code.

Table 0-4 Predefined Themes Developed Inductively with Examples from Drug Reviews

Predefined codes	Sub-codes	Examples from patients' reviews
Perceived effectiveness	Positive	"Anyway, my life is on track, I have nothing to be depressed or sad about."
	Negative	But I still havnt noticed any change in my energy or anxiety.
Perceived Necessity	Positive	"All in all, I love it. I have not have a depressed moment since I've been on it, approx. 8 month. I do not think I cannot live without it."
	Negative	"There's no doubt that Effexor XR saved my life, however long-term use is not the best."
Perceived concern	Concern about side-effects	"Pure poison!" "I was so scared to take it that I delayed it because I read all the comments on this site."
	Concern about withdrawal symptoms	"I am worried about trying to cut down gradually as I read so much scary stuff about going off it." "The withdrawal is hell. You would think you were going through heroin withdrawals."
ADRs		"Typical with effexor XR- Dizzy, jaw tight, teeth grinding."
Perceived Distress from side-effects	Tolerable (low)	"I had tolerable, minimal side effects." "I didn't really experience any side effects..." "The side effects are far less hassle than being in my dark depression."
	Intolerable (high)	"The side effects are hell." "No wonder I suffered so many side effects."
Physician patient - interaction	Positive	"My pharmacist convinced me that it was safe for me to use along with my topamax, and encouraged me to try it. I'm really glad I did."
	Negative	"I am suffering and should of not trusted the NP who prescribed the pills without first checking it out."
Healthcare setting *		"Respondents did not explicitly mention in which healthcare setting they received treatment for depression."
Affordability *		"If the depression doesn't eventually cripple you, the costs will. It is too expensive if you don't have insurance!"
Stigma and cultural related Factors		Non of the respondents did not explicitly complain about stigma or did not present any information indicating cultural factors affecting on patients attitudes antidepressants.
Partner support*	Positive	No example
	Negative	"My husband has no compassion to my withdrawal symptoms."
General concern and necessity*		"Do NOT believe the MDs and pharm corp. hype and lies! FLEE!"
Demographic information		"Basic demographic information including age and gender was reported by respondents."
Knowledge about medication		"No one ever told me even as I was seeing my psych and psychologists weekly, and my physician, that what I was experiencing could be from the drug."
Education level*		Patient did not disclose their education level in their comments. The website also did not provided a field to request responded educational level.
Patient insight about depression *	Positive insight	Mental illness (i.e. lack of serotonin) is just like any other defect the human body can have.
	Negative insight	No example
Severity of depression/symptom		Some of the respondents in the field of "reason of prescription" or their comments mentioned severity of depression or their symptoms, such as: "The depression is so terrible that I am very keen to give Lexapro a fair try."
Type of depression		Some of the respondents mentioned type of their depression in the field of "reason of prescription" or in the comments.
Duration of depression		Most of the respondents in the filed of "duration/dosage" reported duration of antidepressant's usage.

3.3.3 Generating codes (themes) using inductive approach (open coding)

Our regular team meetings enable us to discuss the proposed ideas, discover patterns, and an impression for alternative viewpoints, thus ensuring that one researcher's particular perspective did not dominate and that agreement upon codes (themes) is a team conclusion. We revised the initial framework by defining new themes and refining the existing codes to incorporate unfitted passages of comments. We revised the initial framework by defining new themes to incorporate unfitted passages of reviews.

Patients in the review posts stated their experiences with withdrawal symptoms (WD) and perceived distress from WDs. They also included experiences with withdrawal and discontinuation. Moreover, patients might express their overall attitude and pre-treatment concern with medications. They also may have some suggestion for readers about the quality of the drugs.

Table 3.5 demonstrates new themes developed through the process of the inductive approach (open coding).

3.3.4 Developing the final analytical framework by refining the themes

The process of refining, applying, and re-refining themes was repeated until no new themes (codes) was generated. To develop the final analytical framework, we took the following steps:

- 1) Each theme in the predefined framework that did not fit with information provided in reviews was excluded from the final framework. For example, we removed “healthcare setting,” “stigma and cultural related factors”, “educational level”, and “dosage of medication”.

Table 0-5 New Themes Generated Through the Process of Inductive Approach (Open Coding)

New Themes	Sub-Themes	Examples
Withdrawal symptom		<p>"I weaned slowly from 150mg to 75 to 37.5 and off. I feel nauseous alot and my depression and social anxiety has returned almost 100%."</p> <p>"Do not wean off effexor too soon as i had one very bad day and experienced mild halloucinations."</p>
Perceived distress from Withdrawal Symptoms	Low (tolerable)	"Could miss a dose or two and no problem."
	High (intolerable)	"Trying to come off of the medication is very difficult."
Overall attitude	Positive	"I think that this drug is great, but only for certain people."
	Negative	<p>"This drug is just a poison."</p> <p>"This drug ruined my life."</p>
Drug indication		<p>"My depression mostly manifested in an inability to start new projects, rather than any feelings of sadness."</p> <p>"I used to be so depressed and anxious, that I didn't want to leave the house. I have a long history of depression and anxiety."</p>
Recommendation to others	Positive	"I would definitely recommend Effexor XR."
	Negative	"Do not take this medication!!!!"
Experience of Withdrawal	Discontinuation	"I stopped this drug after two days."
	Weaning off	"Cutting the dose in half, When I went down to taking none, experience withdrawal effects."
	Missing dosages	<p>"Can not miss a single dose or I feel awful."</p> <p>"I experience when missing a dose."</p>
	Switching	<p>"I have insisted on stopping the effexor, and now the doctor is pushing pristiq (the "new and improved" effexor)."</p> <p>Switched to Luvox which I am finding much more beneficial."</p>
	Decision about discontinuation	"I think I m going to quit."
Dosage/duration		"I started out on Effexor XR 75 mg, and was slowly raised from there to Effexor XR 300 mg. Then I was changed to the generic Venlafaxine XR 300 mg."
Experience with Other medications		<p>"I was given zoloft from my pyschiatrist and started at 50mg per day and was steadily increased to 100mg. the doctors increased my dose to 150mg."</p> <p>"Started taking it after Prozac and (can't remember the name) made my stomach hurt so bad."</p>
Problem with financial support		" I did a rapid decline on my cymbalta because I lost my insurance."
Problem with social support		" Thankfully I have a wonderful husband who helped me past the side-effects."
Not applicable		<p>"Hope the comments help and good luck."</p> <p>"I honestly can't tell a difference when I am on or off of Effexor XR."</p>

- 2) If a predefined or a new theme fit in less than 5% of information provided in reviews, we excluded them from or merged them with other themes in the final framework. Similarly,

“affordability”, “partner support” and “patient insight about depression” were excluded and “general concern and necessity” was merged with “**overall attitude**”.

- 3) All codes that were conceptually related and subjectively difficult to distinguish between their applicability to a passage of text were combined in the final framework.

Accordingly, “perceived necessity” and “perceived effectiveness” are conceptually related and subjectively difficult to distinguish in application to passages of text, therefore they were combined in the final framework as “perceived effectiveness”. For example, the following passage can be labeled as both “perceived effectiveness” and “perceived necessity”.

“All in all, I love it. I have not have a depressed moment since I’ve been on it, approx. 8 month”.

- 4) Although ADRs and withdrawal symptoms are conceptually related and both reflect adverse drug reaction, we distinguish between them as two separate codes in the final framework. A patient may receive minimal distress from an antidepressant’s side effects, but perceive severe withdrawal symptoms during the process of discontinuation, which may negatively affect the patient’s attitude. Distinguishing ADRs from withdrawal symptoms is important from the perspective of clinical trials and interventions designed specifically to help patients manage the process of a drug’s discontinuation.
- 5) Perceived concern for the prescribed drug includes passages of text, in which reviewers expressed their perceived distress from ADRs, WDs, or the overall distress they experienced with the drug. Thus, we defined two new codes “perceived distress from ADRs” and “perceived distress from WDs” to express perceived express towards ADRs and WDs. For the passage of text in which a patient expressed overall perceived distress

from a drug without distinguishing it for WDs and ADRs, we assign it to the code “overall attitude”.

- 6) Some patients directly mentioned how the information obtained from online sources including online messaging boards influenced their attitude toward antidepressants and in turn their behavioral reaction. This information can provide important evidence about the impact of an online sources on shaping attitudes toward medications. However, because of the paucity of the information, we did not define any new theme covering this type of information in the final framework. The following provides examples of such sentences:

- *“I don’t want all those withdrawals I m reading about”.*
- *“I was so scared to take it that I delayed it because I read all the comments on this site”.*

Table 3.6 presents codes and sub-codes in the final framework, each with a brief explanatory description and examples from reviews.

Table 0-6 Themes (items) in the Fianl Analytical Framework with Description and Example from Patients Comments

Code	Sub-codes	Description	Example
Adverse Drug Reaction	Presence	If patient reported explicitly he experienced side-effects with or without listing the ADRs in the sentence/comments	"I couldn't take Effexor XR. It gave me horrible nightmares and I kept waking up."
	Absence	If patient reported "they did not experience any ADRs"	"I did not have any side-effect."
Perceived distress from ADRs (ADR-PD)	High	<ul style="list-style-type: none"> - Explicit mentions: If patient explicitly mentioned that they suffered from ADRs, - Functional problems: If patient reported functional problems associated with ADR, such as limitation in daily functioning, social activities, and work performance - Qualifiers indicating severity: If patient used any qualifiers indicating severity of the symptoms. Such as severe, debilitating, intolerable. - Severe ADRs: If patient reported severe ADRs having negative impact on patient quality of life, such as suicidal ideation/attempt, self-harm, bed-ridden. Patient hospitalization or emergency visit also shows the presences of ADRs causing high-perceived distress. 	<p>"The side effects are intolerable."</p> <p>"Have been able to work (software developer) if attempting this drug during work week."</p> <p>"Severe nausea and dizziness."</p> <p>"That drug caused nausea and increased suicidal thoughts."</p>
	Low	<p>Explicit mentions: If patient explicitly mentioned that the ADRs were tolerable.</p> <p>Qualifiers indicating mildness: using Qualifiers indicating mildness of ADRs, such as slightly, mild.</p> <p>Qualifiers are indicating non-persistency: Using qualifiers indicating non-persistency of ADS, given that ADRs are NOT associated with qualifiers indicating severity of ADRs.</p> <p>No experience of ADRs: If patient explicitly mentioned they did not experienced any ADRs.</p>	<p>"Any side effects were, for me, tolerable compared to the benefits."</p> <p>"Mild headache"</p> <p>"Headache for two days, but severe headache for two days indicates high-perceived distress."</p> <p>"The withdrawal made me very dizzy."</p>
Withdrawal symptoms (WDs)	Presence	If patient complained about occurring new symptoms occurred in the process of dosage reduction, discontinuation, or missing dosages (unintentional withdrawal) of the medication, with or without listing the WDs symptoms	<p>"I weaned slowly from 150mg to 75 to 37.5 and off. I feel nauseous alot and my depression and social anxiety has returned almost 100%."</p> <p>"Do not wean off effexor too soon as i had one very bad day."</p>
	Absence	If patient reported that they did not	
WD-perceived distress (WD-PD)	High	If patient mentioned they a) suffered from withdrawal symptoms, b) and/or they reported functional problems associated with the WDs, c) and/or they used qualifiers indicating severity, d) and/or they mentioned severe WDs, the WD is high.	<p>"The withdrawal symptoms are horrible."</p> <p>"I was in bed for about one week."</p> <p>"I missed a dose yesterday, and now I'm nauseous."</p> <p>" I can not function. Feel I am poisoned."</p> <p>"The withdrawal side effects especially brain flashes or brain zap was VERY PAINFUL."</p>
	Low	a) If patients explicitly mentioned that withdrawal symptoms were tolerable, b) used indicators showing low perceived distress, c) used qualifiers showing tolerability of the symptoms, d) using qualifiers showing non-persistency of the symptoms, given that the	<p>"Withdrawal was fine."</p> <p>"When I stopped the drug, I had mild dizziness."</p> <p>"I experienced headache for few days after reducing the dosage."</p> <p>"I had no experience of withdrawal"</p>

		symptom was tolerable, e) explicit mention of no experience of withdrawal symptoms.	symptoms.”
Effectiveness (EF)		A drug is effective, if patients reported that his health condition has been improved or his symptoms were treated after drug consumption.	“For the first few weeks it helped me feel better.”
Ineffectiveness (INF)		A drug is ineffective; if patient reported that his/her health status did not improve, became worse, or still has the same symptoms.	“It did not help me at all.”
Patient – physician interaction (PPI)	Positive (P)	If patient’s express of communication with providers indicated implicitly/explicitly that patient has high level of trust in providers and she/he was satisfied with communication with providers.	“Success with these meds truly depends on staying in touch with your physician.” (Implying trust in physicians) “My doctor and I decided to stop taking it.” (implying providers involved patient in the process of decision making), which can lead to patient satisfaction.
	Negative (N)	- If patients express of communication with providers indicated implicitly/explicitly that patient were not satisfied with healthcare providers. - Patient may complain about provider’s failure in providing sufficient information or non-effective communication, such as provider’s failure to involve patient in the process of decision-making or treatment plan.	“Dr. s do not understand the crazy side effects of starting this class of drugs.” (Implying patient’s complaint about physician’s lack of knowledge) The doctor still claims that 30 mg is not even considered a therapeutic dose, but I know what works. (physician failure to include patient treatment preference in the process of decision making).
Lack of knowledge (~KN)		Some patients may complain that they did not receive sufficient information about ADRs or WDs symptom associated with the drug and and the mechanism of management of the ADRs/WDs. Patient may indirectly complain about lack of knowledge by asking questions in the forum or mentioning that they did research on the web to gain more information.	“Cannot get straight answer from anyone regarding how long these withdrawal symptoms will last.” “No one informed me of the withdrawal nightmare.” “TODAY I'd like to know if my recent muscular twitches are related to this medicine.” “I wish I had been smart enough to do research on Effexor BEFORE I went on it.”
Experience of WD	Unintentional WD (DXD-F)	If patient explicitly mentioned that they forgot to take medication (missing dosages) or run out of medication.	“When I miss a day I feel very spaced out, thick, groggy, sad.” “If I don’t take the medicine at the same time every day or forget a day, I will experience withdraw and its horrible!”
	Intentional WD- Stopping (DXD-S)	If patient explicitly mentioned that they stopped the medication.	“I had to stop taking it.” “I had low blood pressure for about 2 weeks after I stopped taking it.”
	Intentional WD- weaning off (DXD-W)	If patients explicitly mentioned that they are weaning off (tapering off) the medication.	“I have been tapering from 60 mg per day.” “I'm stopping slowly.”
	Intentional decision for WD (DXD-Dec)	If patient explicitly mentioned that they decided to stop or wean off the medication.	“However after dealing with this acne I'm going to try another med.”
General attitude	Positive (P)	If patient’s general attitude towards the drug is positive.	“ I really like this drug, it changed my life.”

toward medication (ATT)	Negative (N)	If patient's general attitude towards the drug is negative (ATT-N).	"I hate the drug, it completely ruined my life."
Patient suggestion to readers about drug (sug)	Positive (P)	If patient's suggestion to a reader (other patients) is positive.	"Zoloft is the best option for depression." "Ask your doctor to give you this drug."
	Negative (N)	If patient's suggestion to a reader (other patients) is negative.	"Please be aware of the drug and its consequences."
Dosage/duration		If patients talked about dosage or duration of the drug, it will be labeled as positive.	"My doctor prescribed 25 ml for me."
Lack of social Support (~SS)		If patient complained that he/she did not receive any social support for managing ADRs or process of discontinuation	"My husband did not understand me."
Lack of Financial Support (FS) (~FS)		If patients reported that because of lack of financial support, they had to stop the medication permanently or temporarily.	"I had to stop the medication because it was expensive."
Other drugs		"If patient discussed other drugs in the review."	"I used triazolam for Insomnia."
Other patients		If a reviewer explained other patient's experience with medication.	"My son was for three years on this drug, but he did not feel better."
Drug Indication (DI)		Any sign or symptom that patient clearly mention that it was the reason for drug indication and/or it was treated by the medication	"my anxiety disappeared because of Zoloft."

Phase 4

3.4 Applying the Final Analytical Framework to the Drug Reviews

The final analytical framework has been applied to all of the drug reviews in the sample. We only conducted data analysis on free text of each review posts that were available in both the side-effects field and comment field for each review (see table 3.1 for the structure of a review). The majority of drug review posts are composed of multiple sentences that each covers various aspects of patients' experiences with drugs, which were explained in **Table 3.6**. Therefore, we used sentences as the basic unit of analysis. In addition, data analysis at the level of sentences ensured that we did not miss any important passage that can be fit to the codes in the analytical framework.

The reviews are written in colloquial English language, in which patients mostly did not follow proper grammatical and punctuation rules. Duplicate punctuation, applying wrong punctuation as a sign of sentence completion, phrase construction irregularities, and

abbreviations are common in the review posts. The reviews may also include personal information, such as phone numbers, emails, or URLs. Therefore, before applying the framework to the data, we pre-processed data to remove the noisy patterns and personal information from the data set.

3.4.1 Data Pre-processing

Processing of the data was composed of text-cleansing and sentence boundary detection. In the first step, we formulated regular expression rules to remove: 1) the personal information including emails, phone numbers, and URLs from the reviews, and 2) the noisy patterns in the reviews' structure. **Table 3.7** shows some noisy patterns in the reviews and the regular expression codes for handling those issues. We did not remove emoticons from sentences, because they may represent perceived distress from a medication's adverse effects or satisfaction with a medication. For example, in the following sentence, the emoticon represents antidepressants effectiveness. "I quit this drug and started on LAMICTAL which saved my life I now feel content and complete. :-)".

To split the reviews into sentences, after removing the noisy patterns, we used the open-source Natural Language Toolkit (NLTK) (Bird, Steven 2006). We also used NLTK for sentence tokenization. Statistics on the review posts, the posts' sentences, and tokens are presented in result section.

3.4.2 Annotating sentences using analytical framework (sentence labeling)

3.4.2.1 Developing guidelines. To maintain consistency and uniformity of sentence labeling across the whole sample of reviews, we developed guidelines using codes introduced in the final analytical framework, their definitions, and examples from the drug reviews (**Table 3.6**).

Table 0-7 Examples of Regular Expression codes for Reducing Noise Patterns in Review Posts

Noisy patterns	Example	Regular expression code
Inserting space after full stop punctuations and start of a new sentence	I have suffered many side effects from the prednisone used to treat the symptoms.I am just now starting to feel better.	comments = re.sub (r'([a-zA-Z])([.?!])([a-zA-Z])',r'\1\2 \3', comments)
Eliminating Emails	com.uk@hotmail.comsevere	
Replacing dot before parenthesis with space	But about a year later I tried others and could tolerate them much better. (Such as Lexapro and Prozac.). I hate it.	Comment = re.sub(r'\.(\s*?([^\s]*)\s*\.)', r'\1', Comment)
Removing repeated punctuations from end of a sentence.	Think menopause is a huge factor in this....i have only been on generic zoloft for about a week and a half	comment = re.sub (r'([.,,:;!/?-])\1+', r'\1',comment)
Replacing “.” with “,” when list of side effects are separated from each other with dot.	Night terrors. Vivid dreams. Nightmares. auditory hallucinations everyday!!!! never take this drug!!!	comment = re.sub(r'\. (\s*(?!(:i she he He She I we We They they it It You you)\b)\w+(?:\s+\w+)?\s*)(?=[^\w\s])\$)', r',\1', comment , flags=re.I)
Replacing hyphen with dot when hyphen indicating full stop of a sentence	Severe teeth clenching/TMJ which caused excruciating headaches - clenching came on right after taking the dose and would gradually wear off	Comment = re.sub (r'(\s+)(-)(\s+)', r'\1.\3 ', comment)
Replacing comma with dot, when comma indicating full stop of a sentence	Effexor People are still having problems with this medication even after being off of it for 8 months or more, Please don't take this medication.	Comment = re.sub (r'(\s+)(,)(\s+)', r'\1.\3 ', comment)
Replacing dot after numbers when numbers indicating order of items in a list.	After two to three hours vomit 2. Stiff muscles (Lack of oxygen?) 3. Spacey, out of sync with time (as in Where did that tree come from? What am I doing?) 4. Weird, (as in How did that tree get there?). 5. Weak muscles	comment = re.sub(r'(\d)\.(\s*[a-zA-Z])',r'\1\2 ',comment)
Inserting space between parenthesis and the followed words.	""Was deadly sick for 2 wks and he kept me on it regardless, upped the dosage three times over a year because it wasn't working(duh, it was my thyroid!)That dr. is history, btw! This time, I was very leary of taking anything.""	comment = re.sub(r'(\s+)(\w+)',r'\1 \2',comment)
Replacing (- ,) before overall with dot	and I was only on 2,5 months - overall - DOES do what it is supposed to do - you need to weight whether it is worth the side effects.	comment= re.sub(r'[, -](\s*[oO]verall?\s+)',r'. \1',comment)
Respondents did not use any punctuation to indicate sentence stop.	I was hoping it would eventually go away so I continued to take it and after the 6th day of the same intense side effects I finally had to just chunk them in the trash.	No solution
Incomplete sentences	i recommend the cookies.. PS.. i have not gained any weight.. and if	No solution

3.4.2.2 Annotation environment. We used Microsoft Excel for developing the framework of annotation. The framework consists of sentences of post reviews arranged in rows and the defined items for sentence labeling arranged in columns. Each sentence was identified by the associated drug name followed by the review post id and sentence index indicating the position of the sentence in the review post. Annotators read the sentences and if a sentence fit into items, they entered 1 in the corresponding cell. The defined items in the framework were not mutually exclusive. In other words, a sentence may be coded as more than one individual item. For example, the following sentence was coded as both “effectiveness” and “perceived distress from ADR- low tolerability”:

“ It really helped me, however I suffered from side effects.”

Therefore, we coded it for both codes. **Table 3-7 and Figure 3-2** presents a snapshot of the coding environment.

Table 0-8 A Snapshot of the Coding (Annotation) Environment

Drug_ID	Sentence index	Sentences	ADR	WD	EF	INF	Others
effexorXR .217	4	“I started on on a low dose and gradually increased to 150 mg.”	0	0	0	0	1
effexorXR .217	5	“I feel so much better taking this drug.”	0	0	1	0	0
lexapro.40	1	“Lack of care for anything important in my life.”	1	0	0	0	0
lexapro.40	2	“Lack of feeling for anything, just lived day by day.”	1	0	0	0	0

3.4.2.2 Coding (annotation) Process. Four coders (annotators) participated in the annotation process. All the coders possessed a minimum qualification of M.Sc. degree in health-related studies. Two coders were pharmacy students and two had backgrounds in public health and health sciences.

We performed an interactive training for “sentence labeling” guidelines that involved mutual discussion between annotators to address the vague definition of items and conflicts in understanding each rule. We followed this step by refining the guidelines’ definitions, rules and examples.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
	com	drug_id	sentence	sentences	ADR	EF	INF	PPI-P	PPI-N	KN	SS	ADR-PA	ADR-NA	WD-PA	WD-NA	others
1			1	Lack of care for anything important in my life.	1											
2	1	lexapro.40	2	Lack of feeling for anything, just lived day by day.	1											
3	1	lexapro.40	3	Inability to Ejaculate during sex.	1											
4	1	lexapro.40	4	Compared to previous medication I was taking, Lexapro did in fact take away my physical pain I was suffering, but has caused so much more in emotional pain.												
5	1	lexapro.40	1	Nausea (for about 1 week), headaches, brain "zaps" occasionally, extreme sleepiness, vivid dreams, loss of sex drive.										1		
6	2	lexapro.41	2	After being on Wellbutrin for about 6 months with good results, I went off medication altogether, Unfortunately, my depression returned with a vengeance after about 7 months.	1											
7	2	lexapro.41	3	Went back to see a therapist.												
8	2	lexapro.41	4	Still not too helpful.						1						
9	2	lexapro.41	5	Went to see my Doc, was put on Lexapro 10mg.												
10	2	lexapro.41	6	Some good results.		1										
11	2	lexapro.41	7	Went to 20mg after some really tough life circumstances, Side-effects were unmanageable!												
12	2	lexapro.41	8	Went back down to 10mg after 1 week on the higher dose.										1		
13	2	lexapro.41	9	After 3 months on Lexapro and therapy, I will be switching to something else .												
14	2	lexapro.41	10	perhaps Wellbutrin again?												
15	2	lexapro.41	11	"BOTTOM LINE: you DON'T have to suffer through your depression, GET HELP!												
16	2	lexapro.41	12	The most important thing you can do is see a therapist.												
17	2	lexapro.41	13	Drugs are only a secondary issue.												
18	2	lexapro.41	14	If you do need drugs, be sure to closely work with your Doc, report side-effects promptly, etc, Don't give up on yourself.												
19	2	lexapro.41	1	Possible loss of libido.	1					1						
20	3	lexapro.42	2	This medication has done absolutely NOTHING for me.												
21	3	lexapro.42	3	I am so glad to be off of it.												
22	3	lexapro.42	4	I started at 5 mg, went up to 10 mg, 15, 20, 30, and even to 40 mg!												
23	3	lexapro.42	5	At that dosage I was so tired and irritable.	1										1	
24	3	lexapro.42	6	I do not see the point of taking Lexapro-type SSRIs for anxiety, which was my primary issue for taking it.												
25	3	lexapro.42	7	I do not think Lexapro gave me any seriously adverse side effects.												1
26	3	lexapro.42														

Figure 0-3 an expert of final analytical framework and comment analysis based on the framework

The review sample contains 887 posts for the four drugs, Zoloft (210), Lexapro (218), Cymbalta (230), and Effexor XR (229). According to the method by which the sample was created, for each drug, the reviews were divided into five subsets according to the rating value provided by the patients, which ranged from 1-5. Overall, we had 20 documents (spreadsheets) that were divided into three non-overlapping subsets, subset 1 (7 documents), subset 2 (7 documents), subset 3 (6 documents). Three of the coders proceeded with each of these sets. The fourth coders coded subset 4, which contained all documents. The review post sentences in each document were annotated by two annotators. The sentences in each document were coded by strictly applying the guideline's rules.

3.4.2.3 Calculating inter-annotator agreement. To test reliability of themes defined in the final framework, we calculated inter-coder agreement between pairs of annotators annotating each file. Assessment of inter-annotator Agreement (IAA) demonstrates consistency among observational ratings provided by multiple coders. In addition, the inter-coder agreement indicates to what extent codes are defined firmly and precisely to exclude all unrelated passages.

To calculate IAA, Kappa statistics are used to measure the observed level of agreement between coders for a set of nominal ratings and corrects for agreement that would be expected by chance, providing a standardized index of IRR that can be generalized across studies. The degree of observed agreement is determined by cross-tabulating ratings for two coders, and the agreement expected by chance is determined by the marginal frequencies of each coder's ratings. Kappa is computed based on the equation where $P(a)$ denotes the observed percentage of agreement, and $P(e)$ denotes the probability of expected agreement due to chance" (Hallgren, 2012):

$$k = \frac{P(a) - P(e)}{1 - P(e)}$$

Possible values for the Kappa coefficient ranges from -1 to 1, where 1 indicates complete agreement, 0 indicates completely random agreement, and -1 indicates complete disagreement. Moreover, according to the instruction for interpreting Kappa data provided by Landis and Koch (1977), 0.0 to 0.2 indicates “slight agreement”, 0.21 to 0.40 indicates “fair agreement”, 0.41 to 0.60 indicates “moderate agreement”, 0.61 to 0.80 indicates “substantial agreement”, and 0.81 to 1.0 indicates “almost perfect” or “perfect agreement”.

Hayes and Krippendorff (2007) introduced more conservative instruction for interpretation of Kappa values. They specified that Kappa less than 0.67 agreement should be discounted from analysis, variables with Kappa between 0.67 and 0.80 should be tentatively used in analysis, and variables with Kappa above 0.80 can be used for definite conclusion. Despite this instruction, in practice, variables with Kappa less than 0.67 are often retained in studies. Accepting or discounting a variable with Kappa less than 0.67 for testing hypothesis in qualitatively analysis strongly depends on a study’s research question and methodology.

To report IAA agreement calculated using Kappa, the following items should be considered:

- 1) The metric in which a variable was coded (e.g., nominal vs. ordinal, interval, or ratio);
- 2) design of the study (e.g., whether all subjects vs. a subset of subjects are rated by coders);
- 3) Intended purpose of the IAA estimate (e.g., to estimate the reliability of individual coders’ ratings vs. the reliability of the mean ratings from multiple coders);
- 4) Reporting both the statistic and its computational variant (Hallgren, 2012).

For the purpose of our study, we calculated Kappa statistics based on the following criteria:

- 1) The metrics in which all variables (codes) were coded were nominal variables.
- 2) Each file was double coded.
- 3) The intended purpose of the IAA estimate is to calculate reliability of individual coders.
- 4) To calculate the IAA agreement, we used `cohen_kappa_score` from module of `sklearn.metrics` in python. This function uses the statistical method introduced by Artstein and Poesio (2008).

Because the defined codes are not mutually exclusive (a passage may be coded into more than one individual theme), we coded them such that each individual code represented a separate binary variable (we dichotomized each variable (themes) into a binary variable, i.e., coded as ‘1’ for present and ‘0’ for absent). Using this procedure, we reduced the comparison into a 2*2 table based on whether each of the two coders reported the code as present or absent for each variable.

Table 3-9 shows the IAA for each code in the final framework, and examples of disagreement. For ADRs and withdrawal symptoms (WD), we decided to exclude all sentences with general mentions of these two items, such as "side-effects" in the sentence: "I really suffered from side-effects." For both drug effectiveness and ineffectiveness, we decided to include only sentences that explicitly referred to a drug's success or failure in improving or deteriorating a patient's symptoms and general health. For example, the sentences: "I feel much worse after using the drug" should be labeled as "Ineffectiveness", and the sentence "I had a really bad headache and dizziness, after my doc increased the dosage", should be labeled only as "ADRs", not both ADRs and Ineffectiveness. However, some annotators labeled sentences with mentions of ADRs as ineffectiveness, leading to a lower IAA for this item compared to other items.

Table 0-9 Computed IAA for each Themes in the Sentence Classification Component with Examples of Disagreement

Items	IAA	Example of disagreement	Annotator 1	Annotator 2
ADR	0.81	"I stopped taking the lexapro, the anxiety has lightened a little bit, but the crying hasn't."	ADR	WD
WD	0.69	"However, I have tried to taper and quit several times and CAN'T."	WD	Others
EF	0.82	"Helped a great deal then put on the generic and had a totally negative reaction."	EF	EF & INF
INF	0.65	"At the end of 6 weeks I still felt no difference but my best friend said they noticed a difference."	EF	INF
DI	0.91	"I really didn't think that I was depressed, but I think since taking the lexapro, my mod has lifted."	EF	EF & DI
ADR-PD-high	0.84	"I'd rather be normal sized and anxious than overweight."	ADR, SS, ADR-PD-high	ADR, SS
ADR-PD-low	0.89	"Night sweats, trouble sleeping, but all of these gone away after a few weeks."	ADR, ADR-PD-low	ADR
WD-PD-high	0.85	"When i miss a day i feel very spaced out, thick, groggy, sad."	WD WD-PD-high	WD
WD-PD-low	0.83	"I only experience a little nausea/dizziness if I miss a dose."	WD WD-PD-low	WD
PPI-P	0.50	"My doctor switched my dosage time to morning and I noticed a world of change." "I called my dr. office they told me to drop it down to 2.5 until they could see me."	PPI-P PPI-P	Others Others
PPI-N	0.59	"Every time I mentioned this to the doctors their only advice was to up the dose."	PPI-N	PPI-A
~KN	0.64	"This drug should not be given to anyone without the doctor notifying family/friends."	KN	Others
DXD-Dec	0.61	"I am going to stop the medication, because of severe side-effects"	DXD-Dec	DXD-S
DXD-S	0.75	"The withdrawal was horrible, I had to stop working for two weeks."	DXD-S	Others
DXD-W	0.65	"I am in the process of discontinuation, be careful about withdrawal."	DXD-W	DXD-S
DXD-F	0.88	"I run out of medication for few days, I felt horrible."	DXD-S	DXD-F
ATT-P		Not double coded– Removed from the final analytical framework		
ATT-N		Not double coded – Removed from the final analytical framework		
SUG-P		Not double coded– Removed from the final analytical framework		
SUG-N		Not double coded– Removed from the final analytical framework		
Dosage/duration		Not double coded– Removed from the final analytical framework		
~SS		Not double coded– Removed from the final analytical framework		
~FS		Not double coded– Removed from the final analytical framework		
Other drugs		Not double coded– Removed from the final analytical framework		
Other patients		Not double coded– Removed from the final analytical framework		

Most of drug review posts do not include any explicit indicators that could be used to measure patients' satisfaction with providers. For example, this sentence: "I talked to my doctor and he immediately suggested I come off of it", does not clearly imply the patient's attitude towards his or her provider. A patient may conclude that his or her provider did not have sufficient knowledge about treatment or may conclude that the provider considered his or her concern. This vagueness led to high rate of observational error and low rate of IAA.

For the variable, "complaining of lack of knowledge", we decided to include patients reports for searching on the WEB for ADRs or WDs or advise to readers to gain information about ADRs or WDs as the sign of lack of knowledge. However, some of the coders did not take into account this indicator as sign of patients' lack of knowledge, causing a lower IAA for this variable.

3.4.2.4 Resolving disagreement. Annotators' observational errors occurred due to differences in annotators' interpretations of the guidelines and the differences in their interpretations of the review posts. A program was written to scan for instances of disagreement. Instances of disagreement were then reviewed and discussed by the same annotators who annotated the respective document earlier. For a specific item, annotation was added or removed if it was marked by any of the annotators, given that they both agreed on the decision. Otherwise, the sentences were labeled as others. The harmonization was performed over the complete corpus in the presence of annotators.

Phase 5

3.5 Entity Identification

The focus of the first phase of the data analysis was on examining the **qualitative aspects of knowledge** (e.g., features, properties) using analytical frameworks. While the focus of this

phase, entity identification, is to extract medical entities including adverse drug reactions (ADRs), withdrawal symptoms (WDs), and drug indicators (DIs), and qualifiers representing the severity (QS) and persistency (QP) of ADRs, WDs, and DIs.

Identification of ADRs and WDs from consumer health posts can directly provide a list of various types and frequency of ADRs and WDs associated with psychiatric medications. The ADRs and WDs may not be detected by clinical trials or post-marketing surveillance systems established by regulatory agencies, such as the Adverse Event Reporting System (AERS) of Food and Drug Administration (FDA). In addition, it also enables the generation and testing of hypotheses related to the association between different type of ADRs and WDs and discontinuation, adherence, and attitude.

In the sentence-labeling phase, we labeled sentences for presence of ADRs, WD, and DIs. In this phase, we were able to select sentences with mentions of these entities and then proceed with the process of identification.

3.5.1 Guidelines for entity identification

Guidelines of entity identification includes the entities' definitions and rules for proper identification of the entities. **Table 3-10** includes definitions of entities and rules for entities identification. The rules are related to patients' certainty in linking ADRs or WDs with drugs, identifying patients' subjective complaints and functional problems, as well as excluding unnecessary context such as simile and metaphors from the span of ADRs, WDs, and DIs.

Identifying patients' subjective complaints is important because they may reflect subtle physiological, psychological, or cognitive ADRs associated with drugs. For example, “*feel like I could not stop moving*” reflects patient restlessness, which is a sign of akathisia (ADR of psychiatry medications). As another example, “*body move in coordination with other people's*

bodies” (Echopraxia) indicates an extrapyramidal symptom. Any of these symptoms have pharmacologically related affective components that may lead to a patient’s negative perception of drugs and eventually drug non-adherence.

Identifying functional problems in drug review posts is also significant, not only for understanding how ADRs influence patients' normal daily activities and interpersonal relationships, but also for estimating the indirect cost associated with the ADRs. Overlooking functional ADRs when patients suffered from them has significant negative affects on expected clinical outcomes from medications. Moreover, collecting this information also enhances clinicians' abilities to predict the impact of ADRs on patients’ functionality, such as limitations on daily activities and social participation and restriction on work performance. Hence, they will be able to design more effective interventions targeting patients' attitude and adherence towards medications. The significance of representing signs or symptoms as functional problems is explained in detail in the International Classification of Functioning, Disability and Health (ICF) documents (World Health Organization 2010) (Giannangelo et al., 2005) . The identified ADRs, WDs, and DIs were categorized as physiological (Phys), Psychological (Psycho), Cognitive (Cogn), and functional problem (FP).

3.5.2 Entity identification process

Four annotators participated in the process of identification and extraction of the three entities explained in **table 3-10**. We first selected the sentences with labels of ADRs, WD, and DIs and exported them to new excel spreadsheet documents. In the second step, the documents were divided into three sets and each set was reviewed by an annotator for entity identification.

Table 0-10 Entities' Definitions and Identification Rules with Examples

Entity	Definitions	Example	Rules for identification	Example
ADR	Any sign or symptom that the patient explicitly associated it with drug consumption, except the phase of dosage reduction and discontinuation.	My doctor increased my dose and I felt <u>severe dizziness</u> (ADR).	<p>1. Certainty: If a patient is not confident about the association between ADR/WD and drug, the ADR/WD is not extracted.</p> <p>2. Subjective complaints: If ADR or WD was expressed as subjective complaints, it should be extracted with the entire necessary context.</p> <p>3. Functional problems: if patient mentioned his complaint as a functional problem, such as problem with daily functioning and social activities, it should be extracted and labeled as ADR/WD.</p>	<p>1. It caused hair loss and <u>stomach bloating</u> (ADR), however I am not sure that hair loss (not ADR).</p> <p>2. It certainly erased the anxiety, but I <u>hardly feel human anymore</u>. (ADR)</p> <p>3. I would just stay around and <u>do nothing all day</u> (ADR)</p> <p>4. “Very hard to take a deep breath (ADR) like someone is squeezing my lungs. (Simile –non necessary)</p> <p>5. The <u>anxiety</u> (ADR) was debilitating. I also had <u>severe headache</u> (ADR), but the <u>anxiety</u> (ADR) was worse.</p> <p>6. Anxiety is now though the roof (qualifier-severity)</p> <p>6.2. Constant (qualifier-persistency) <u>bad</u> (qualifier-severity) headaches</p>
WD	Any sign or symptom that patient explicitly associated it with the phase of dosage reduction and discontinuation of a drug.	“Trying to come of has gotta be worse than heroin, <u>Uncontrollable rage</u> ,(WD) and <u>lots of emotional attachment with partner</u> (WD).”	<p>4. Excluding simile and metaphor: If a patient used a simile or metaphor to provide information about his/her feelings towards ADR/WD/DI that simile or metaphor should not be extracted.</p> <p>5. Duplicates: Duplicate ADR/WD/DI in a sentence should be independently extracted, that is, all the occurrences of the entities are identified.</p> <p>6. Qualifiers: If ADR/WD/DI is associated with qualifiers presenting severity or persistency of that entity, the qualifier needs to be identified.</p>	
DI	Any sign or symptom that patient mentioned as the reasons for drug consumption/prescription.	“My <u>depression</u> (DI) disappeared after drug usage.”	<p>1. Treated symptoms: a patient may also mention a sign/symptom that was relieved by drug consumption. That sign or symptom is also a drug indicator.</p> <p>If a patient’s sign/symptom became worse because of drug consumption, the sign/symptom should be labeled as both DI and ADR of the drug.</p>	This drug reduced my <u>sense of guilt</u> (DI).

For the purpose of calculating inter-annotator agreement for entity identification, the entire dataset was reviewed by the fourth annotator. Entity identification and extraction was

conducted based on the rules specified in the guidelines. We did not extract general mentions of entities, such as “side-effects” and “withdrawal” in the sentences. For example, in these sentences, “*I really suffered from side-effects,*” and “*withdrawal seems impossible*”, side effects and withdrawal were not extracted.

Prepositions and possessive adjectives were excluded to improve consistency in spans of identified entities. For example, in “*my anxiety became worse,*” *my* was not extracted as part of anxiety. In the last step, the disagreement between annotators for the identified entities were resolved by Dr. Fontelo (a PhD committee member).

3.5.2.1 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):

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

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. Since entities were identified from the sentences with predefined labels (tags), by default, they have the label of the sentences. Therefore, we excluded parameter β from the calculation. For parameter α , we set the span on strictness. Therefore, the identified entities must match exactly. For example, for identifying ADRs from this sentence: “*the headache was terrible*”, if annotator A identified “headache” as a side-effect, and annotator B identified “headache was terrible” as a side-effect, then the matching between annotators A and B is “0”.

The computed pairwise agreement for strict match are 0.86 for SFX , 0.81 for WD and 0.91 for SS, with the average 0.86 for the entire corpus. The reason for applying pairwise agreement (PA) for computing IAA rather than conventional measures, such Kappa, is that PA is calculated at the level of entities. Since the identification task requires identifying the entities and determining their correct boundaries, the chance agreement is effectively zero.

Phase 6

3.6 Terminology Association (Entities Normalization)

While sentence classification and entity identification in drug review posts have significant implications for text mining and machine learning systems focusing on information retrieval, mapping patient expressions of these entities to the language of health professionals fills the gap between a layperson expressions and professional expressions of medical entities. This translation may benefit the generation and testing of medical hypotheses by providing unambiguous and standard information for statistical data collection and analysis.

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). An example of such normalization is CADEC corpus, in which the identified entities from drug review posts were mapped to SNOMED-CT, MEDRA, and ATM (Karimi et al., 2015).

To normalize the entities in our corpus, we mapped the identified entities including ADRs, WDs, and DIs to their corresponding concepts in both the Unified Medical Language System Metathesaurus (UMLS Meta) and the Systematized Nomenclature of Medicine—Clinical Terms (SNOMED-CT).

UMLS Meta is a compendium of many standard medical vocabularies that provides a mapping structure among vocabularies, allowing one to translate among various terminology systems. UMLS Meta is organized by concepts. Each concept is assigned one or more semantic types (categories) and a Concept Unique Identifier (CUI). Semantic type covers a broad category of medical concepts, such as sign/symptoms, mental or behavioral dysfunction, and intellectual products. Different lexical representations of the same concept can be identified by CUIs. UMLS Meta has been used for identification and retrieval of medical entities from clinical and biomedical literature and clinical notes in Electronic Health Record (EHR) Systems . Automatic extraction systems built on UMLS Meta, such as MetaMap, have demonstrated high recall and precision for identifying medical entities in biomedical literature and clinical notes (Fung, Jao, & Demner-Fushman, 2013). However, as Nikfarjam et al. (2015) showed, MetaMap performance in entities identification (ADRs and drug indications) from consumer health posts showed 47% precision and 39% recall on corpuses from the “Daily Strength” forum and Twitter. Mapping ADRs and drug indications to UMLS Meta, in addition to the normalization benefit, reveals consumer health vocabulary that have not been covered by medical terminologies.

SNOMED CT is considered to be the most comprehensive, multilingual clinical terminology in the world. The primary purpose of SNOMED CT is to encode the terms available in health information to support effective clinical recording for improving patient care. Accordingly, it is intended to provide a core general terminology for EHR systems. Mapping layperson expressions of medical entities to SNOMED-CT supports seamless information sharing between Personal Health Record (PHR) systems and EHR systems. In addition, Mapping to SNOMED-CT as the most comprehensive standard terminology supports providing an unambiguous data collection for statistical analysis. Moreover, mapping data to both UMLS

Meta and SNOMED-CT reveals the deficiency of SNOMED-CT in covering patients' expressions of medical terms compared with other terminologies in UMLS Meta.

3.6.1 Guidelines for Terminology Association

The process of mapping laypersons' expressions of medical entities to professional expressions is based on an assumption that consumers' and professionals' terms express the same underlying concepts. For example, the consumer term "feeling sick to my stomach" expresses the same concept as the professional term as "nausea". By considering this assumption, most research on normalizing consumer health vocabularies has focused on the terms rather than the underlying concepts that the terms carry (Keselman et al., 2008). However, proper mapping involves both syntactic and semantic mapping between terms. For semantic mapping, we need to understand the conceptual models in both professional and consumer terms, which can be mostly identified in a term's definition and a term's context. Accordingly, for this study, we applied the process of entity expression mapping to both UMLS Meta and SNOMED-CT by creating a guideline taking into account their potentially different conceptual models.

The guidelines have been constructed based on the reviews of clinical trial studies targeting ADRs of psychiatric mediations and qualitative studies investigating the themes of patients' experiences with the drugs. We reviewed literature concerning three types of ADRs associated with psychiatric medications, including physiological, psychological, and cognitive. Each set of these studies contains descriptions and clinical features associated with a specific ADR or a group of ADRs that were mostly developed based on clinical trial reports. Studies that focus directly on patients' experience reports mostly attempted to classify the experiences by broad themes. The identified themes are usually equivalent to medical concepts used in professional settings. Here, we provide examples for each set of studies associated with a type of

ADR and how they can be used for translating patients' reported experiences to professional terms.

- 1) Example of Physiological ADRs: Psychiatric medication can cause Akathisia (Salem, Nagpal, Pigott, & Lucio Teixeira, 2017), which is described with the following clinical features: feeling of inner restlessness, compelling need to be in constant motion, as well as motor restlessness, such as "inability to sit" or "feeling the need to pace". These underlying concepts of Akathisia, patients' expressions of ADRs indicating inner restlessness, such as "feel like I was inside screaming" and motor restless, such as "I am not able to sit" could be translated to Akathisia. Another example of ADRs is "brain shivers". Brain shivers are a sensation of "sudden shake, vibration, tremor, jolt" in addition to electric shock that patients feel in the brain and that mostly occur due to missing dosages or discontinuing the drugs abruptly. Other terms that have been used to describe this symptom are "electric brain thingies," "brain zaps," and "brain flips" (Aronson, Jeff 2005). Hence, any mentions of these terms with the same underlying concept, such as "feeling electrical shock in the brain" in the review posts can be translated to "brain shivers".
- 2) Example of Psychological ADRs: Psychiatric medications are associated with a general syndrome of indifference as a behavioral syndrome and as an emotional syndrome (Sansone & Sansone, 2010). Behavioral indifference, a clinical feature of "apathy", is manifested by patients' lack of desire to engage in activities, lack of desire to make any changes, not caring about anything, or similar symptoms. Therefore, any terms reflecting behavioral indifference can be translated to "apathy". For example, "just don't care", and "just lived day by day" are equivalent to apathy. The emotional aspect of indifference is characterized by a restricted range in expressing emotions and feelings, a sense of blunting

the emotions, or of feeling numb. Accordingly, clinical features of emotional indifference could be described by two main concepts, “blunted affect” and “lack of capacity to feel emotions”. Therefore, any patients’ expression reflecting the restricting range of emotion could have the same underlying concepts. For example, “very slow to excite”, “dull mood”, “want to express myself and cry but cannot” can be translated to blunted affect.

- 3) Example of Cognitive ADRs: Executive dysfunction as a cognitive ADR of psychiatric medication is associated with the 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, any patient’s complaints reflecting the underlying concepts of executive dysfunction, such as “cannot follow through on simple tasks” can be translated to “executive dysfunction”.

3.6.2 Creating a dictionary

To reduce the time to accomplish the mapping the identified ADRs (SFX and WD) and drug indications to UMLS Meta and SNOMED-CT, we first developed a dictionary using 2,100 mentions of ADRs extracted from 240 review posts. First, we looked for syntax mapping of the identified entities with UMLS Meta concepts using the UMLS API. If a syntax match was not available, we translated the colloquial expressions of ADRs to proper professional medical terms using the guidelines, and then we selected a proper UMLS Meta concept for the term. To ensure the selected UMLS Meta concept reflects the original term, we checked the parent terms of each concept to ensure that the UMLS Meta concept and the SNOMED-CT concept associated with it do not carry additional meaning. The additional meaning would be inherited from the parent terms that are not related to the patients' experience with the medication. This dictionary was reviewed for accuracy and was set as the reference code for mapping the identified entities to

both UMLS Meta and SNOMED-CT. We updated the guidelines and the dictionary periodically to include new expressions of ADRs and indicators.

3.6.3 Procedure of mapping

The mapping procedure and selecting the proper UMLS Meta and SNOMED-CT concepts was conducted based on the following procedure. Annotators used the UMLS Terminology Services, UTS browser (2017) for the finding proper UMLS and SNOMED-CT concepts.

- 1) Search for syntax matching by setting “search type” on “normalized word” and sources on “all sources”.
 - a) In the case of a unique concept result, evaluate the concept using the requirements for a proper concept (**Table 3-11**).
 - b) In the case of multiple concepts results, select the proper concept using the requirements for a proper concept (table 3-11).
 - c) In the case of no concept returned, go to step 2 if the original term is associated with a qualifier. Otherwise, go to step 3.
- 2) Search for a partial match by removing qualifiers associated with the original term.
 - a) Follow sub-steps “a” to “c” specified in step 1.
- 3) Search for a synonym using the guideline of the terminology mapping
 - a) Follow sub-steps “a” to “c” specified in step 1.

Figure 3-3 shows the detail of the procedure for identifying proper UMLS Meta and SNOMED concepts for the identified entities. **Table 3-11** presents the requirement for selecting proper UMLS Meta and SNOMED-CT concepts. In the case of availability of multiple UMLS Meta concepts for the original term, the proper concept needs to include the expressions of the most recent versions of SNOMED-CT. If multiple UMLS Meta candidates met the mentioned requirements, the proper UMLS Meta concept has a SNOMED-CT expression that is closest to the original term. Using the flowchart (**Appendix A**) and requirements for finding proper concepts, we mapped identified entities to both UMLS and SNOMED-CT concepts.

Table 0-11 Requirements for Selecting Proper UMLS and SNOMED CT Concepts

Requirements	Description	Example
(1) Definition	<ul style="list-style-type: none"> - Definition of a proper concept should cover patient's specific physiological, behavioral, emotional, cognitive, or functional problem. - If the patient did not specify the severity or type of a reported adverse effect, we used the most general concept (code) that represents the patient's reported problem. 	For patient complaint "takes a long time to get to sleep", UMLS concept: "initial Insomnia". "Sleeplessness" is not correct, because it includes all phases of insomnia.
(2) Semantic type	The semantic type of proper concept includes "finding", "sign or symptom" or "mental or behavioral dysfunction". However, in some cases, other semantic types, such as "individual behavior" for concept "aggressive behavior" is a proper map.	
(3) Hierarchical structure (ancestors and children)	The SNOMED-CT concept should not convey additional meaning inherited from ancestors (parents) that are not related to patient's complaint.	For the patient complaint "inability to eat", Aphagia [C0221470] is not a proper map, because the concept is linked to ancestors of swallowing finding in SNOMED-CT. While the patient with depression does not have any problem swallowing, they have a problem with loss of appetite. So [C1971624] "Loss of appetite" is a proper map.
(4) Including SNOMED-CT concept	In a case of existing multiple UMLS candidates, the preferred concept is a concept that includes the most recent SNOMED-CT concept.	For the patient complaint "memory loss", "Amnesia [C0002622]" compared with "Memory loss [C0751295]" is the proper concept, because memory loss includes the obsolete version of SNOMED-CT concept.
(5) Syntax match with SNOMED-CT concept	In the case of existing multiple UMLS candidates meeting requirement (1), (2), and (3), the preferred match has SNOMED-Concept that syntactically match with the original term.	For example, for the patient complaint "restlessness", the UMLS concept candidates are "Agitation [C0085631]", Restlessness [C3887611], Psychomotor Agitation [C3887612], and Akathisia [C0392156]. The proper match would be "Restlessness [C3887611]" due to including the SNOMED-Concept preferred term (PT) with the same syntactic match.

There were some cases in which a UMLS Meta concept that is the best match for an original term lacks SNOMED-CT concept. In this case, we searched for another UMLS Meta synonym concept that is expressed in SNOMED-CT. For example, for the original term "*Zombie like*", the closest UMLS Meta concept is "Felt like a zombie [C0857486]", which has no SNOMED-CT expression. In that case, we used UMLS concept "Emotionally detached [C0233754]" that include SNOMED expression.

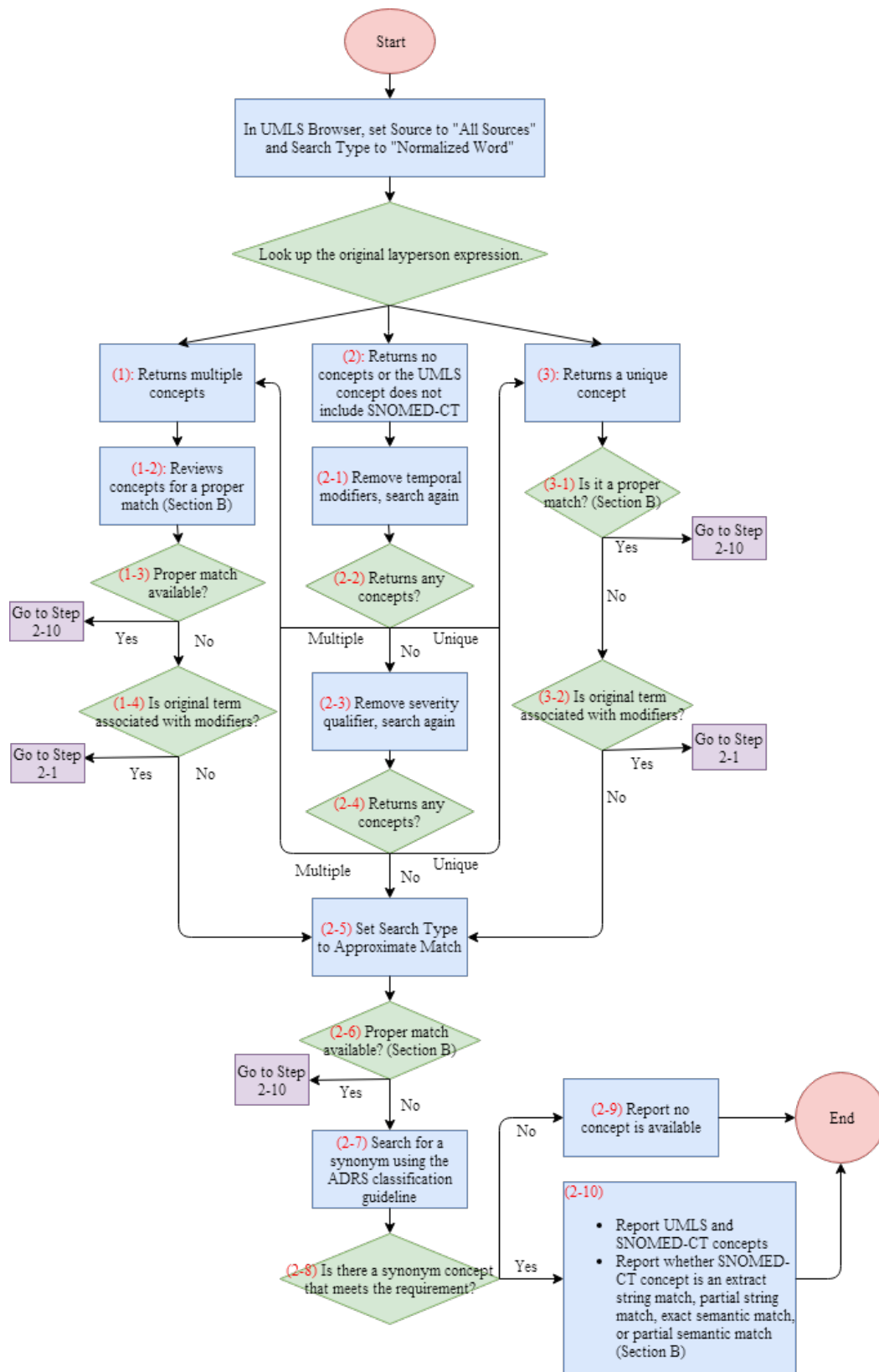


Figure 0-4 Flowchart of Finding Proper Concept for Layperson's Expression of Medical Entities

Table 3-12 provides example of mapping layperson expressions to both UMLS and SNOMED-CT.

Table 0-12 Examples of Mapping Entities to both UMLS and SNOMED-CT Concepts

drug_id	Sen _Id	original Term	UMLS (1)	SNOMED-CT (1)	UMLS (2)	SNOMED-CT (2)	Qualifier
cymbalta. 124	1	Felt sick	C0857027 / Feeling Sick /Sign or Symptom	No code	C0231218/ M alaise / Sign or Symptom	Undifferentiate d illness: Vague ill health (finding) [A3771172/SN OMEDCT_US/ FN/248282002]	
lexapro.1 2	3	"zombie" like	C0857486/ Felt like a zombie/ Finding	no code	C0233484 / Emotionally detached / Finding	Emotionally detached (finding) [A3413164/SN OMEDCT_US/ FN/24936000]	
cymbalta. 12	2	loose my friends	C0524322 / Personal relationship breakdown / Human- caused Phenomenon or Process	Personal relationship breakdown (finding) [A3626907/SNOME DCT_US/FN/105414 008]			
cymbalta. 131	1	Excessive sleepiness	C0013144/ Drowsiness/ Finding	Drowsy (finding) [A3406202/ SNOMEDCT_US/ FN/271782001]			Excessive (QS)
Effexor.7 8	1	minor muscle spasms	C0037763 / Spasm/ Sign or Symptom	Spasm (finding) [A3712892/ SNOMEDCT_US/ FN/45352006]			Minor (QS)
effexor.9 7	2	Sweating like crazy all the time	C0700590 / Increased sweating / Sign or Symptom	Excessive sweating (finding) [A3440966/ SNOMEDCT_US/ FN/52613005]			All the time (QP)
effexor.1 1	7	Brain zap	no concept	no concept			

Phase 7

3.7 Usability of the Dataset (Corpus)

Development of structured data from drug reviews in consumer health posts has several implications in the area of natural language processing, such as improving performance of text mining algorithms, and generating and testing medical hypotheses. In the next phase we show

how the dataset can be used for testing hypotheses concerning patients attitudes toward antidepressants.

3.7.1 Identifying the underlying factors affecting patient attitude towards drugs

The following hypotheses were formulated based on the availability of information in the dataset generated through phase 3 (developing the analytical framework) and phase 4 (applying the analytical framework) and literature review focused on measuring association between pharmacological treatment factors, healthcare system factors, social-cognitive and psychological factors, patient-related factors, and depression factors with patients attitudes towards antidepressants. The list of hypotheses is as follows:

- Testing association between personal variables and attitude

1. **Age:**

- a) H_0 : There is no association between age and attitude.

2. **Gender:**

- b) H_0 : There is not association between gender and attitude

Testing association between clinical variables and attitude:

1. **Duration:**

- c) H_0 : There is no association between duration and attitude

2. **Drug Effectiveness:**

- a) H_0 : There is no association between drug effective-ness and attitude.

3. **Drug ADR:**

- a) H_0 : There is no association between drug adverse reaction and attitude.

4. **ADRs-PD:**

- a) H_0 : There is no association between perceived distress received from ADRs (ADRs-PD) and attitude.

5. **LK:**

- a) H_0 : There is no association between Lack of knowledge and attitude.

If this null hypothesis was rejected, the following null hypothesis will be tested:

6. **WD-EXP:**

- a) H_0 : There is no association between experience of withdrawal (WD-EXP) and attitude. Experience of WD includes any experiences of intentional (Weaning off or stopping) or unintentional (missing dosage) withdrawal of a drug.

7. PPI:

- a) H_0 : There is no association between patient physician interaction (PPI) and attitude.

In addition to the hypotheses, the relationship between variables modeled to specify the significance of each variable in forming a patient's attitude towards antidepressants. We also hypothesized that:

- 1. ADR-PD and drug effectiveness are the most significant factors affecting patients attitudes toward antidepressants.

3.7.1.1 Preparing the dataset for testing hypotheses

To increase the accuracy of data analysis and enhance the level of inter-annotator agreement (IAA) between coders, the unit of analysis was set at the sentence level rather than the whole review. Therefore, for an individual review, there may be multiple occurrence of the same variable (code). For testing hypotheses, we reduced the multiple occurrences to a single one indicating presence or absence of a variable in the drug review.

A single drug review may be coded for contradictory codes, such as both "Effectiveness" and "Ineffectiveness". That is, a patient may find medications effective for a certain point of time, such as acute phase of treatment, but in the maintenance phase, they may not be satisfied with the drug. Therefore, for reviews coded as both "effectiveness" and "ineffectiveness", we recoded both effectiveness and ineffectiveness.

In addition, a review may be coded for both ADR-PD as high and ADR-PD as low. That is, in one part of his or her review, a patient may report an ADR was severe and has debilitating effects on quality of life, while another ADR may be mild and does not have such an effect. For such reviews, I only include ADR-PD-high, because overall, ADR-PD for patient is high.

Table 0-13 Lists of the Variables, Possible Values, and Limitations Associated with each of them for Measuring Patient Attitude toward Antidepressants

Variables	Possible values	Limitations for measuring attitude
ADR	<ul style="list-style-type: none"> - Presence - Absence - Unknown: no explicit report on presence or absence of ADR. 	<ol style="list-style-type: none"> 1) The patient might fail to distinguish between ADRs and WDs and report all sign/symptoms associated with drugs as ADRs 2) For some reviews, there is no explicit report on presence or absence of ADR.
ADR-PD	<ul style="list-style-type: none"> - High - Low - Moderate: when patient did not use any indicators showing the perceived distress was high or low, then it is moderate. 	<ol style="list-style-type: none"> 1) AD-PD involves only three level of distress. These categories do not include nuance difference for ADR-PD for patients. 2) The definition for ADR-PD may be incomplete and do not include all the cases for ADR-PD-high. For example, persistency of ADRs indicates ADR-PD for some patients, while in definition we did not include it as sign of ADR-PD.
WD	<ul style="list-style-type: none"> - Presence - Absence - Unknown, if patient reported they experience intentional or unintentional WDs, but no explicit report on presence or absence of ADR. 	<ol style="list-style-type: none"> 1) Patient may fail to distinguish withdrawal from ADRs and therefore did not report all WDs. 2) Patients who experienced the WD, there is no explicit report on the presence or absence of WD.
WD-PD	<ul style="list-style-type: none"> - High - Low - Moderate: patient did not use any indicators showing the perceived distress was high or low, then it is moderate. 	WD-PD is subject to the same limitations explained for ADR-PD
EF	<ul style="list-style-type: none"> - EF: drug was effective, - INF: drug was ineffective - EF-INF: patient reported both effectiveness and ineffectiveness of the drug during the process of treatment. - Unknown: no explicit report on EF or INF of the drug. 	<ol style="list-style-type: none"> 1) Patients may fail to realized drug effectiveness because of the high perceived distress received from ADRs. 2) A patient may fail to report drug effectiveness, because there is no request for reporting it in the healthcare forum. 3) The definition for drug effectiveness may not be complete. For example, a patient might attempt to show drug ineffectiveness through reporting ADR or ADR-PD.
LK	<ul style="list-style-type: none"> - Presence - Absence 	<ol style="list-style-type: none"> 1) A Patient might fail to report their lack of knowledge or conducted search to gaining information about ADRs/WDs. 2) Definition for complaining about lack of knowledge may be incomplete.
WD-EXP	<p>DXD-Dec: existence of any indicators showing the patient's decision for discontinuation, such as "I am going to stop"</p> <p>DXD-WD: existence of any indicators showing patient experienced process of discontinuation, such as "weaning off" or "tapering off".</p> <p>DXD-S: existence of any indicators showing patient already stopped the medication, such as "stopped".</p> <p>DXD-F: existence of any indicators showing patient experienced withdrawal unintentionally, such as "missing a dose"</p>	Patient might fail to report intentional or unintentional withdrawal from drug.
PPI	<ul style="list-style-type: none"> - Patient perceived communication-positive - Patient perceived communication-negative - Unknown: no explicit report on patients' satisfaction with healthcare providers. 	<ol style="list-style-type: none"> 1) Most of the patient did not declare any information about their perspective with providers. 2) Patient mentions of communication did not clearly convey their attitude toward providing, causing high disagreement between annotators (coders) for this variable. 3) This variable was defined as binary variable, which may not reflect the slight differences between patient attitude toward medication.

Table 3-13 lists the variables, possible values, and limitations associated with each variable for measuring patient's attitude toward antidepressants.

3.7.1.2 Strategy for handling missing values

To handle the missing values, we adopted different imputation methods regarding the nature of the missing values for each variable. There are three assumptions for handling missing data:

a) Missing completely at random (MCAR): “When observations of a variable are missing completely at random, the missing observations are a random subset of all observations; the missing and observed values will have similar distributions”(Bhaskaran & Smeeth, 2014). In other words, the absence of data is unrelated to other variables in the model.

b) Missing at random (MAR): “Missing at random means there might be systematic differences between the missing and observed values, but these can be entirely explained by other observed variables” (Bhaskaran & Smeeth, 2014) . In other words, MAR means that the variable having missing data can be fully accounted for variables on which we have full information.

c) Missing not at random (MNAR): MNAR is data that absence is neither MAR nor MCAR. For example, if a patient fails to report their review for drug because of difficulty of filling the form or fear of breaching confidentiality, the missing values are MNAR.

The following strategies have been used for handling the missing values from the dataset.

- **Elimination of drug reviews**
 - 1) Eliminating reviews with no text; accordingly five reviews were eliminated
 - 2) Eliminating reviews that did not include any information for the variables for testing the hypotheses; accordingly seven variables were removed from the sample.

- 3) We removed the variable “PPI” from the analysis, because of the low IAA calculated using Kappa method and high number of missing values. The IAA for PPI-Positive (0.5) and PPI-Negative (0.59), with overall 0.55 for PPI.

- **Imputation of missing values**

- 1) Age: The assumption for missing values for imputing missing values for this variable is MCAR. It was assumed that the missing value has the same distribution of the observed values. Therefore, we used mean of data to fill in the missing values.
- 2) Gender: The assumption for imputing missing values for this variable is MCAR. Because this data is categorical, we used the mode of the variable to impute the missing values.
- 3) Duration of usage: The assumption for imputing missing values for this variable is MACAR. Because the data is very skewed, the median of data for this variable was used for data imputation, rather than the mean of the data.
- 4) ADR: The assumption for imputing missing values for this variable is MNAR. If patient did not explicitly reported any ADRs, I assumed that ADR is presence for the patient.
- 5) ADR-PD: The assumption for imputing missing values for this variable is MNAR. According to the definition, a perceived distress from ADR is high, if ADR is associated with qualifiers indicating severity or debilitating effect on patient’s daily life. On the other hand, the perceived distress is low, if ADR is associated with qualifiers indicating mildness of the ADR. The detail definition of these two variables is presented in **Table 3-6**. Accordingly, if the indicators are not available, the ADR-PD was coded as “NA” (missing value), which may imply that patients found the ADR neither severe, nor mild. Therefore, we imputed the missing values as “moderate”.
- 6) **EF**: The assumption for imputing missing values for this variable is MAR. For imputing missing values for this variable, the following procedure was conducted:

1. Drug reviews with missing values were eliminated from the dataset.

2. The association between EF and ADR-PD, WD-experience, LK, age, gender, and duration were computed using the Chi-square test.
 3. Regarding the variables associated with variable EF, K-nearest neighborhood (KNN) was used for the whole dataset to impute the missing values.
- 7) WD-Experience: WD-experience includes the patient's intentional and/or unintentional withdrawal of WD, including variables "DXD-S", "DXD-W", and "DXD-F". The assumptions for coding these variables were, if the patient explicitly reported experience of intentional or unintentional discontinuation, then WD-Experience is "presence", otherwise the WD-Experience is "absence". Accordingly, there was no missing value for this variable.
 - 8) LK: For variable lack of knowledge (LK), if a drug review includes any indicators showing lack of knowledge (**table 3-6**), the review was coded for LK-presence, otherwise the LK-absence. Accordingly, there was no missing value for this variable.

3.7.2 Testing Association between ADRs and Attitude

Identifying ADRs that are significantly associated with negative attitude towards antidepressants is important from the clinical perspective. It is important that healthcare providers understand whether exposure to different type of ADRs can increase the risk of developing negative attitudes and consequently non-adherence behavior in patients.

To measure the association between antidepressants' ADRs and attitude or adherence, the Antidepressant Side-Effect Checklist (ASEC) (Uher et al., 2009) has been used. The ASEC was constructed as a self-report scale to measure twenty-one physiological antidepressants' ADRs. For each item, the patient can specify whether a symptom (if present) is likely to be linked to antidepressants adverse effects (yes or no).

Patients' expression for ADRs in a healthcare forum are associated with wide semantic variation, leading to limitations in statistical analysis for testing hypothesis related to different

ADRs and attitude. This study normalized the semantic variability of ADRs by mapping them to UMLS Meta concepts. The details of the normalization was provided in “3.6 Terminology association”. Expressions of ADRs were mapped to the closet syntax match in both UMLS and SNOMED-CT. Therefore, some ADRs maybe syntactically different, but semantically related UMLS Meta concepts. For example, “headache” (mapped to the UMLS Meta concept “[C0018681] Headache”) and “severe headache” (mapped to UMLS Meta concept “[C2957106] severe headache”) are semantically related. For each ADR mentioned in the ASEC, we grouped semantically related UMLS Meta concepts. In the next step, we calculated the total frequency of items in each group, and then tested associations between each group of ADRs and patients’ attitudes toward antidepressants.

3.7.3 Statistical Methods

For testing hypothesis for this study, we used Chi-square, ANOVA, and ordinal logistic regression.

Chapter 4: Results

4.1 Dataset Statistics

The original sample size was calculate using formula for calculating sample size, which was 892 reviews, with 432 in the SSRI class (Zoloft (213) and Lexapro (219)) and 460 in the SNRI class (Cymbalta (231) and Effexor XR (229)). Five posts were excluded from the sample, because they lacked any content. **Table 4.1** shows statistic for sample of this study.

The final sample is almost equally divided between SSRI and SNRI classes. Table 4 shows demographic information of the whole sample, as well as each class of drugs separately. Almost 50% of the reviews was recorded by patients who were satisfied with drugs, while 33% of patients were dissatisfied with the drugs, indicating that patients with negative experience are not dominant in this healthcare forum. 68% of the posts were posted by patients with age less than 40. This is possibly due to higher access to Internet in young patients and their willingness to report their experience with medications. The gender proportion in the sample for female is significantly higher than for male for both classes of drugs. Duration of usage is highly skewed due to the effect of outliers. Posting reviews after 1 day of treatment may indicate high concern of patients about the treatment.

4.2 Statistics on Developing and Applying the Analytical Framework

For the first phase of this study, the framework method was used to provide structured data from text-based consumer reviews of the antidepressant drugs specified in this study. In the framework method, a hybrid approach of inductive and deductive analysis was applied for coding and theme development. In the deductive approach, a comprehensive set of articles and self-report scales (questionnaires) that has been used in studies with focus on measuring patients' attitudes towards antidepressants and psychiatric medications have been reviewed to identify the important factors affecting patients attitudes towards antidepressant medications. The identified

factors were categorized into five major categories, including pharmacological treatment, healthcare systems, psycho-social factors, patients-related factors, and disorder related factors.

The factors were used for building the initial Framework of analysis.

Table 0-1 Dataset Statistics

Dataset statistics	Dataset	SSRI	SNRI
Sample Size	892	432	460
No. of reviews with text	887	428	459
Time span	Feb 2001 Sep 2016	Feb 2001 Aug 2016	July 2004 Sep 2016
Attitude	3.16	3.33	3
Attitude rated as 1	195 (23%)		
Attitude rated as 2	104 (12%)		
Attitude rated as 3	152 (17%)		
Attitude rated as 4	209 (24%)		
Attitude rated as 5	219 (25%)		
Gender	F 669 (76%) M 212 (24%) Missing value (11)	F 310 (72%) M 118 (28%) Missing value (4)	F 359 (79%) M 94 (21%) Missing value (7)
Age	Avg. 37 Med. 35 Sd = 12.03 Missing values (10)	Avg. 35 Med. 34	Avg. 38 Med. 37
Age range	14-83	14- 73	14- 83
Age categories			
<20	49 (6%)		
>= 20 <30	242 (27%)		
>=30 <40	249 (28%)		
>=40 <50	200 (23%)		
>=50 <60	106 (12%)		
<=60	33 (4%)		
Duration of usage	Avg. 18 months Med. 5 month SD. 31.7 (month)	Avg. 19 months Med. 5 months	Avg. 17 months Med. 5 month
Duration of usage (range)	1 day - 20 years	1 day- 16 years	1 day - 20 years
Duration of usage categories			
< 1 month	215 (24%)		
>= 1 month < 3 months	116 (13%)		
>= 3 months < 6 months	120 (14%)		
>= 6 months < 1 year	125 (14%)		
>=1 year <2 year	82 (9%)		
>=2 years <5 years	128 (15%)		
>=5 years <10 years	66 (7%)		
>= 10 years	27 (3%)		

300 drug reviews were coded using the codes (themes) at initial analytical Framework. Using inductive approach (open coding), passages of text that did not match into the codes in initial framework were discussed for more analysis in weekly meeting. Overall, eleven new themes were identified, including withdrawal symptoms, perceived distress from withdrawal symptoms, experience of withdrawal (intentional and unintentional discontinuation), decision for discontinuation, recommendation to others for the drug usage, experience with other medications, drug indications, problem with financial support, problem with social support, overall attitude toward medications, and recommendation to others.

The codes (themes) generated using deductive and inductive approach were refined by availability of patient expression for each theme, difficulty in differentiating between themes, and also relevance of the theme to measuring attitudes towards antidepressants. The final analytical framework includes 29 themes. **Table 4-3** lists the themes in the final analytical framework. To improve consistency of annotating (coding) using the analytical framework, the reviews were split to sentences. The detail of splitting drug reviews to sentences was explained in the methodology section.

Table 4-2 presents statistics on posts, sentences, and tokens. In average, patients on SNRI drugs had longer posts, 7.1 vs. 6.4 for number of sentence, and 117.5 vs. 103.3 for number of tokens. In average, patients on SNRI drugs had longer posts, 7.1 vs. 6.4 for sentence length, and 117.5 vs. 103.3 for tokens. However, the range of sentence number in SSRI drug is higher than SNRI drugs.

Table 0-2 Statistics on Frequency of Posts, Sentences, and Tokens

	Corpus	SSRI	SNRI
No. Posts	887	428	459
No. Sentences	6004	2749	3255
Avg. posts length (sentence)	6.77	6.42	7.1
No. Sentences (range)	1-35	1-35	1-25
No. Tokens	98186	44237	53949
Avg. Posts length (word)	110.7	103.36	117.53

Table 0-3 Frequency of Themes after Resolving Disagreements for the Whole Dataset

Themes (variables)	Corpus	SSRI	SNRI
ADRs	2004	985	1019
WD	279	64	215
EF	1078	575	503
INF	308	143	165
DI	806	422	384
PPI	280	121	159
PPI-P	55	26	29
PPI-N	87	34	53
KN	70	18	52
ADR-PD-low	404	193	211
ADR-PD-high	976	455	521
WD-PD-low	61	27	34
WD-PD-high	355	66	289
Decision for Discontinuation (DXD-dec)	49	17	32
Intentional Discontinuation (DXD-S)	212	103	109
Intentional Discontinuation-(DXD-W)	93	38	55
Unintentional discontinuation (DXD-F)	58	15	43
ATT-P (overall attitude towards drug-Positive)	58	34	24
ATT-N (overall attitude towards drug-Negative)	223	72	151
Suggestion to others (positive)	39	24	15
Suggestion to others (negative)	84	33	51
Dosage	277	151	126
Other patient	10	6	4
Lack of social support (SOS)	2	0	2
Lack of financial support (FS)	18	4	14
Other drugs	374	181	193
others	778	309	469

Table 4-3 shows the counts of annotated variables (themes) for the complete corpus and for both psychiatric categories of drugs separately. 33% of sentences included information about adverse drug reactions in the sample. 17% of the sentences provided information about drug

effectiveness, and only 5% of the sentence included information for patients' perceptions about healthcare providers. This information indicates that this data source is a rich source to evaluate pharmacological aspects of psychiatric medications, but not a rich source for patients-physician interaction. Comparing the frequency of sentences including negative recommendation about drug usage (84) to readers with frequency of sentences providing negative suggestion (39) indicates that patients with negative experiences with drugs are more likely to suggest others not to use the drugs. 12% of the sentences did not provide any informative information for the items (codes) defined in the final analytical framework.

To reduce the complexity of final analytical framework, themes that were directly related to patients' attitudes were eliminated from the Final analytical framework. Therefore, ATT-P, ATT-N, Suggestion to others (positive) and Suggestion to others (negative) were removed from the framework. In addition, themes "dosage" and "other patients" were removed from the final analytical framework because they do not provide any information for the patients' attitudes towards medications. Lack of social support (SOS) and Lack of financial support (FS) were also removed from the analytical framework because of high missing values for the themes.

Using the analytical framework, all the sentences were double coded by participation of four annotators with health-related background. The inter-annotator agreement (IAA) was calculated using Kappa. The Average IAA for the whole dataset was 0.75 with the lowest IAA for patient-physician interaction-positive attitude (0.5) and highest IAA for Drug indication (0.91).

4.3 Statistics on Summarized Dataset

Because annotation of comments were conducted at sentence level, an individual drug review maybe annotated several time for availability of a theme (code). To summarize annotation at

comment level, multiple expressions of a theme were reduced to a single one for a review posts. For example, several sentences in a drug review maybe annotated for the presence of ADRs, but at the level of a review (comment), we reduced it to the single presence of the theme (ADR). We summarized the whole dataset at the level of drug review using this strategy. All themes were considered as variables. **Table 4-4** shows the variables and their frequency. Limitations of each variable for measuring patients' attitude toward drugs and the strategy for filling missing values were discussed in the methodology section.

Table 0-4 Statistics on Variables Used for Testing Hypotheses

Variables	Frequency
ADR	
Presence	823
Absence	56
ADR-Perceived distress	
High	483
Low	166
Medium	230
Drug Effectiveness	
EF	524
EF-INF	120
INF	235
Patient-Physician Interaction	
Negative	47
Positive	62
Negative-positive	4
Missing value:	766
Complain of lack of knowledge	
Presence	60
Absence	819
Experience of Withdrawal (Intentional and/or Unintentional)	
No Experience	508
Experience	371
Unintentional withdrawal	
No report	831
Reported	48
Intentional Withdrawal	
No report	639
Existence of report on intentional withdrawal	240
Report on Intentional withdrawal and decision for withdrawal	
Existence of report	275
No-report	604

4.4 Testing Hypotheses

4.4.1 Measuring association between variables and levels of attitude

To measure association between variables and levels of attitude, Chi-Square has been used for categorical variables and Anova was used for continuous variables. **Table 4-5** lists the variables, related hypothesis, the type statistical test used for testing hypothesis, and P-value. For this study, P-value was set on 0.05. All the analysis was conducted using RStudio, version 1.0.153.

We reject the null hypotheses for variables ADR experience, ADR-Perceived distress, Drug Effectiveness, Complaint of lack of knowledge (LK), Duration of usage, Experience of Withdrawal (Intentional and/or Unintentional), and Intentional Withdrawal (weaning off and stopping the medications). In fact, these variables are significantly associated with patients attitude towards antidepressants.

For two variables age and gender, there is not enough evidence to reject the null hypothesis. So we conclude that these two variables are not significantly associated with attitudes towards antidepressants specified in this study.

Table 0-5 List of Variables, Hypothesis, Statistics Test, and computed P-values

Variables	Hypothesis	Statistic method	Values	P-value
ADR Experience	ADR and attitude are independent variables.	Chi-square	X-squared = 31.11	2.907e-06*
ADR-Perceived distress	ADR-Perceived distress and attitude are independent variables.	Chi-square	X-squared = 231.6	< 2.2e-16*
Drug Effectiveness	Drug Effectiveness and attitude are independent variables	Chi-square	X-squared = 548.52	< 2.2e-16*
Complaint of lack of knowledge (LK)	LK and attitude are independent variables.	Chi-square	X-squared = 59.358	3.957e-12*
Experience of Withdrawal (Intentional and/or Unintentional)	WD-Experience and attitude are independent variable.	Chi-square	X-squared = 55.624	2.404e-11*
Intentional Withdrawal	Intentional Withdrawal and variable are	Chi-square	X-squared = 64.009	4.161e-13*

	independent variable.			
Gender	Gender and attitude are independent variables.	Chi-square	X-squared = 2.6812	0.6125
Age	Age and attitude are independent variables.	ANOVA	F-value= 0.7183	0.3969
Duration of usage	Duration of usage and attitude are independent variables.	ANOVA	F-value= 43.665	6.756e-11*

Figure 4-1, Figure 4-3, Figure 4-4, Figure 4-5, Figure 4-7, Figure 4-9, Figure 4-11, Figure 4-13 shows distribution of “gender”, “age”, “duration”, “experience of ADRs”, “perceived distress from ADRs”, “drug effectiveness”, “withdrawal experience”, “complaint of lack of knowledge” for each level of patients attitude respectively.

Figure 4-2, Figure 4-6, Figure 4-8, Figure 4-10, Figure 4-12, Figure 4-14 shows frequency of “gender”, “experience of ADR”, “perceived distress form ADRs”, “drug effectiveness”, “withdrawal experience”, and “complaints for lack of knowledge” for each level of attitude.

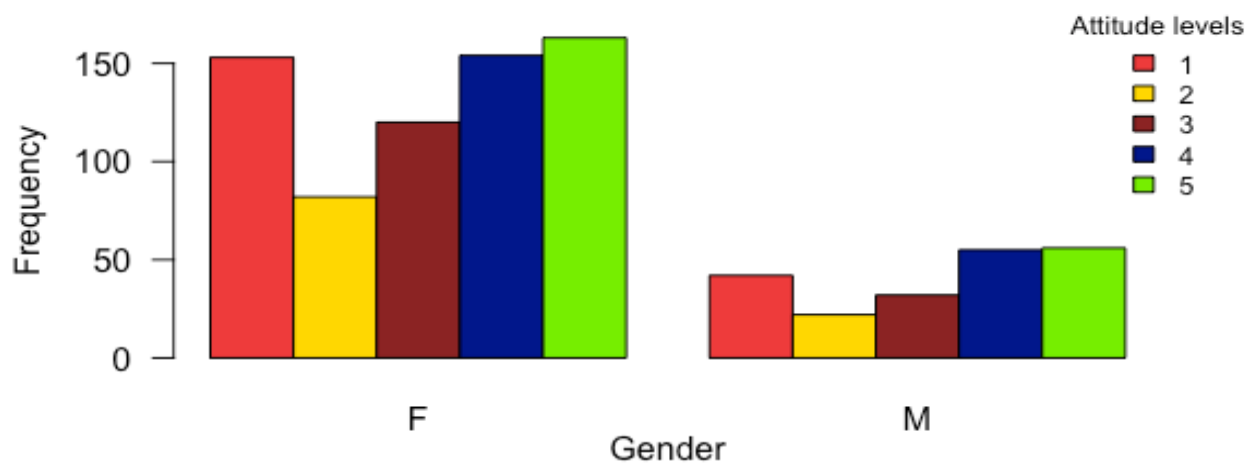


Figure 0-1 Distribution of gender for each level of patients attitude

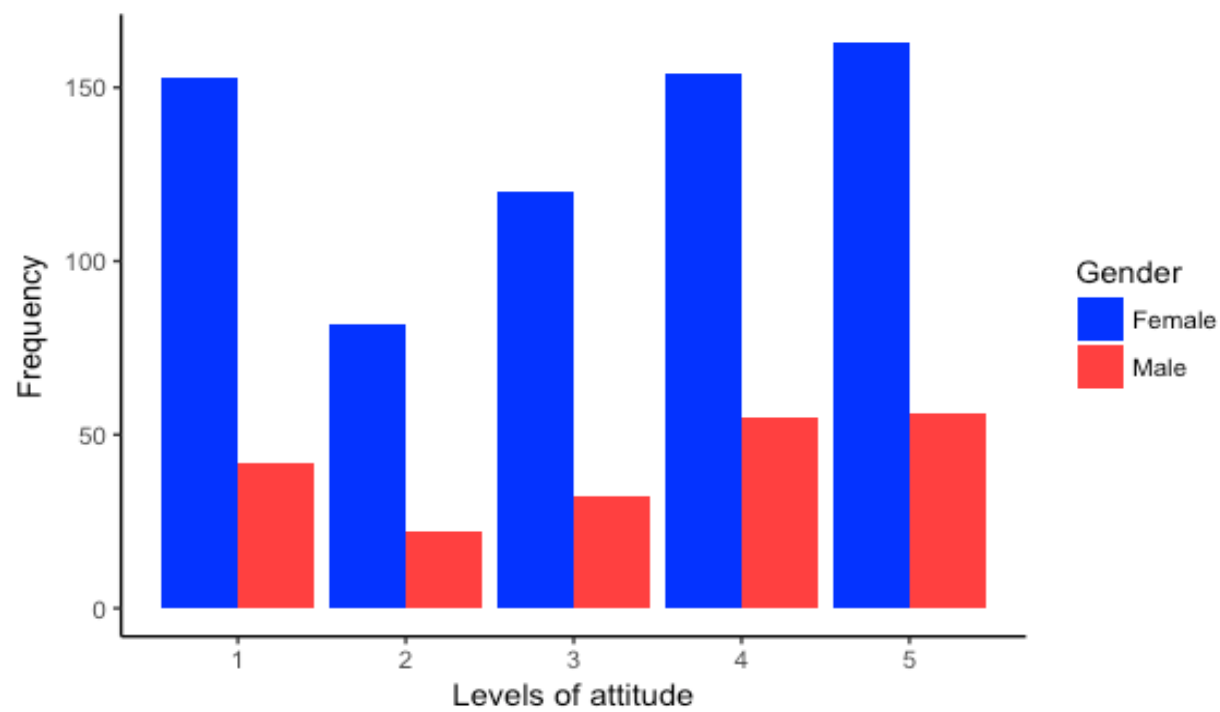


Figure 0-2 Frequency of gender for each level of attitude

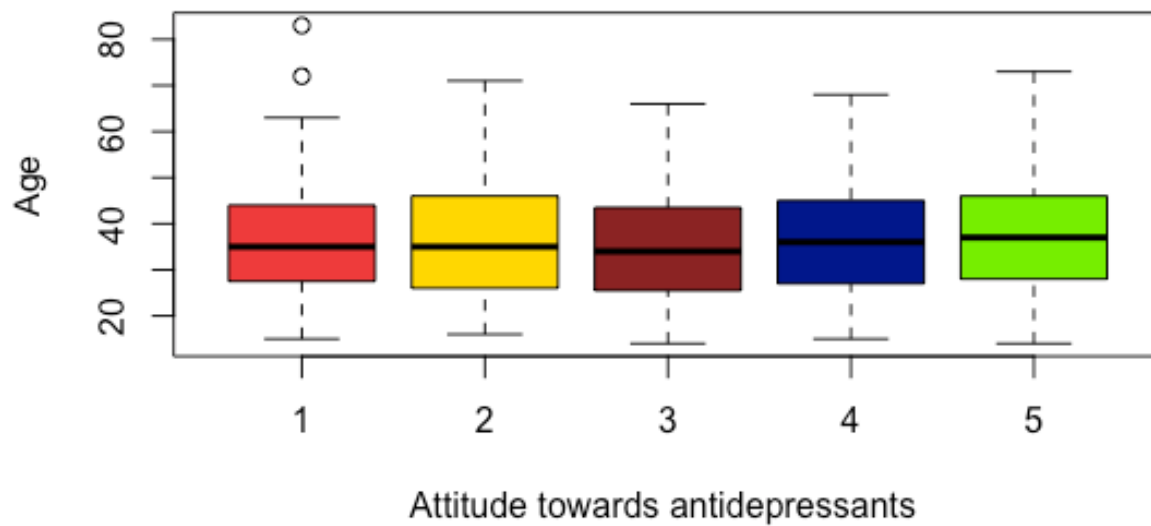


Figure 0-3 Distribution of age for each level of attitude

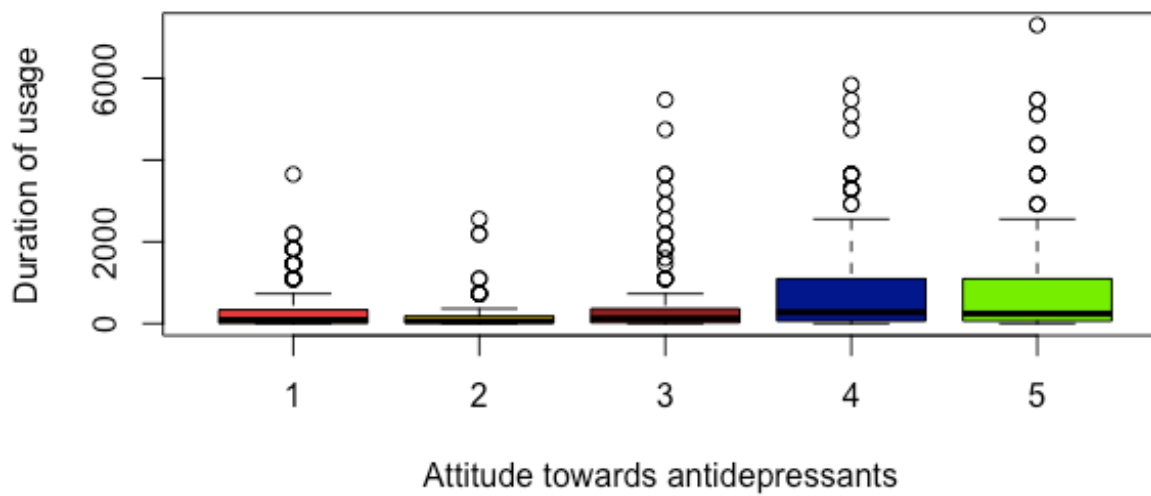


Figure 0-4 Distribution of duration of usage for each level of attitude

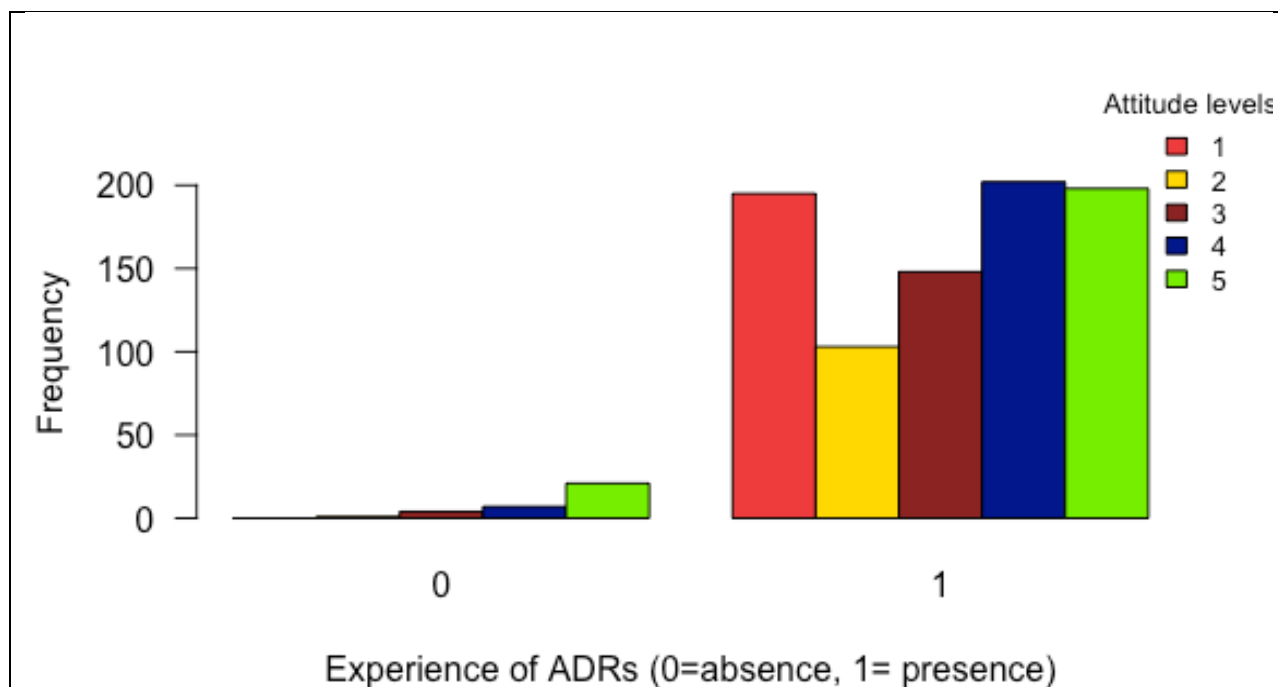


Figure 0-5 Distribution of presence and absence of ADRs for attitude levels

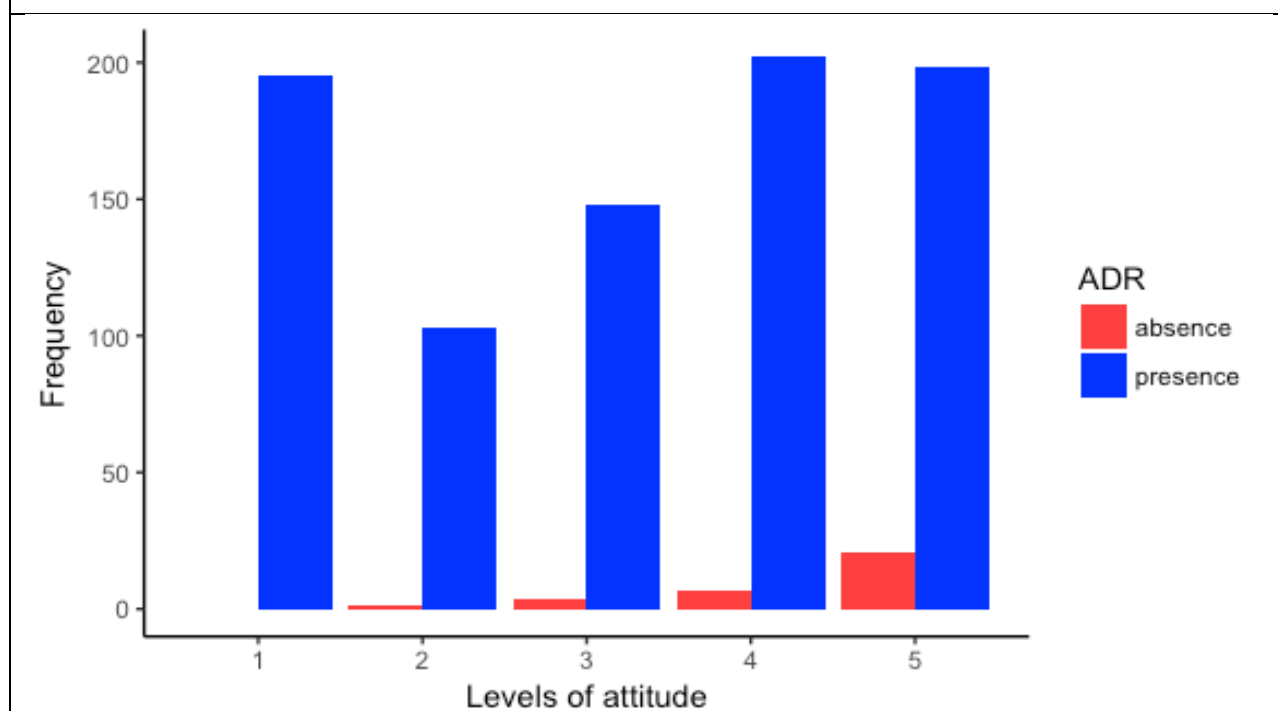


Figure 0-6 Frequency of ADR for each level of attitude

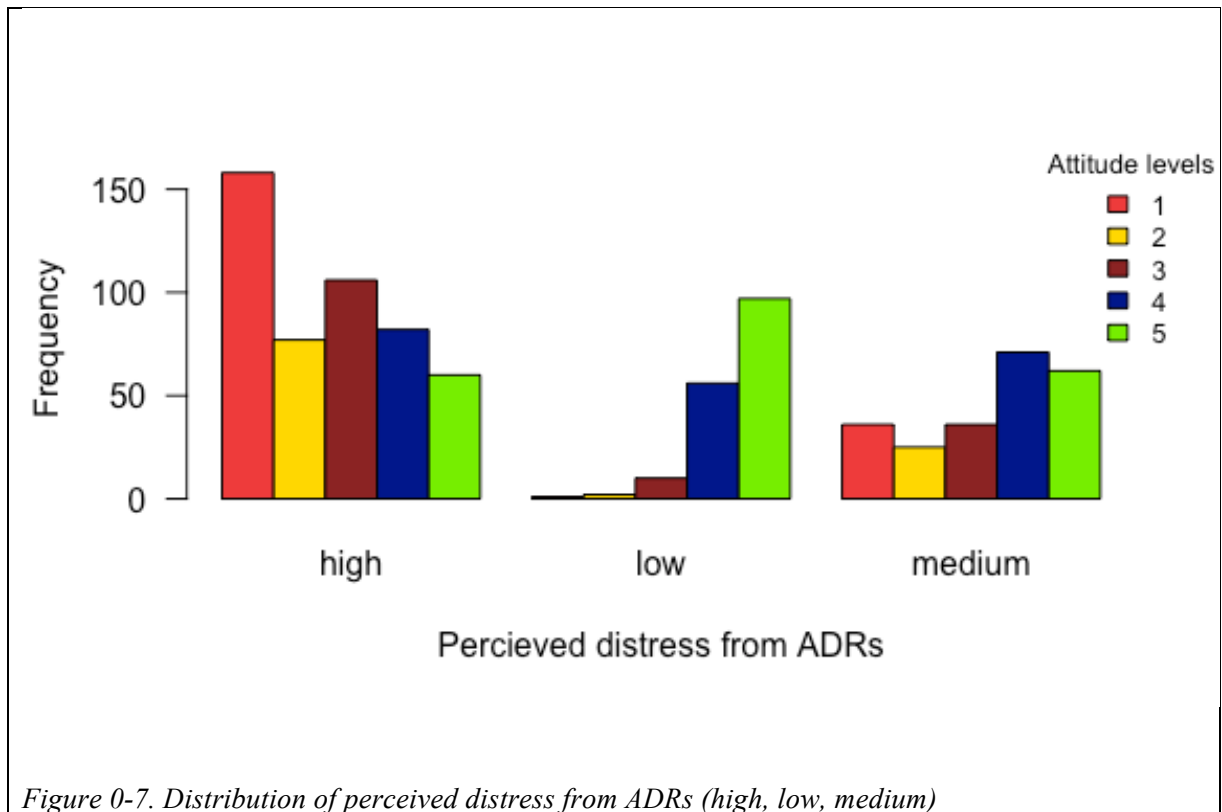


Figure 0-7. Distribution of perceived distress from ADRs (high, low, medium)

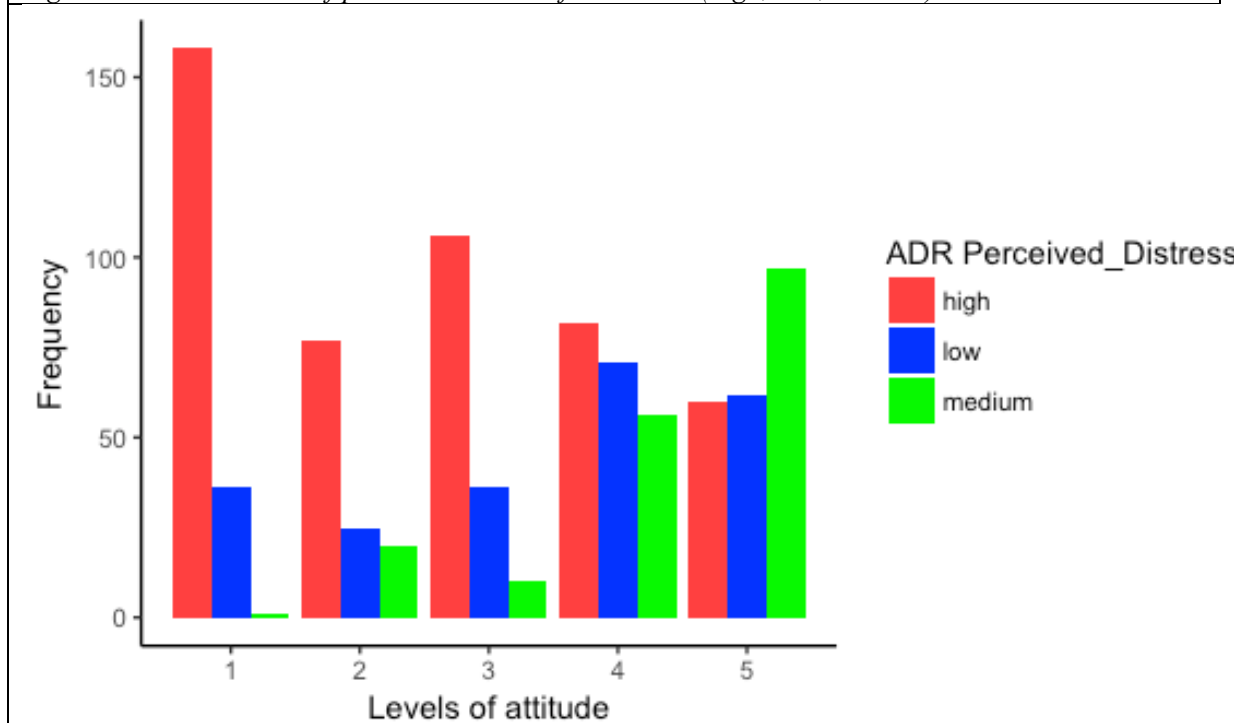


figure 0-8 Frequency of perceived distress from ADRs

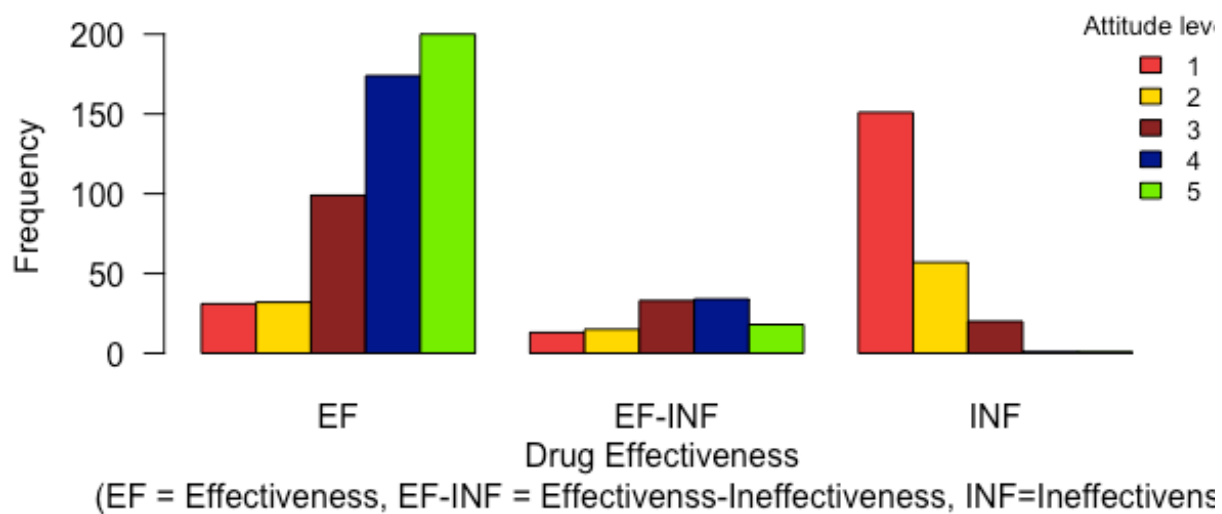


Figure 0-9 Distribution of drug effectiveness

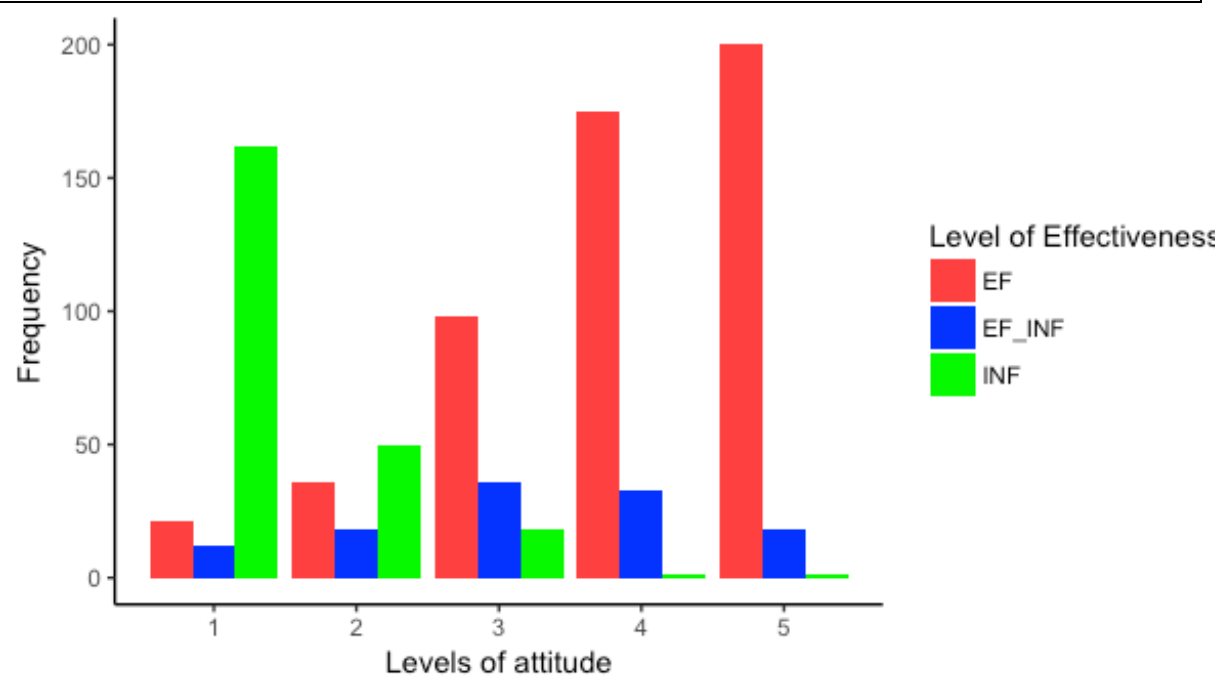


Figure 0-10 Frequency of drug effectiveness

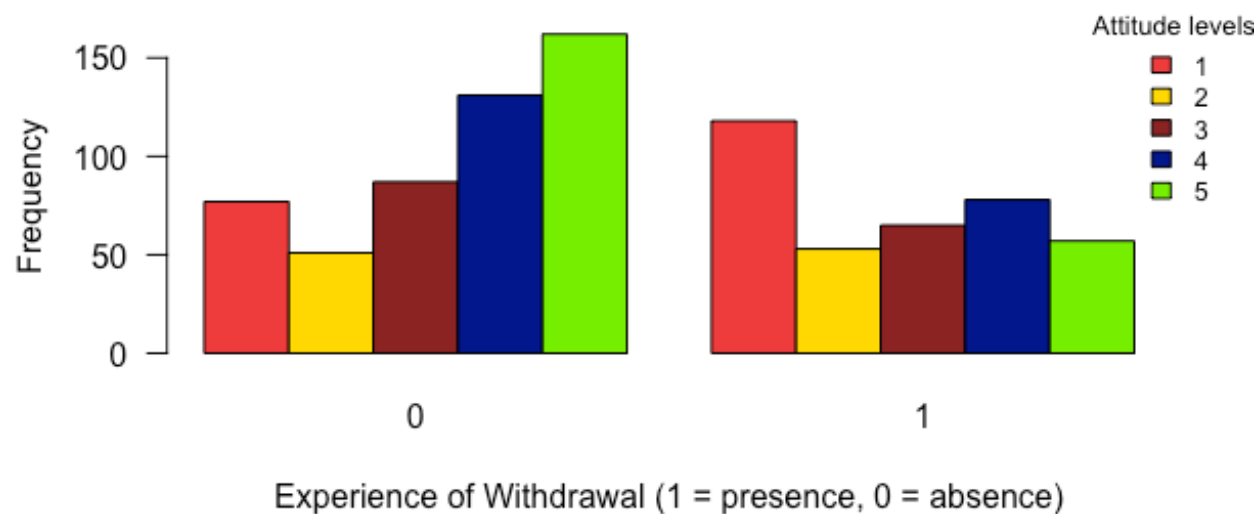


Figure 0-11 Distribution of withdrawal experience for levels of attitude

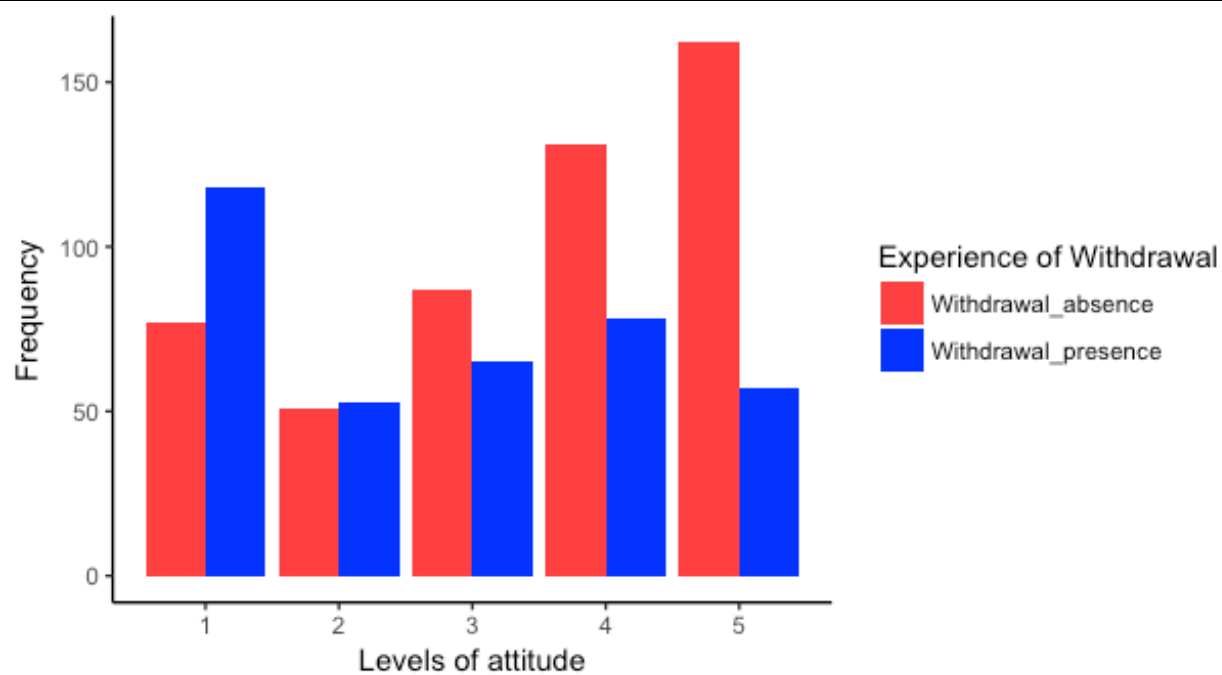


Figure 0-12 Frequency of withdrawal experience for each level of attitude

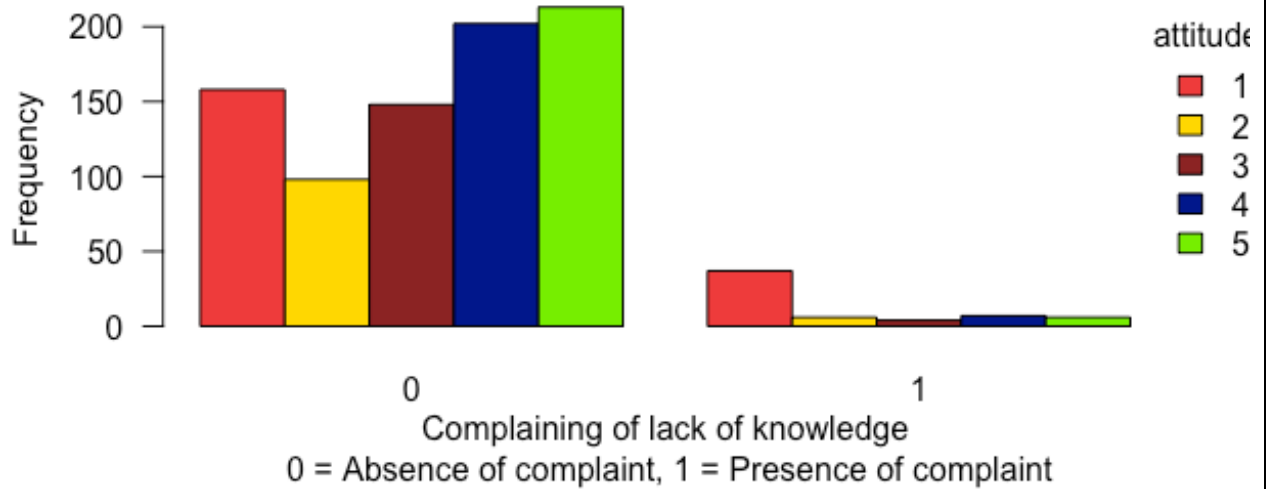


Figure 0-13 Distribution of complaints for lack of knowledge

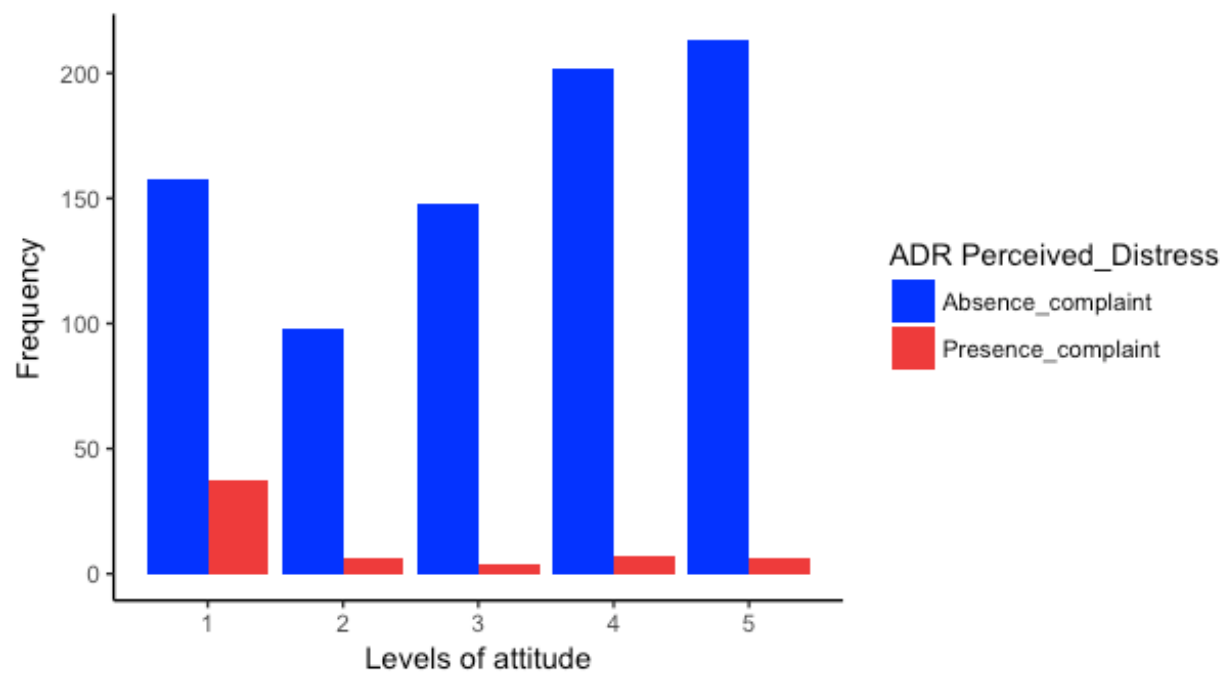


Figure 0-14 Frequency of complaints for lack of knowledge for each level of attitude

4.4.2 Developing a predictive model

To find the best fitting model to describe the relationship between the independent variables (outcome variables) and attitude (dependent variable), we used ordinal logistic regression. We used polr command from the MASS package in R to estimate an ordered logistic regression model. The command name comes from proportional odds logistic regression, highlighting the proportional odds assumption in the model. “Polr” uses the standard formula interface in R for specifying a regression model with outcome followed by predictors.

For developing the model we did not include age and gender, because they are not significantly associated with patients attitude toward antidepressants. The following equation shows the relationship between independent and dependent variables:

$$\text{Attitude} \sim \text{Experience of ADR} + \text{ADR_ Perceived distress} + \text{Effectiveness} + \text{Experience of WD} + \text{duration of usage} + \text{lack of knowledge}$$

Table 4-6 shows the coefficient, standard error, and p-value for the outcome variables for this model.

Table 0-6 Values of statistical model

Variable	Co-efficient	Std. Error	P-value
ADR_ Experienced (1)	-0.511989943	1.170552e-01	1.220362e-05
ADRPD (low)	1.938705530	1.876554e-01	5.091863e-25
ADRPD (medium)	0.804059699	1.584470e-01	3.882502e-07
EFINF (EF-INF)	-0.870736633	1.945260e-01	7.598449e-06
EFINF (INF)	-3.979180295	2.123201e-01	2.274740e-78
ExperienceWD (1)	-0.700315050	1.382074e-01	4.038538e-07
Lack of knowledge 1	-0.433608481	3.215736e-01	1.775311e-01
Duration	0.000259638	8.441442e-05	2.099712e-03

Table 4-7 shows the confidence interval for the outcome variables in the model.

Table 0-7 Confidence Interval for the Outcome Variables in the Model

Outcome variables	Minimum	Maximum
ADR Experienced (1)	NA	NA
ADRPD (low)	1.5397784239	2.34744346
ADRPD (medium)	0.4932958996	1.1160166506
EFIN (FEF-INF)	-1.2550505954	-0.4878371728
EFINF (INF)	-4.4199855660	-3.5554091696
Experience WD (1)	-0.9723098874	-0.4297849025
Lack of knowledge 1	-1.0830989795	0.1830980650
Duration	0.0001209877	0.0004024088

Analysis of coefficients (**Table 4-7**) showed that “drug effectiveness” and “perceived distress from ADRs” are the most significant factors affecting patients’ attitudes toward antidepressants.

Comparing our findings with findings of literature showed that in line with literature, “drug ineffectiveness”, “experience of ADR”, and “lack of knowledge” and “duration of treatment (usage)” are associated with negative attitude towards antidepressants. Demyttenaere et al. (2004) and (Reilly JL, 2011) showed drug effectiveness is strongly associated with patient attitude towards antidepressants. For “experience of ADR” several studies found that experience of ADR is associated with negative attitude towards medications (Dougherty et al., 2009; Murata et al., 2012; Ng et al., 2012). (Haslam et al., 2005) and (Gabriel & Violato, 2010) found that patients with more knowledge about their illness and their treatment are likely to have positive perception about treatment and, in turn, to be more adherent. Association between variable “patient-physician interaction” and attitude were not tested, because of low IAA for this variable and high rate of missing values.

As the results in table 4-5 shows, age and gender are not associated with levels of attitude. Therefore, we did not include them in the predictive model. The findings are in agreement with findings of studies conducted by Jacob et al. (2015), (Murata et al., 2012), and

(Ng et al., 2012). However, our findings for age and gender are contrary to the findings of study conducted by Prins et al. (2008).

We found “Experience of withdrawal” associated with negative attitude towards antidepressants. However, this variable was not discussed by literature.

4.5 Statistics on Entities Identification

In the second phase of the study, ADRs, WDs, and DIs were identified and extracted from sentences that in the first phase of this study were annotated as ADRs, WDs, and DIs. All the entities were identified by strictly following guidelines covering the rules of identification, such as patients’ certainty in linking ADRs and WDs to the drugs. In addition to identifying the physiological, psychological, and cognitive patient complaints about the drugs, the impact of the ADRs on patients’ daily functioning and social participation were also extracted. Identifying functional problems is significant for understanding the impact of ADRs on patients’ quality of life. It is also important for estimating the indirect cost associated with the ADRs (Giannangelo et al., 2005). Collecting this information enhances clinicians' abilities to predict impact of ADRs on patients’ functionality and work performance. Hence, they will be able to design more effective interventions targeting patients' attitude and adherence towards medications.

Four annotators (coders) identified the boundary (span) of the ADRs/WDs/DIs in the sentences. All the entities were double coded by strictly following the guidelines of entities identification. The computed pairwise agreement for strict match was 0.86 for ADRs, 0.81 for WDs and 0.91 for SSs, with the average 0.86 for the whole corpus. The reason for applying pairwise agreement (PA) for computing IAA rather than conventional measures, such as Kappa, is that PA is calculated at the level of entities. Since the identification task requires identifying the entities and determining their correct boundaries, the chance agreement is effectively zero.

Table 4-8 lists frequencies of identified entities for the total corpus, for each class of drugs, and for each type of entity separately. From 6,534 identified entities, 73% are ADRs, 12% are WDs, and 18% are drug indicators. In total, 41% of the entities were duplicates, with the highest frequency of duplicates for drug indicators (89%), and lowest frequency for withdrawal symptoms (20%), indicating that patients mostly use the same medical terms to describe the reasons for drugs prescription. This result may be due to similar information **given** by providers to patients. Physiological entities constitute 59% of the total entities, followed by psychological entities (35%), Cognitive entities (5%) and Functional problems (2%), showing that the review posts are not a rich of the impact of ADRs or WDs on patients' quality of life. For the purpose of designing more effective medication adherence interventions, it would be useful if healthcare forums also ask patients to report the impact of drugs on their daily functioning and social activities. For ADRs and WDs, psychological and cognitive expressions have higher variability than physiological expressions. That is due to level of subjectivity of these types of entities leading to creating various phrases by patients to express their subjective complaints. For two classes of psychiatric medications, SSRI and SNRI, the distribution of ADRs and drug indicators is almost similar.

Table 0-8 Statistics on Entity Identification Component for Each Class of Drugs and Type of Entities Separately

	Total		Physiological		Psychological		Cognitive		Functional	
	Entities	Unique	All	Unique	All	Unique	All	Unique	All	Unique
Corpus-entities	6,534	59% (All) (3856)	3,834	63% (All) 2422	2266	51% (All) 1165	297	79% (All) (236)	136	131 (All) 96%
ADR Corpus	4774	66% (All) 3176	3380	62% (All) 2098	1061	76% (All) (804)	247	77% (All) (191)	85	96% (All) (82)
ADR-SSRI	2236	1374	1553	880	521	372	119	81	42	40
ADR-SNRI	2537	1802	1826	1218	540	432	128	110	43	42
WD-Corpus	592	79% (All), 470	355	74% (All) 263	180	83% (All) (151)	26	100% (All) (26)	31	97%(All) 30
WD-SSRI	129	91	72	45	51	40	5	5	1	1
WD-SNRI	463	379	283	218	129	111	21	21	30	29
SS-Corpus	1168	17% (All) 210	99	61% (All) 61	1025	20% (All) 210	24	79% (All) 19	20	95%(All) 19
SS-SSRI	621	153	30	20	568	114	14	10	9	8
SS-SNRI	547	157	69	41	457	96	10	9	11	11

Table 4-9 provides examples of the top 5 identified entities for each type of entities and class of drugs separately. Overall, ADRs and WDs can lead to debilitating functional problems, such as restricted work performance, loss of job, poor educational performance, emergency visit, and hospitalization.

Table 0-9 The Top Five Identified Entities for Each Class of Drugs and Type of Entities Separately

	Physiological –top (5)	Psychological- top (5)	Cognitive- top (5)	Functional - top (5)
SSRI - ADRs	Weight gain (55), nausea (38), dry mouth (30), insomnia (30), fatigue (25)	vivid dreams(19), nightmares (9), increased anxiety (9), suicidal thoughts(9), anxiety (9)	Memory loss(6), brain fog (4), inability to concentrate (3), hard to focus (4), forgetfulness (2)	Called in sick(2), impossible to do my job (1), struggling to learn new things (1),unable to function(1),struggle just to comb my hair(1)
SNRI- ADRs	Insomnia(47), constipation(37), weight gain(34), nausea(32), dry mouth(28)	vivid dreams(19), anxiety (14), nightmares (13), no sex drive (12), loss of sex drive(8)	Memory loss (6), confusion(4), spacey(3), inability to concentrate(3), couldn't focus on anything (2)	Hospitalized (2), unable to function (1), loss of friends(1) unable to work(1), cannot drive (1)
SSRI- Withdrawal symptoms	Dizziness (7), upset stomach (3), brain zap (2), nausea (2), crying spell (1)	Irritability (4), suicidal thoughts (2), aggression (1), crying spells (1), very anxious (1)	Slight confusion (1) Severe mental confusion (1), Couldn't concentrate (1), Memory sucks (1)	Ended up in the er (1), 1/2 weeks in psych ward (1)
SNRI Withdrawal symptoms	Dizziness(16), nausea (11), brain zaps (8), headaches (6), dizzy (6)	Mood swings (5), nightmares (3), moody (2), feel like a walking zombie (2), aggressive (2)	Dissociative episodes (1), foggy (1), lack of concentration (1), confused (1), cannot think (1), feel spaced out(1)	can't function (1), could not drive (1), Difficulty tidying house (1), Difficulty performing shopping activities (1), Difficulty performing educational activities(1)
SSRI- Drug Indicators	Insomnia (4), upset (2), impatient (1), fibromyalgia (1), physical pain (1)	Depression (163), anxiety (127), anxiety (32), depressed (15), panic attacks (14)	intrusive thoughts (3), obsessive thoughts (2), confused (1), racing thoughts (1), inner critic (1)	lost my job (2), isolating (1), barely functional (1), dysfunctional (1)
SNRI- Drug Indicators	Pain (12), fatigue (6), insomnia (2), fibromyalgia (4), migraines(2)	Depression (185), anxiety (103), anxious (9) suicidal thoughts (6),social anxiety (6)	obsessive thinking (4) Mental clutter(2), memory loss (1), rumination (2), memory loss (1)	Loss everyone in my life (1), could not function at work (1), lost everyone in my life (1), harsh edge to my inter-action with others

4.6 Statistics on Terminology Association

In the third phase of this study, the identified entities were normalized by mapping to both the UMLS Metathesaurus and SNOMED-CT. This mapping benefits the generation and testing of medical hypotheses related to associations between ADRs, WDs and patients attitudes and discontinuation behavior of patients by providing unambiguous and standard information for statistical data collection and analysis.

To improve accuracy and consistency of mapping, guidelines was developed using clinical trials addressing the ADRs of the drugs specified in this study and qualitative studies investigating the themes of patient complaints about the drugs. Each ADR concept is associated with a set of attributes reflecting patients' problems. For example, "executive dysfunction" is a term taken from the clinical trials literature and mapped to the Metathesaurus and SNOMED-CT. The concept of executive dysfunction as a cognitive ADR is associated with inability to initiate and follow processes of completing a task, so a patient complaint of "cannot follow through on simple tasks, can be mapped to "executive dysfunction" as a general concept. The guidelines also includes the requirements for selecting proper/preferred UMLS and SNOMEDCT concept, and flowchart for mapping to both UMLS and SNOMED-CT.

Table 4-10 shows statistics for normalization component. The final set contains 698 UMLS concepts for ADRs, showing that out of 3176 unique identified ADRs and WDs, only 14% of them are unique standard medical concepts. On average, for each standard ADR concept, there is 4.5 layperson expression of ADR/WD, reflecting the challenge of automatic identification of ADR/WD using standard medical lexicons. From 210 unique identified drug indicators, 81% are unique standard concepts, showing that patients mostly use the diagnosis results provided by healthcare professionals to report the reason for drug prescription.

Overall, all three types of entities (ADRs, WDs, and Drug Indicators) were mapped to 811 UMLS concepts, from Which 154 concepts did not include SNOMED expressions. Therefore, we attempted to use equivalent medical concept that semantically match the primary concepts and include expressions of SNOMED-CT concepts. **Table 4-11** lists the most frequent UMLS concepts in the corpus that did not include SNOMED-CT concepts.

This study also identified qualifiers representing severity (QS) and persistency (QP) of the entities. Identifying the qualifiers can help healthcare providers to estimate the debilitating effects of ADRs/WDS on patient quality of life and whether they need to use any specific interventions to improve patient adherence to medication.

Table 0-10 Statistics on Entities Normalized with UMLS and SNOMED-CT Concepts

Entities	Physiological	Psychological	Cognitive	Functional	
ADRs	698	462	197	42	31
The five most frequent ADRs after normalization	Sleeplessness (171), nausea (169), weight gain (148), lack of libido (138), headache (106)	Anxiety (92), detailed recall of dream (62), depressed mood (41), apathy(38), feeling suicidal(38)	Foggy feeling in head (44), unable to concentrate (30), amnesia (19), memory impairment (15), forgetful (14)	Difficulty in daily functioning (10), emergency room admission (9), social withdrawal (8), hospitalization (6), bed-ridden (5)	
WD	218	109	72	18	19
The five most frequent WDs after normalization	Dizziness (43), nausea (32), headache (27), malaise (10),	Irritable mood (16), depressed mood (11), mood swings (10), nightmares (8), severe anxiety (6)	Confusion (4), unable to concentrate (3), mental suffering (3), Foggy feeling in head (2), actual low self control (1)	Difficulty in daily functioning (4), bed-ridden (3), restricted work performance (3), difficulty driving a car (2), emergency room admission (2).	
Drug indicators	171	46	103	9	13
The five most frequent WDs after normalization	Pain (13), fibromyalgia (9), sleeplessness (14), fatigue (6), tired (4)	Depressed mood (444), anxiety (258), panic attacks (27), feeling suicidal (22), social fear (17)	Obsessive thoughts (13), unable to think clearly (2)	Difficulty in daily functioning (4), difficulty maintaining relationships (2), Loss of job (3)	

Table 0-11 The Most Frequent UMLS Concepts in the Corpus without SNOMED Expressions

UMLS-Primary Concept	SNOMED -CT Primary Concept	Frequency	UMLS-Equivalent Concept	SNOMED-CT Concept
C0859330 / Foggy feeling in head / Finding	No Code	46	C0683369 / Clouded consciousness / Sign or Symptom	Clouded consciousness (finding)
C0857507 / Spaced out / Finding	No Code	34	C0349446 / Dissociative trance / Mental or Behavioral Dysfunction	Dissociative trance (disorder)
C0392703/ Shakes/ Finding	No Code	33	C0040822 / Tremor / Sign or Symptom	Tremor (finding)
C0549209 / Feeling jittery / Sign or Symptom	No Code	29	C0849963 / Feeling nervous / Sign or Symptom	Feeling nervous (finding)
C0857486 / Felt like a zombie / Finding	No Code	24	C0233484 / Emotionally detached / Finding	Emotionally detached (finding)

T

Table 4-12 shows statistics on the qualifiers indicating intensity and persistency of the ADRs and WDs. Identifying the qualifiers can help healthcare providers to estimate the debilitating effects of ADRs/WDS on patient quality of life and whether they need to use any specific interventions to improve patient adherence to medication.

Table 0-12 Frequency of Identified Qualifiers Associated with ADRs and WDs

Category	Frequency	Example
Mild	112	Slight (27), mild (22), a little (14), slightly (10), minor(5)
Moderate	67	Moderate (6), Somewhat (3), semi (2), possibly some (1)
Severe	432	Extreme (52), severe (40), extremely (17), horrible (17), terrible (16)
Persistent	243	Constant (16), always (13), constantly (11), chronic (7), chronic (6)
Not-persistent	317	at first(22), occasional(12), in the beginning (12), sometimes(11), initial(7)

4.7 Testing association between specific type of ADRs and levels of Attitude

After normalizing the ADRs, we tested the association between 21 physiological ADRs with levels of attitude. The 21 physiological ADRs that were specified in the Antidepressant Side-Effect Checklist (ASEC) are dry mouth, drowsiness, insomnia (difficulty sleeping), blurred vision, headache, constipation, diarrhea, increased appetite, decreased appetite, nausea, problems

with urination, problems with sexual function, palpitations, feeling light-headed on standing, feeling like the room is spinning, sweating, increased body temperature, tremor, disorientation, yawning, weight gain (**Table 4-12**). For each ADR mentioned in the ASEC, the semantically related UMLS Metathesaurus concepts were grouped and the association between that group of ADRs with attitude were tested. In total, 128 unique UMLS concepts for ADRs (out of 698 concepts), which were identified in the phase of entity normalization, were grouped with ADRs specified in the ASEC questioner. Association between these ADRs and attitude was evaluated using X-square. For ADRs that the frequency for a level of attitude in the contingency table was less than 5, Fisher-exact test has been used. The result of the study showed that weight gain, yawning, disorientation, palpitation, increased appetite, and dry mouth are associated with patients' attitude towards antidepressants. While other ADRs in the ASEC are not associated with patients' attitudes toward antidepressants.

Table 0-13 Testing Association between ADRs and Attitude

Variables		Statistic test	P-value for X-Square	P-Value for Fisher Exact Test
Physiological ADRs specified in the ASEC Questionnaire	Headache	8.4545	0.07628	
	Insomnia	4.007	0.4051	
	Drowsiness	2.7613	0.5985	
	Blurred. Vision	6.6146		0.1577
	Dry mouth	11.801	0.01889*	
	Constipation	1.9675	0.7417	
	Diarrhea	7.1359	0.1289	
	Decreased Appetite	3.6637	0.4534	
	Increased. Appetite	19.862	0.0005318 *	
	Palpitation	9.2751	0.05458	0.06515
	Sweating	5.4177	0.2471	
	Disorientation	10.995	0.02662	0.02323
	Yawning	19.116	0.0007457	0.002313
	Increased body temperature	4.4946	0.3432	0.4051
	Weight gain	10.488	0.03296	

	Tremor	4.8804,	0.2998	
	Problems with urination	3.2561	0.5159	0.3724
	Problem with sexual functioning	23.689	9.218e-05	
	Nausea or vomiting	0.2873	0.9906	
	Vertigo	3.3951	0.494	0.5618
	Light headed	6.5986,	0.1587	
Psychosocial ADRs specified in the literature	Emotional indifference	21.131	0.0002983	
	Apathy	20.332	0.0004293	2.224e-05
	Emotionally detached	24.477	6.407e-05	
	Mood Swing	X-squared = 11.913	0.01801	
	Anxiety	23.346	0.000108	
	Problem with Motivation	32.532	1.489e-06	
Cognitive ADRs specified in the literature	Difficulty in concentrating	18.661	0.0009162	0.0004507
	Memory problem	25.283,	4.413e-05	4.413e-05

The result of studies conducted by De las Cuevas et al. (2014) using ASEC questionnaire also showed that “weight gain” and “dry mouth” associated with patient adherence toward medications.

In addition to the ADRs specified in the ASEC questionnaire, we also tested the association between psychological ADRs including “emotional indifference”, “apathy”, “emotionally detached”, “mood swing”, “anxiety”, and “motivation” and cognitive ADRs including “Difficulty in concentrating” and “memory problem”. Overall 36 unique UMLS concepts for ADRs that semantically related to the psychological and cognitive ADRs were grouped and tested for the association with attitude. All the psychological and cognitive ADRs were strongly associated with patients’ negative attitude toward antidepressants (**Table 4-13**). The frequency of ADRs (specified in table 3) for levels of attitude (1-5) depicted in from Figure 4-15 to Figure 4-39.

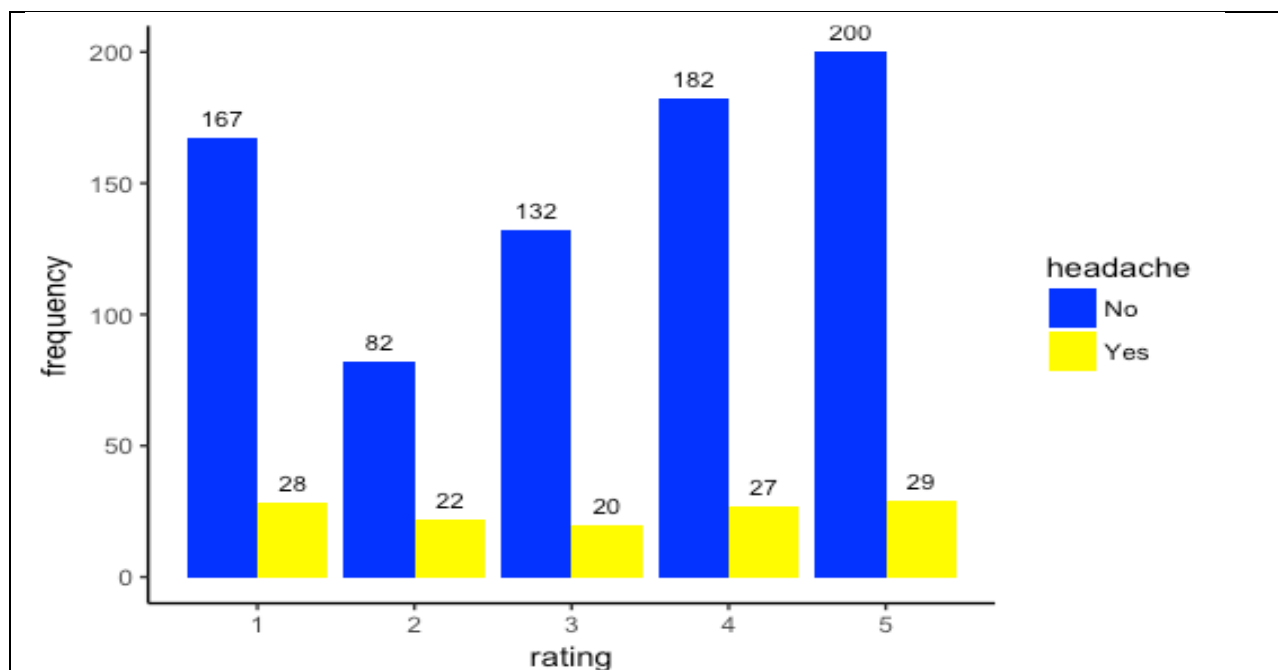


Figure 0-15 Frequency of headache for each level of attitude

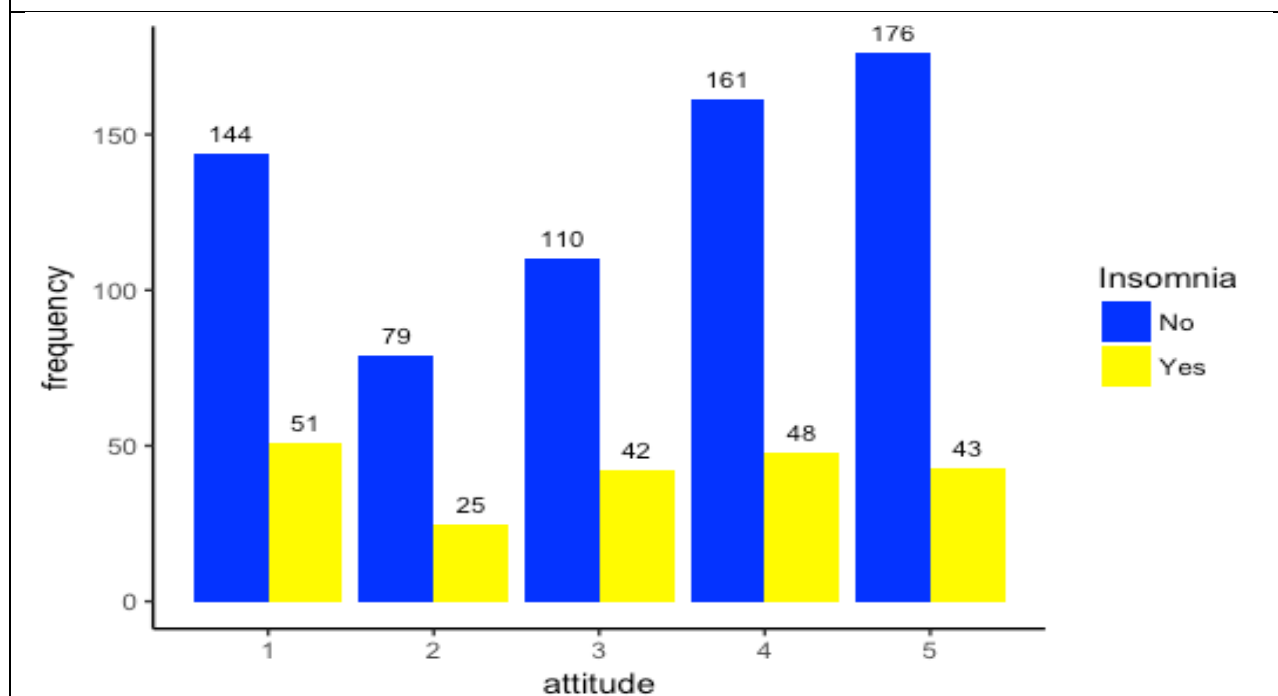


Figure 0-16 Frequency of insomnia for each level of attitude

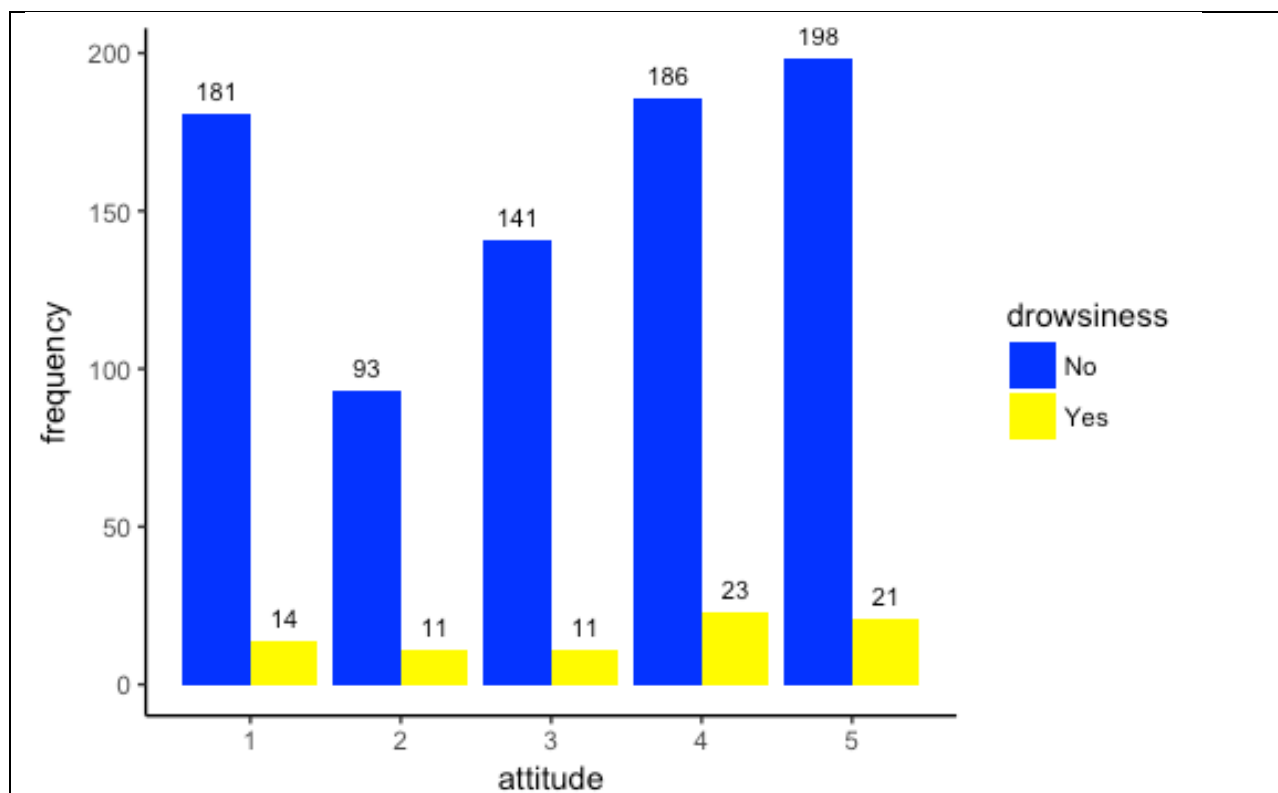


Figure 0-17 Frequency of drowsiness for each level of attitude

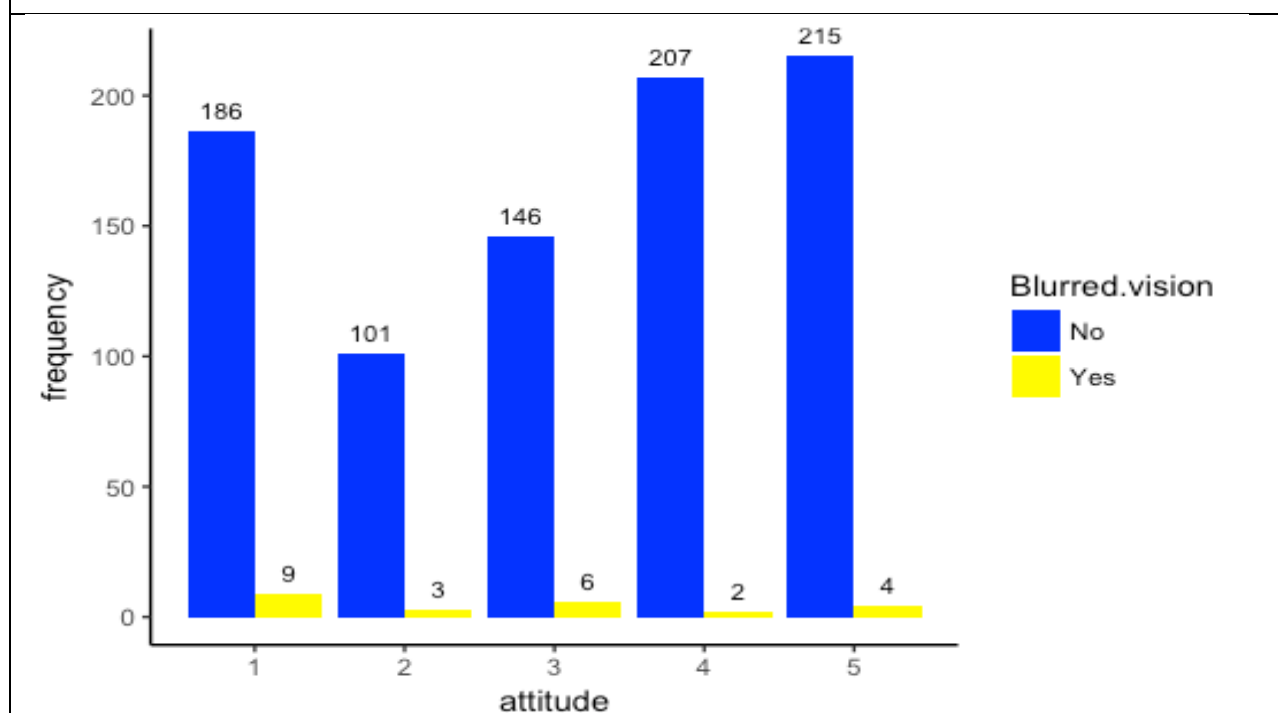


Figure 0-18 Frequency of blurred vision for each level of attitude

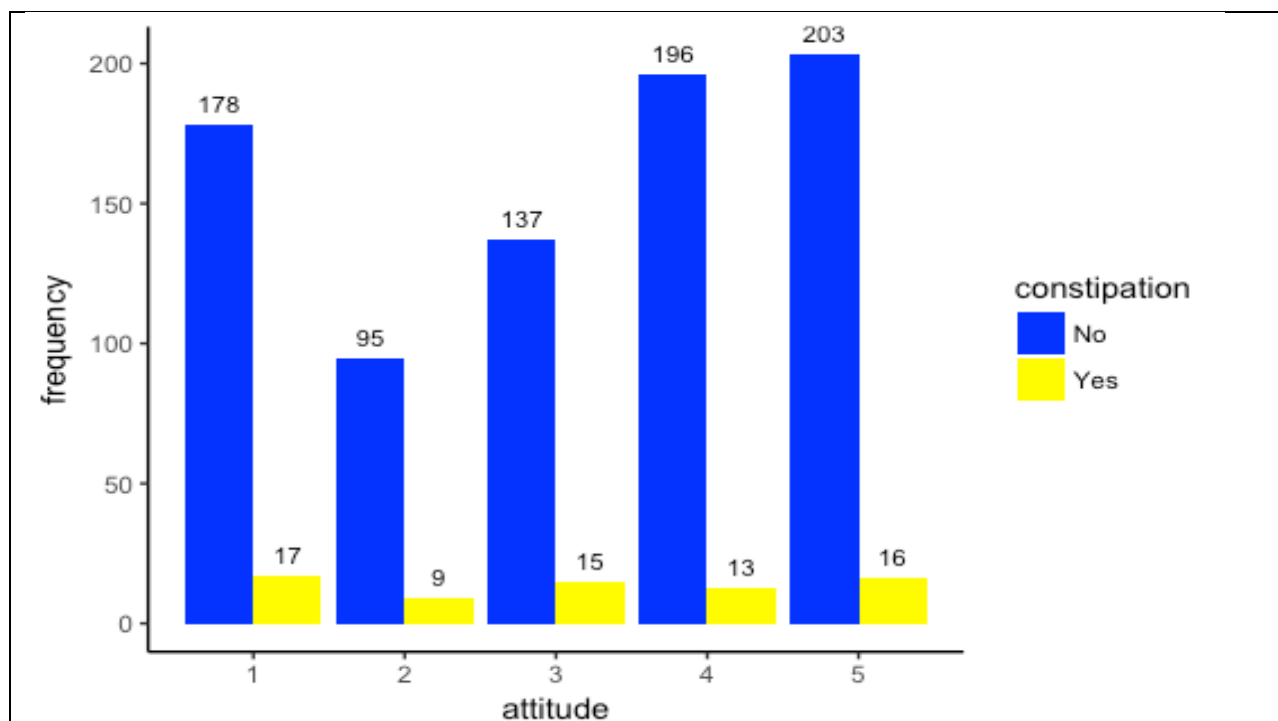


Figure 0-19 Frequency of constipation for each level of attitude

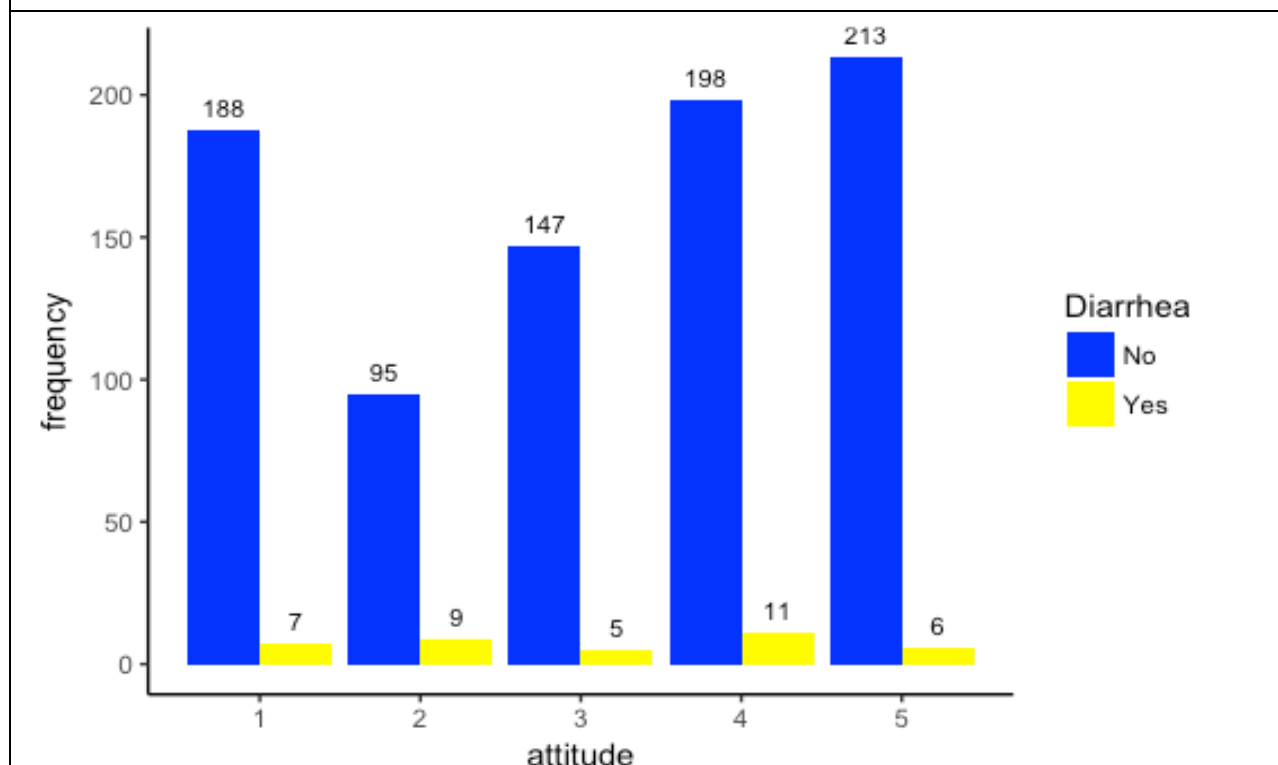


Figure 0-20 Frequency of diarrhea for each level of attitude

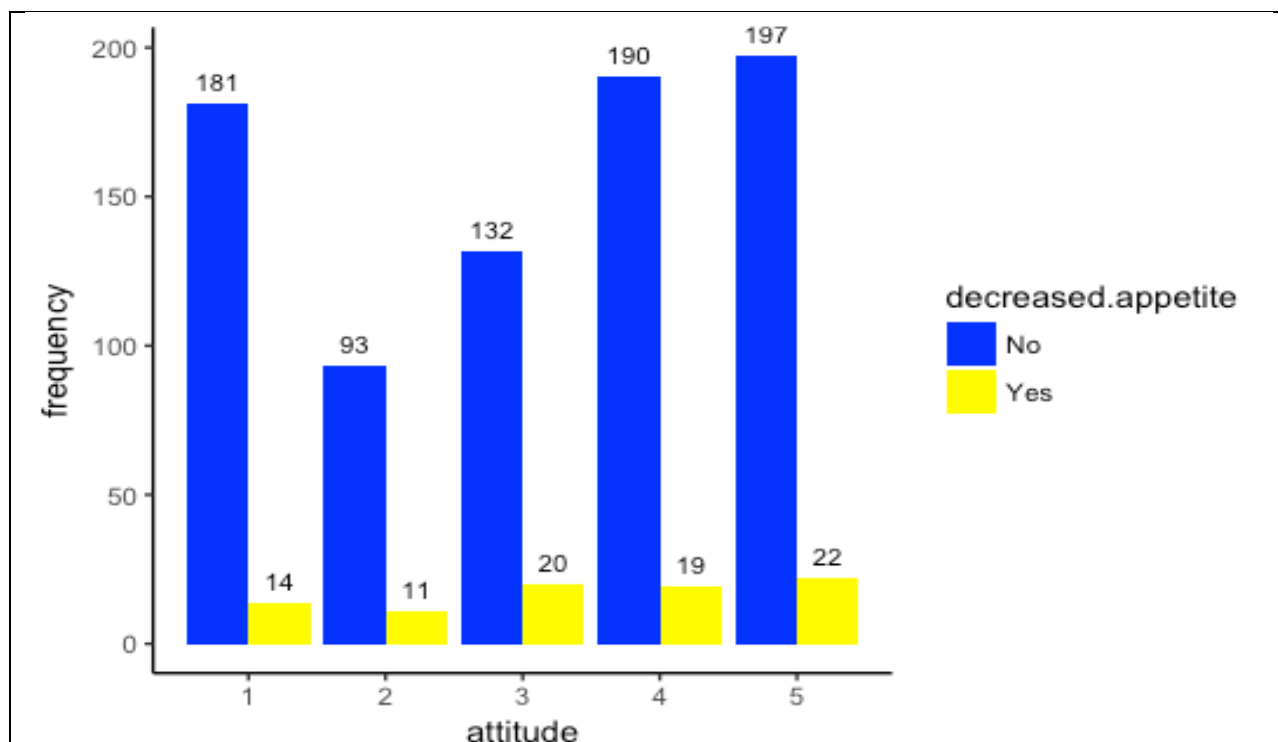


Figure 0-21 Frequency of decreased appetite for each level of attitude

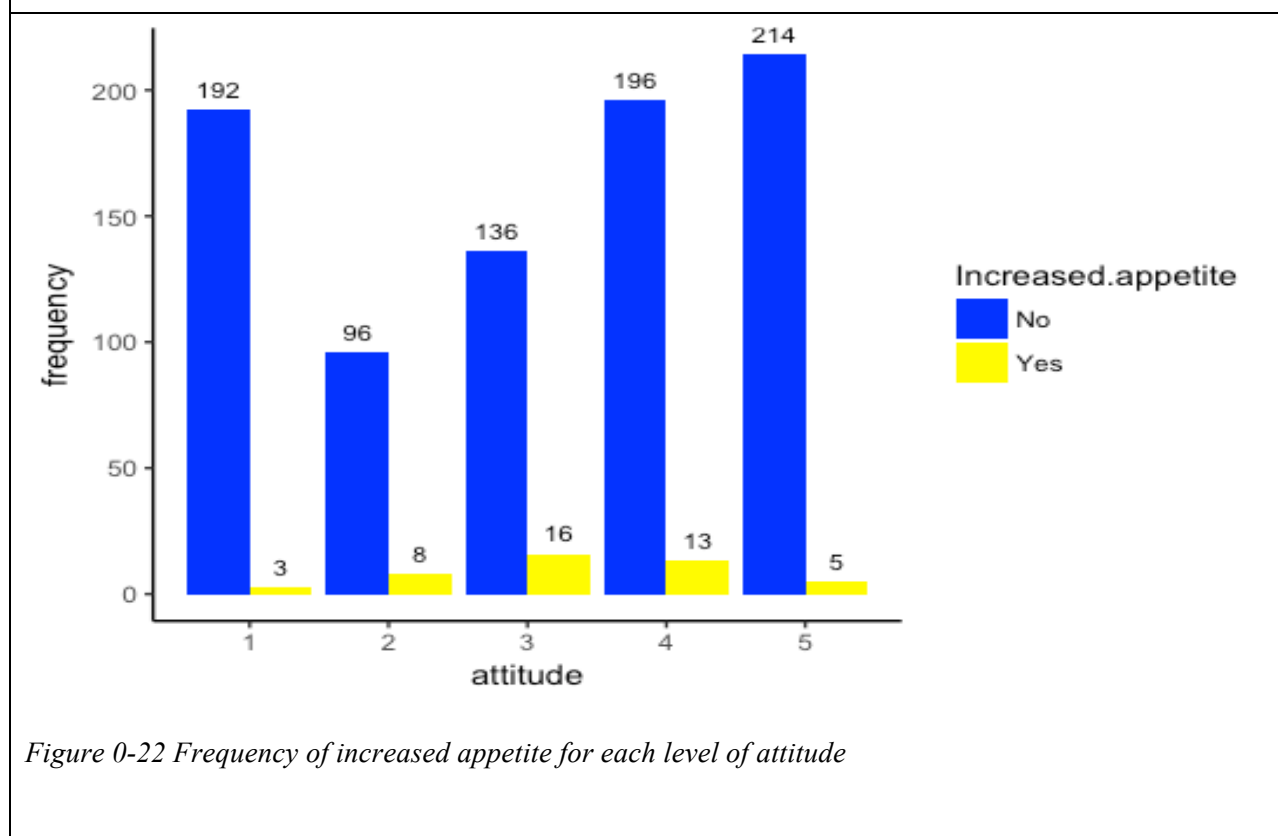


Figure 0-22 Frequency of increased appetite for each level of attitude

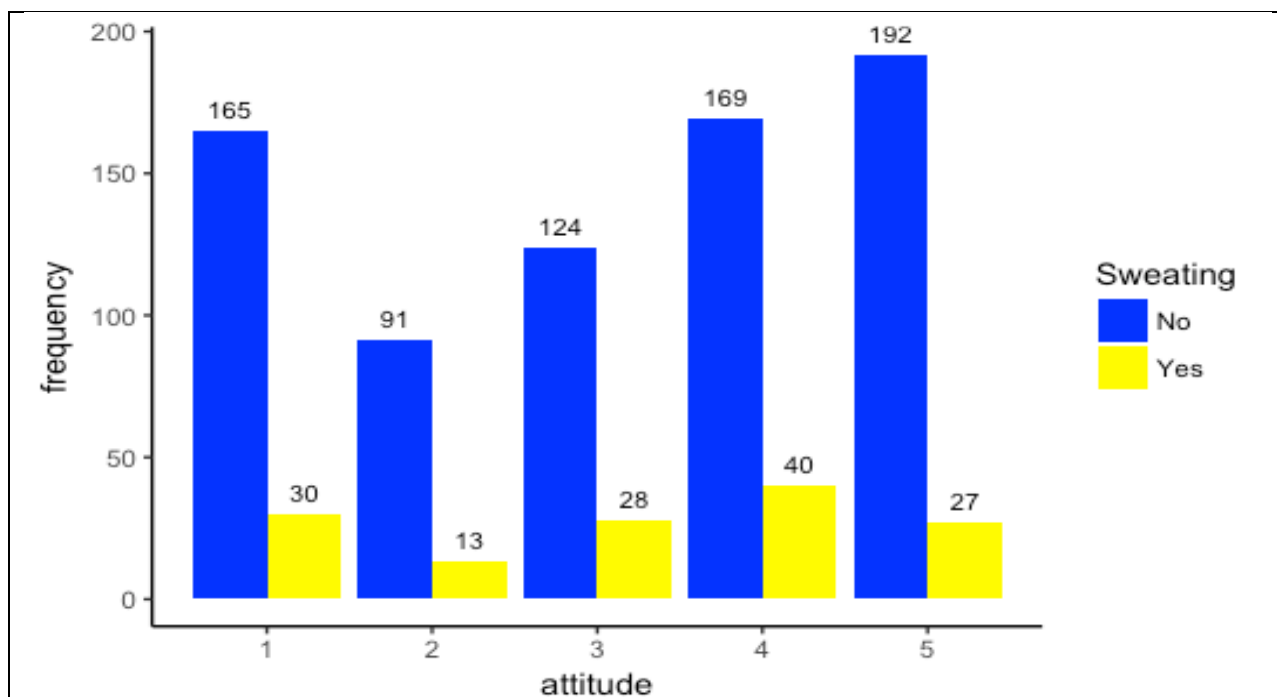


Figure 0-23 Frequency of sweating for each level of attitude

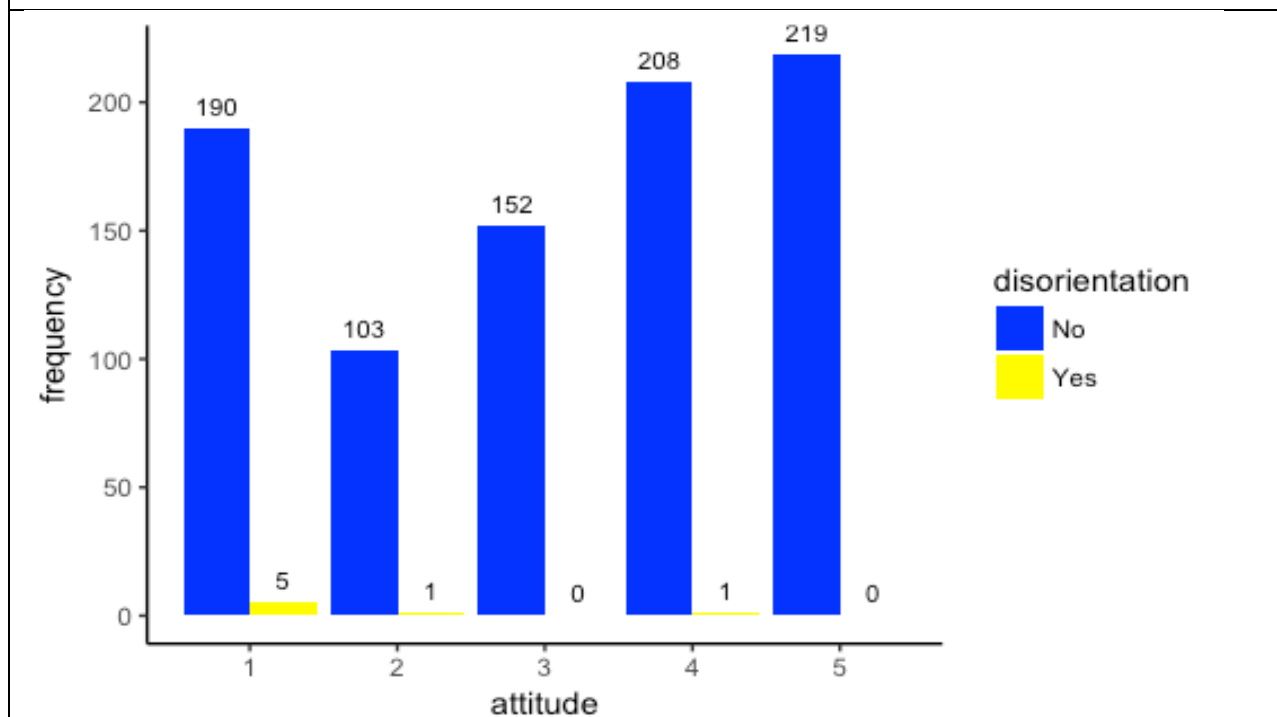
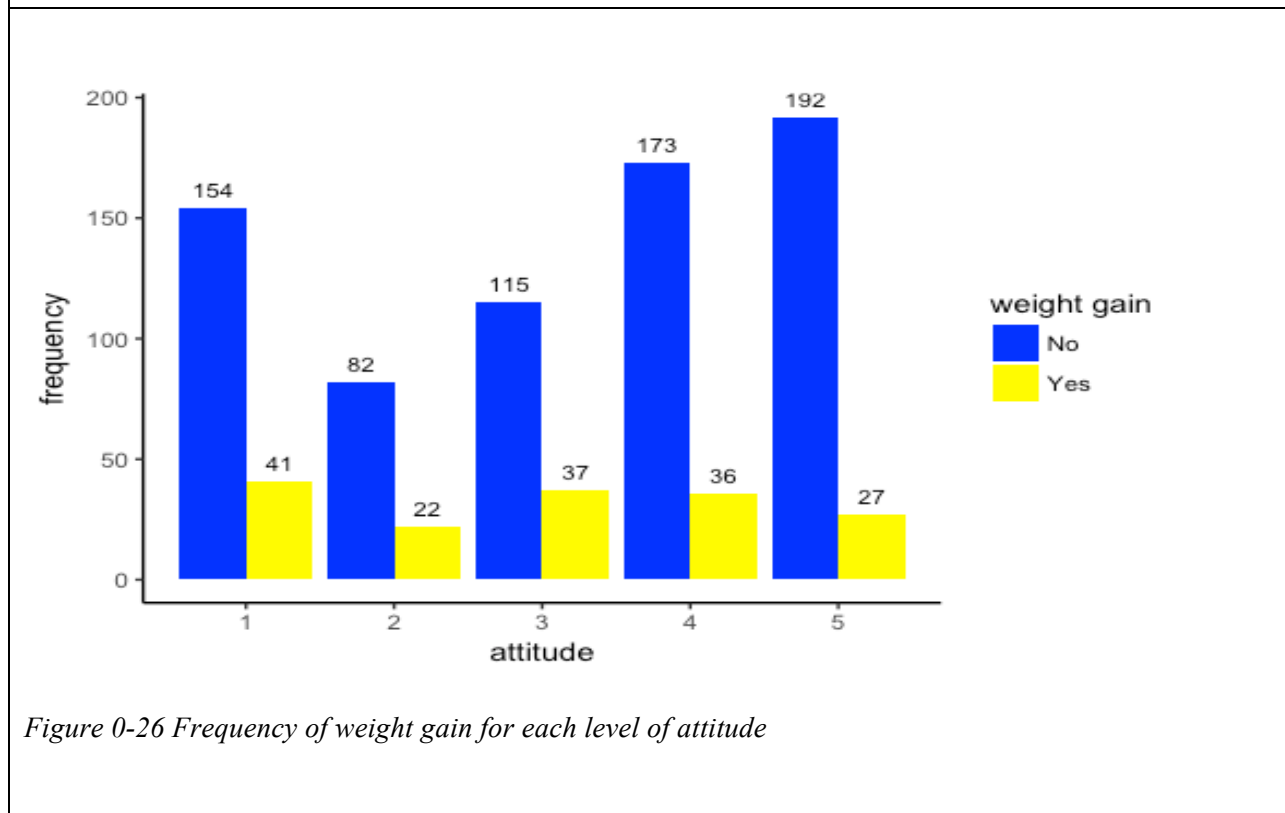
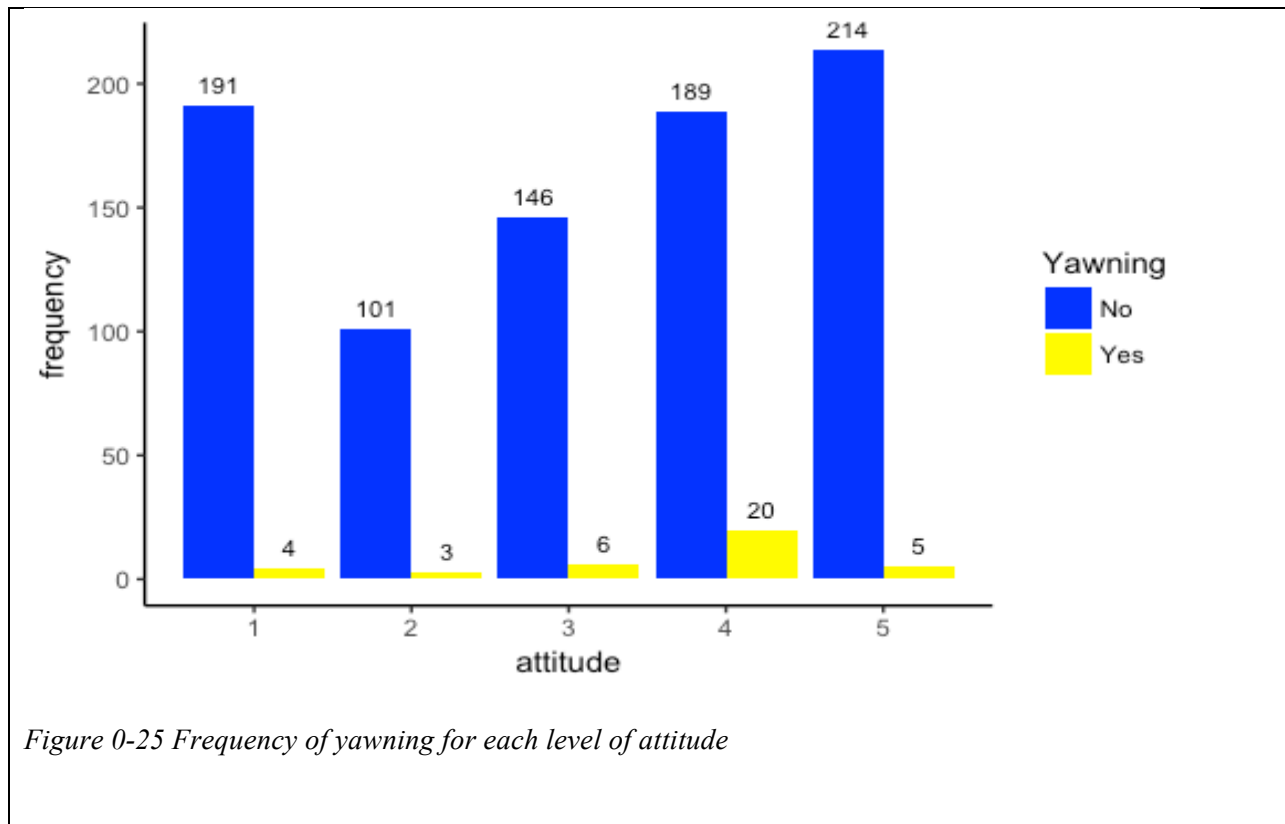


Figure 0-24 Frequency of disorientation for each level of attitude



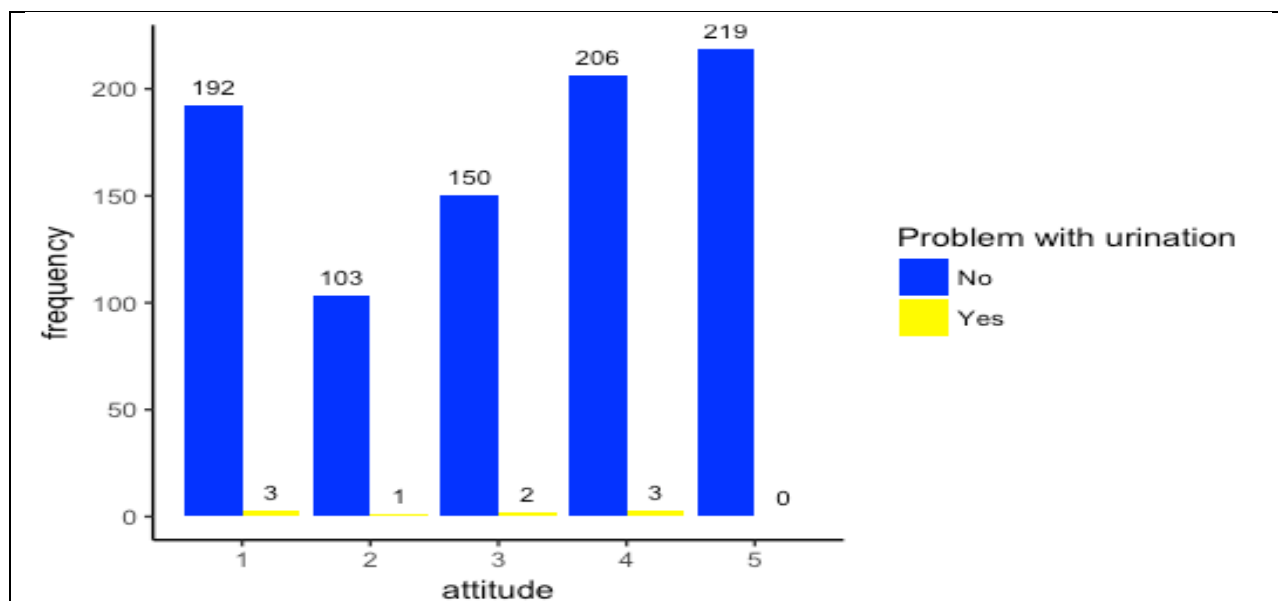


Figure 0-27 Frequency of problem with urination for each level of attitude

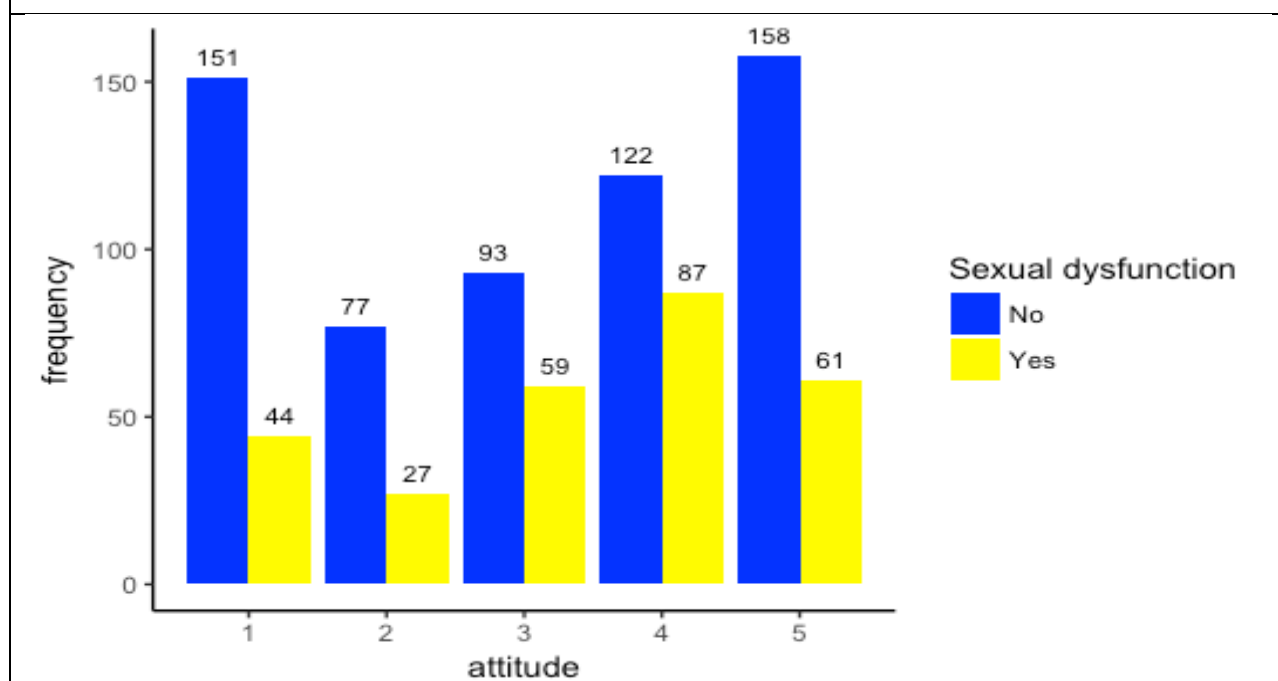


Figure 0-28 Frequency of sexual dysfunction for each level of attitude

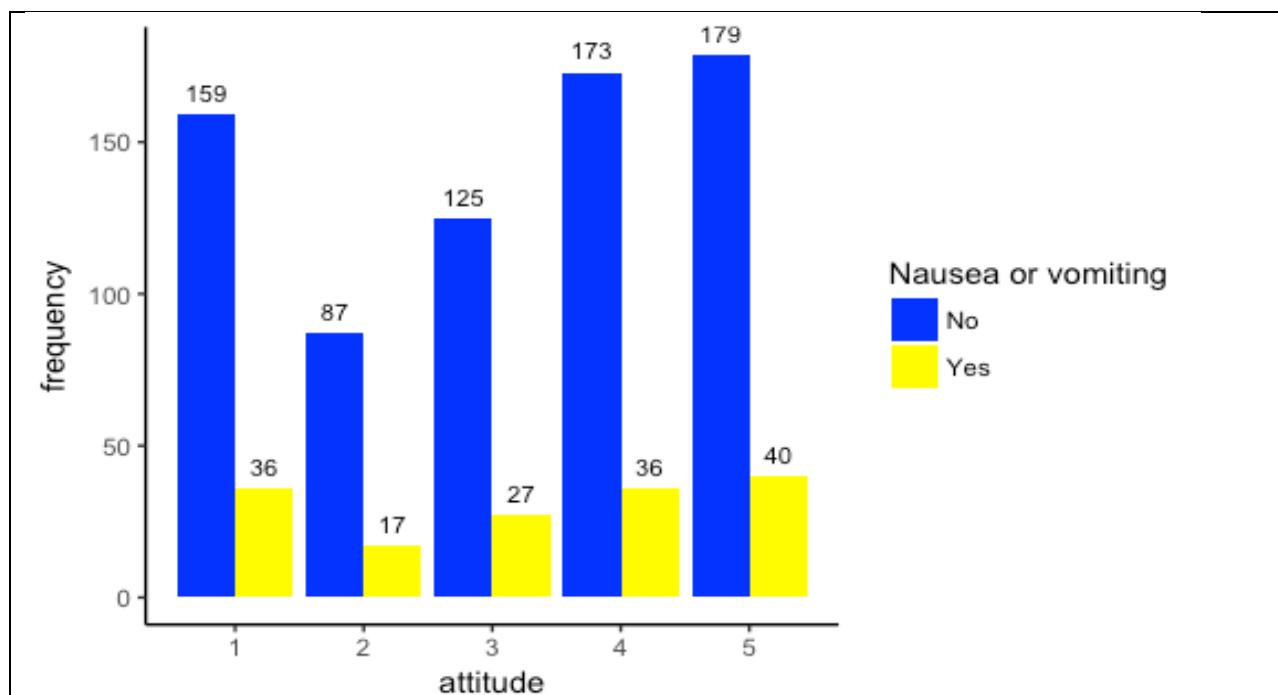


Figure 0-29 Frequency of nausea or vomiting for each level of attitude

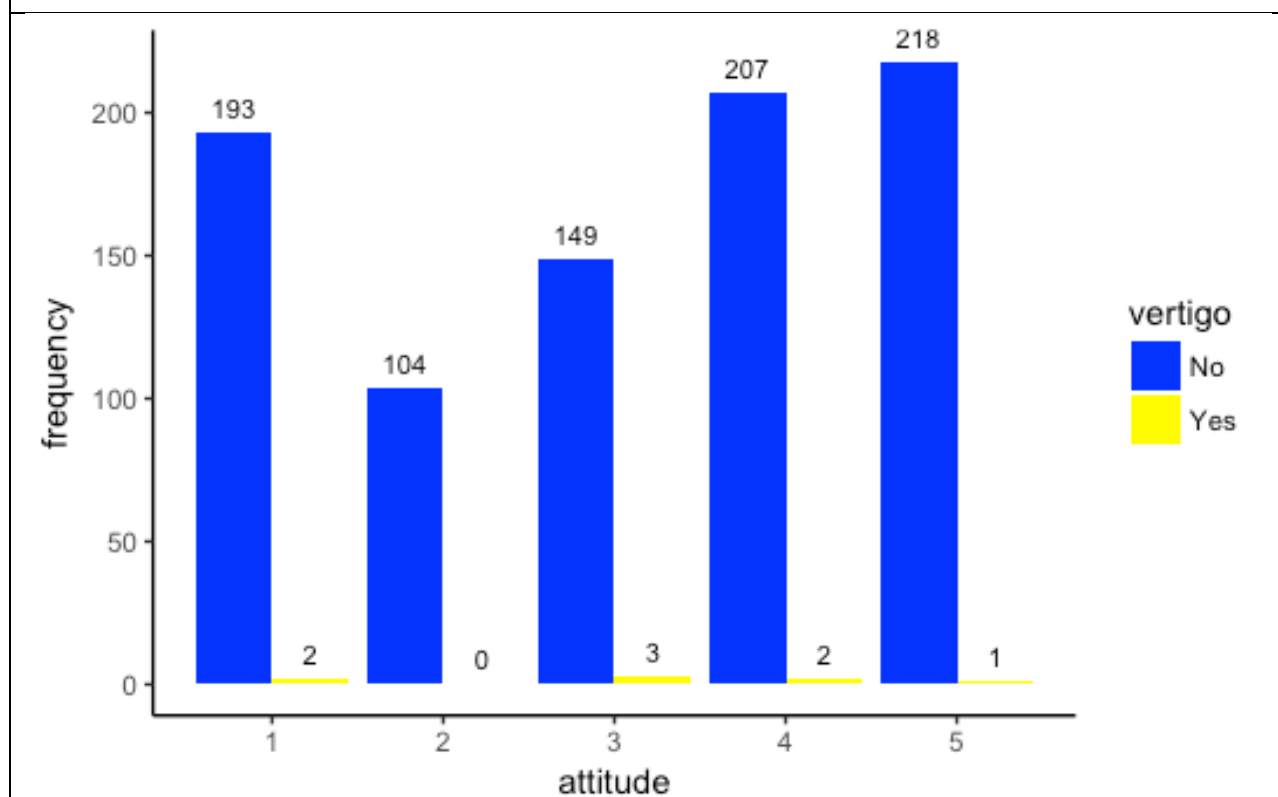


Figure 0-30 Frequency of nausea or vertigo for each level of attitude

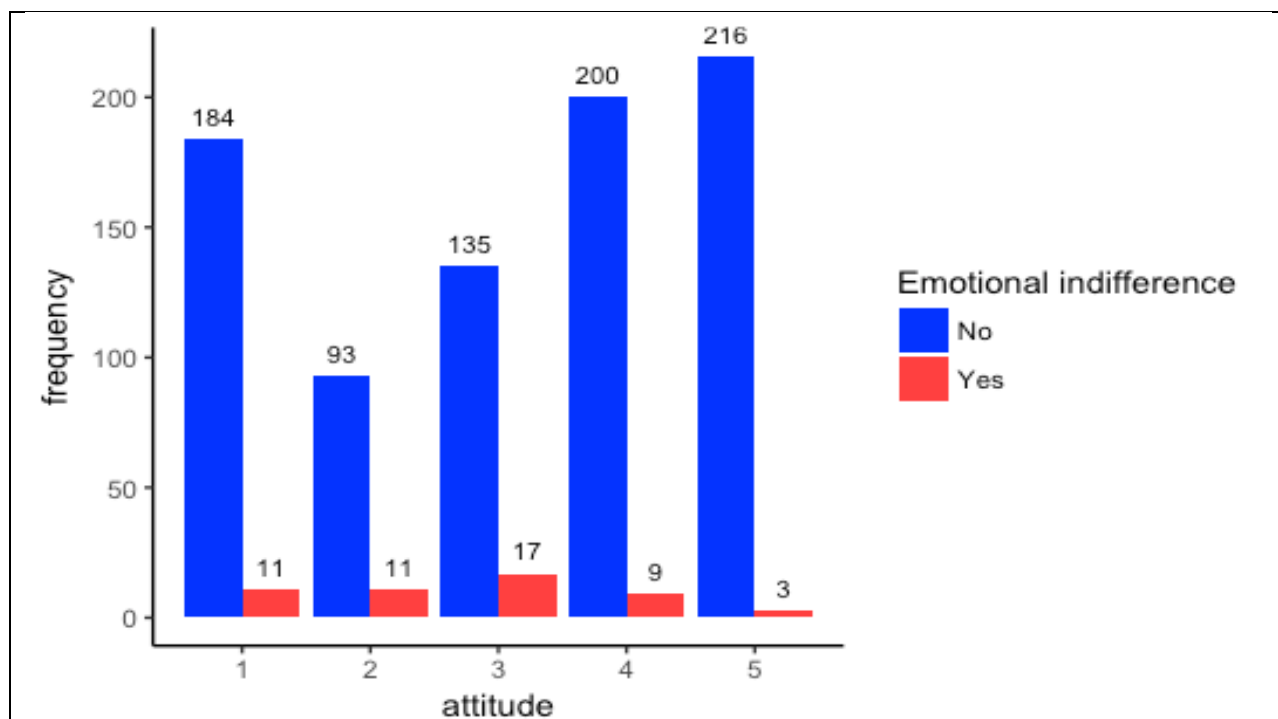


Figure 0-31 Frequency of emotional indifference for each level of attitude

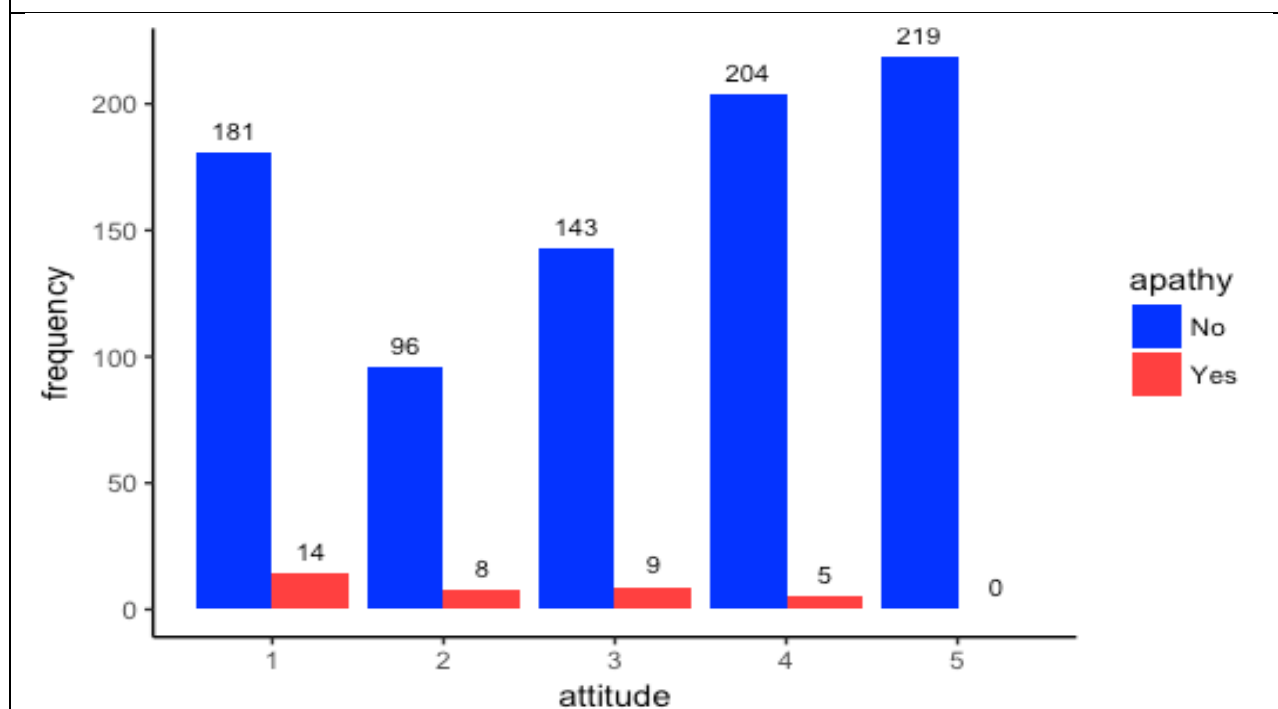


Figure 0-32 Frequency of apathy for each level of attitude

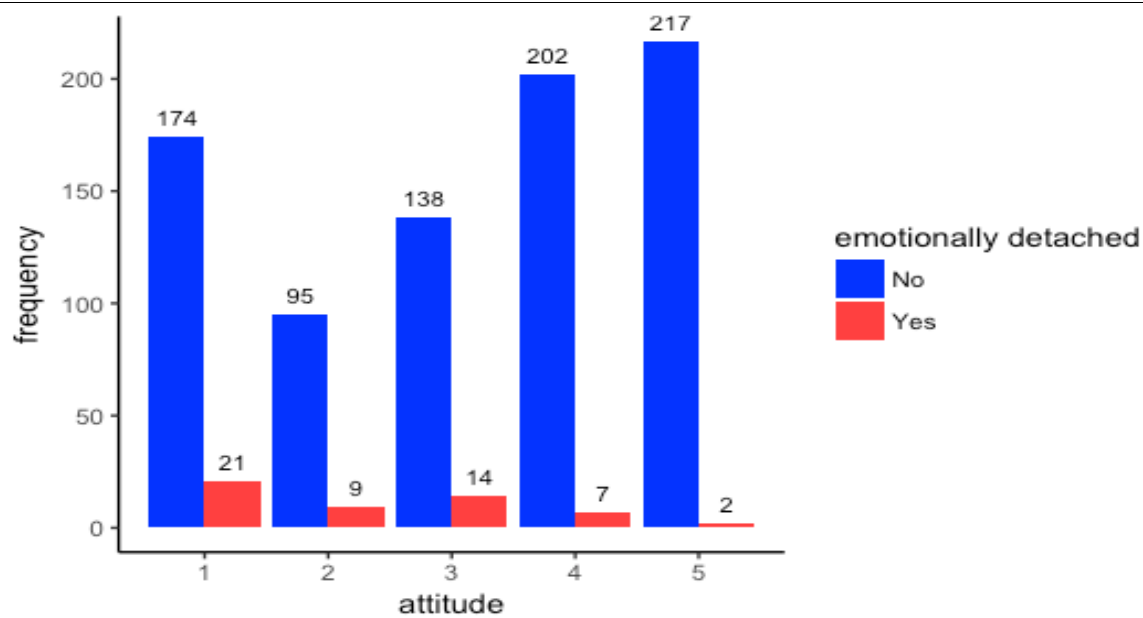


Figure 0-33 Frequency of emotionally detached for each level of attitude

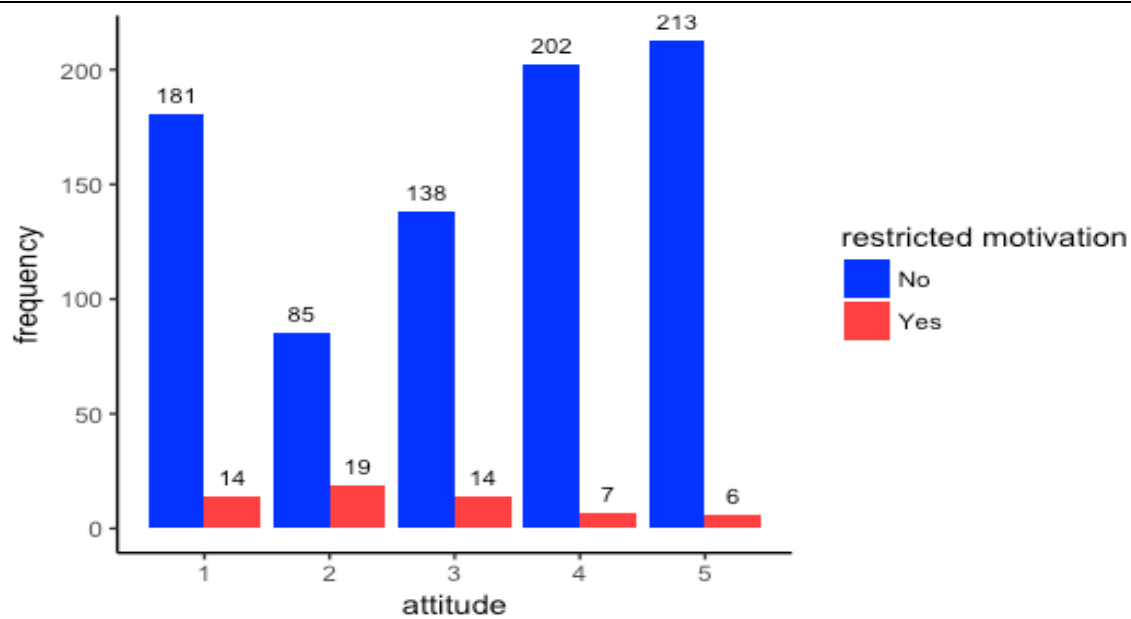
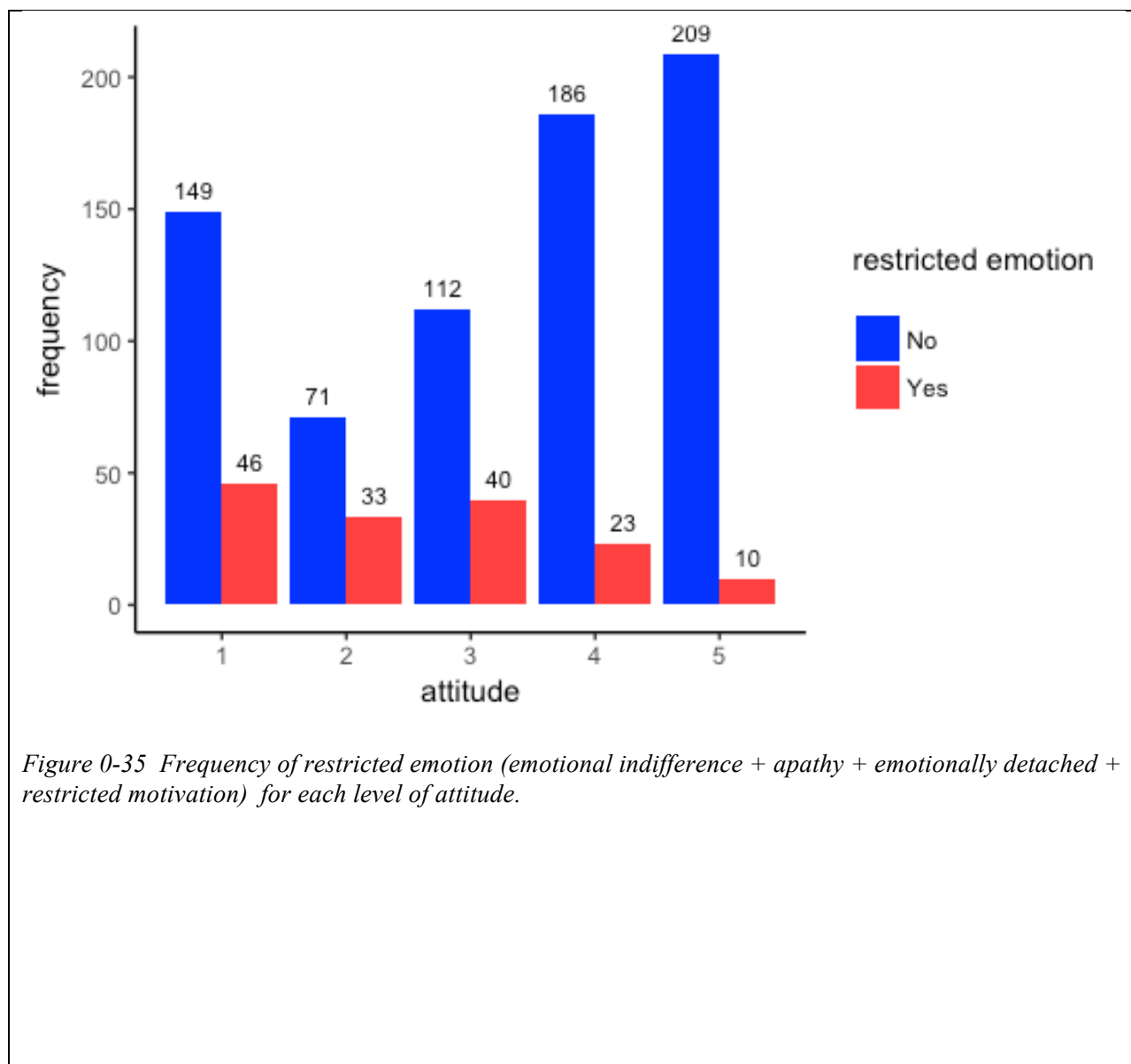


Figure 0-34 Frequency of mood swings for each level of attitude



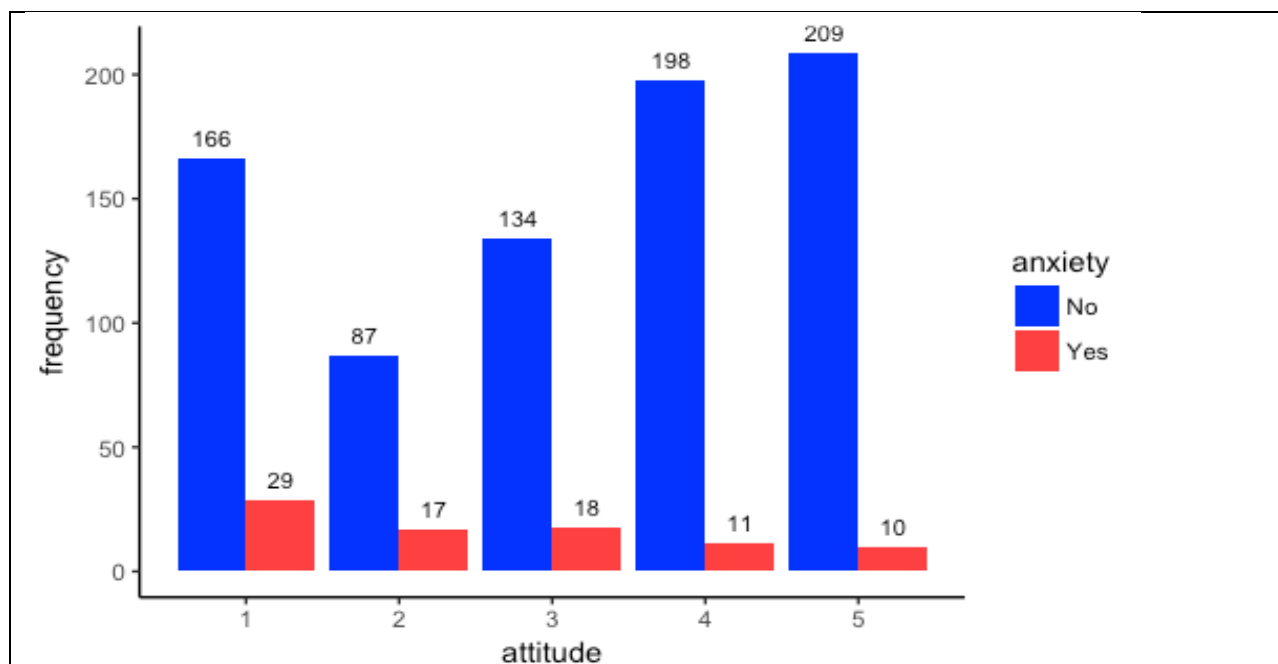


Figure 0-36 Frequency of anxiety for each level of attitude

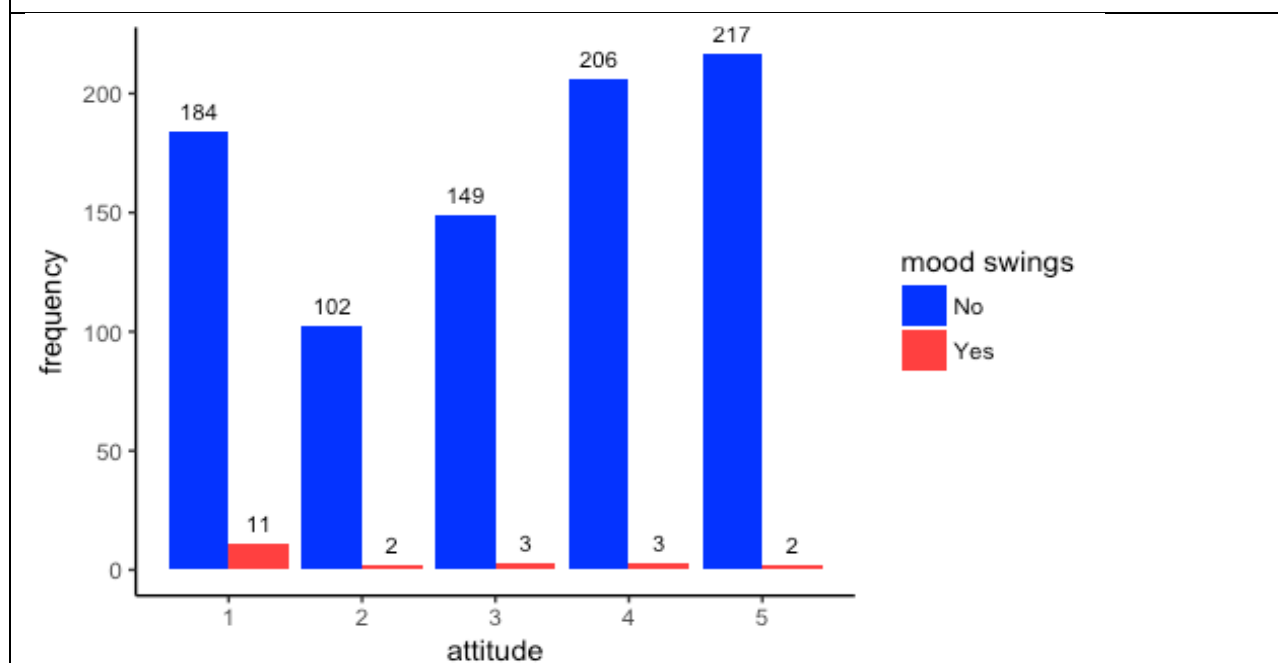


Figure 0-37 Frequency of motivation for each level of attitude

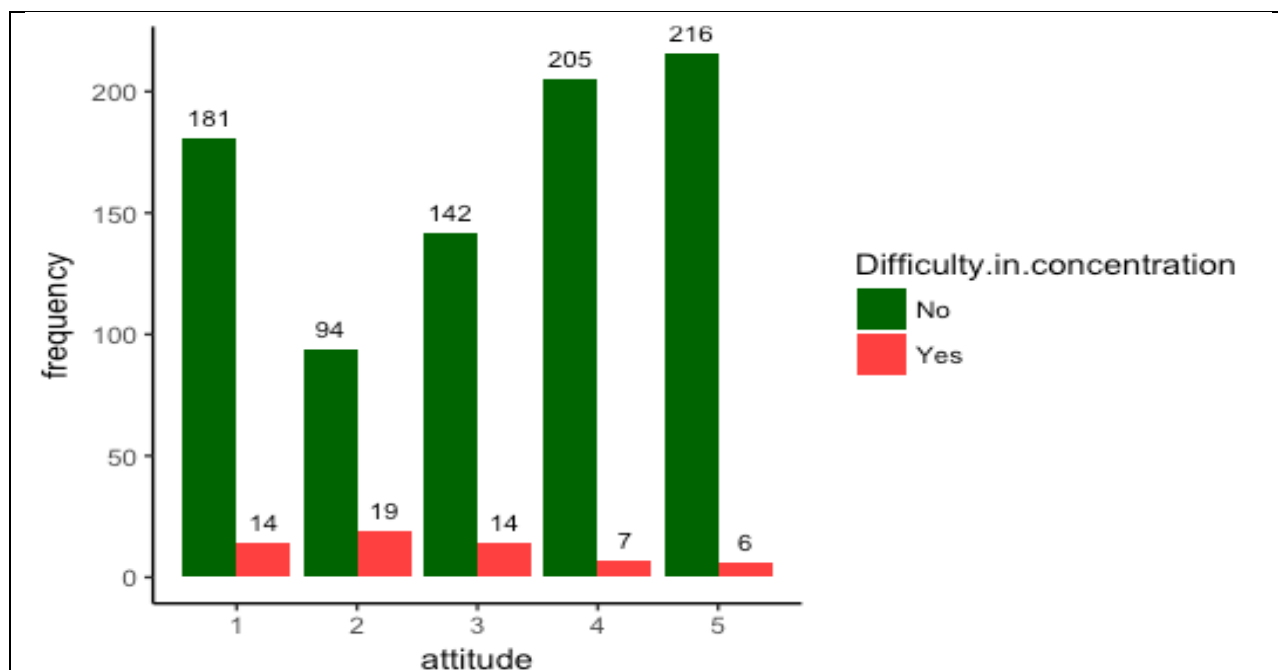


Figure 0-38 Frequency of difficulty in concentration for each level of attitude

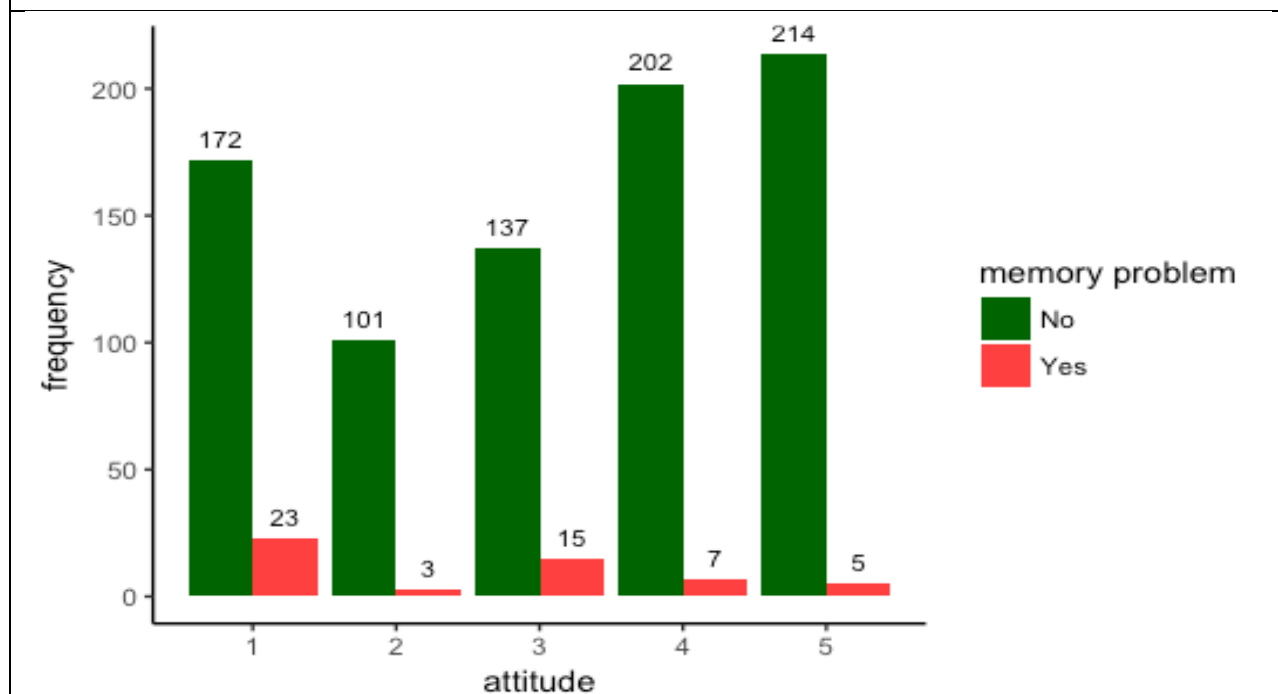


Figure 0-39 Frequency of memory problem for each level of attitude

Chapter 5: Discussion

This study involves both qualitative and quantitative approaches for the purpose of pharmacovigilance for psychiatric medications and generating and testing hypotheses concerning patients' attitudes toward antidepressants. The source of data is a healthcare forum called "askapatient.com" and the drug source are four psychiatric antidepressant medications including Sertraline (brand name: Zoloft) and Escitalopram (brand name: Lexapro) from the Selective Serotonin Reuptake Inhibitor (SSRI) Class and venlafaxine (brand name: Effexor XR) and duloxetine (brand name: Cymbalta) from the Serotonin Norepinephrine Reuptake Inhibitor (SNRI) Class.

5.1 Significance of Social Media for Providing Insight about Attitude and Adherence Behavior

The data analysis using the framework method showed that drug review posts in social media provide significant insight for patients' perception and attitude towards antidepressants and pharmacological aspects of drugs. However, it does not provide any insight for patients' adherence to medications, because the key factor in definition of adherence is a patient's agreement with the healthcare providers' treatment plan and drug reviews do not contain explicit information indicating whether a patient's withdrawal from medication was according to healthcare providers' recommendations or not.

5.2 Significance of Social Media in Identifying Underlying Factors Associated with Attitude

- 1) The analysis of data using analytical framework showed that drug reviews posts can be a significant source for underlying pharmacological factors affecting patients' attitudes. However, it does not provide reliable information to evaluate healthcare system factors, such as patients' perceived perceptions from communication with healthcare providers.

2) In this study, we measured association between patients perceived distresses from ADRs using patients drug review comments. This variable was measured using explicit indications including qualifiers representing severity, negative impacts of ADRs on daily activities and social life, and reporting severe life-threatening ADRs (suicidal ideation, suicidal attempt, and self-harm). The result of the data analysis showed that perceived distress from ADRs is significantly associated with patients' attitudes toward antidepressants. However, currently there is no self-report scale that was specifically designed to measure patient perceived distress received from ADRs associated with antidepressants. Although the Antidepressant Compliance Questionnaire (ADCQ) and the Beliefs about Medications Questionnaire (BMQ), the Drug Attitude Inventory (DAI) includes items that measure patients concern for antidepressants, such as long term affects and its impact on patient quality of life, losing autonomy, possibility of addiction, and control over feelings, they do not measure perceived distress that patients directly received from ADRs. For example, a patient may experience severe dry mouth, but does not have concern about long-term effects of antidepressants on life quality or control over feeling.

In addition, the total ADR score obtained from the Antidepressant Side-Effect Checklist (ASEC) does not represent perceived distress that a patient received from a set of ADRs. For example, a patient may obtain total score "four" from four reported mild ADRs, while another patient may obtain total score "three" from one severe ADR. Apparently, the second patient received higher perceived distress from the first one, however the calculated total number does not represent it.

- 3) Drug reviews also provide information about patients withdrawal experience as significant predictive factors for patients attitude toward antidepressants. However, self-report scale for measuring patients' attitude for antidepressants do not include any items related to patients' withdrawal experience.

5.3 Limitations of Social Media in Identifying Underlying Factors Associated with Attitude

Drug reviews in healthcare forums compared with self-report scales has some limitations:

- 1) First, they are not a rich source of patients' perceptions towards healthcare providers. The ADCQ, a self-report scale that specifically designed for measuring patients attitude towards antidepressants, includes a component that measure three dimensions of patient-physicians interaction: 1) patient's perception of physician's knowledge; 2) patient's perception of sufficiency of knowledge provided by physicians; and 3) patient's perception of communication effectiveness with providers. We attempted to measure patients' satisfaction with healthcare providers using the three dimensions in the drug reviews. However, majority of the patients did not provide explicit indications showing their perceptions towards clinicians, which caused low IAA among annotators.
- 2) Drug reviews posts also do not provide explicit information about patients general concern and necessity towards medications. However, the BMQ self-report scale includes components for measuring general concern and necessity towards medications. In addition, the BMQ includes a component for measuring the perceived necessity of a prescribed medication, while in the data analysis of this study, patients' expressions of perceived necessity for antidepressants was labeled as "drug effectiveness", because of difficulty in differentiation of perceived effectiveness and perceived necessity in sentence classification.

- 3) Furthermore, while the ADCQ self-report scale can capture patients perceived social support, few patients expressed their perceived social support in drug reviews.

By comparing items in the final analytical framework and the self-report scales (the ADCQ, the BMQ, and the DAI), we conclude that drug reviews cannot be used as an alternative source for measuring patients' attitudes towards antidepressants compared to self-report scales. However, they can be used as a supplementary source, due to providing insights for some underlying factors that are not available in the self-report scales.

5.4 Significance of Social Media in Identifying ADRs and WDs Associated with Antidepressants

Identifying ADRs and WDs from drug review posts showed that this data source is a valuable source for identifying ADRs and WDs associated with antidepressants. Although Uher et al. (2009) demonstrated that adverse effects of antidepressants can be reliably assessed through self-report scales, the inherent limitations associated with the scales, such as low coverage of potential adverse effects, particularly the rare ones, may reduce the reliability and validity of the scales. For example, the ASEC scale does not cover psychological, cognitive and functional problems associated with antidepressants, which is the major limitation of studies employing this scale for measuring ADRs associated with antidepressants (Bet, Hugtenburg, Penninx, & Hoogendijk, 2013; Uher et al., 2009).

The Liverpool University Neuroleptic Side Effect Rating Scale (LUNSERS) and the WHO Quality of Life Assessment-Brief version (WHOQOL-BREF) are other self-report scales used for measuring the ADRs associated with antidepressants. The LUNSERS is also associated with some limitations. First, it is not specifically designed for measuring antidepressants' adverse effects. Second, the ADRs indicating emotional problems and cognitive dysfunctions are very

limited and do not reflect a wide range of emotional and cognitive adverse effects of antidepressants.

Items in the WHOQOL-BREF assess individual quality of life from a general perspective that may not be directly related to the impact of antidepressants on quality of life. For example, items such as “level of satisfaction with living place” or “having enough money to support your needs” do not measure the impacts of ADRs associated with antidepressants.

These scales may also suffer from patients’ recall bias, because they do not define a specific time-period during which the adverse effects have been occurred. Hence, patients may just report severe adverse effects they experienced and disregard the mild ones (Bet et al., 2013).

Patients’ self-reports in healthcare forums are not associated with these limitations. First, patients can access the forums at their own convenience. Hence, their reports may not suffer from memory bias. Second, in contrast with structured questionnaires that limit the patients’ reports to a specific set of ADRs, patients in healthcare forums can report any psychological, physiological, and cognitive ADRs using their own language. So, there is no concern regarding the patients’ wrong interpretation of ADRs specified in the questionnaires (self-report scales). In addition, patients voluntarily report their experiences with medications. So, the concern for hiding information or breaching confidentiality would be minimized.

Regarding this information, we can conclude that drug review posts in social media can address limitations associated with self-report scales (questionnaires) and they can be used as an alternative source for measuring ADRs associated with antidepressants.

5.5 Implications of this Study

5.5.1 Developing tools in the area of text mining algorithms and machine learning for extraction of health related information from consumer health posts

The three phases of this study: 1) applying the analytical framework at level of sentences; 2) identifying ADRs, WDs, and DIs; and 3) mapping the identified entities to both UMLS and SNOMED-CT concepts, provides a corpus that has several implications for developing tools in the area of text mining and machine learning for extracting health related information from consumer health posts. First, the generated dataset can be used to improve the recall of dictionary-based systems designed for automatic identification of pharmacological aspects of drugs. Second, the dataset has a significant implication in developing and evaluating text mining and machine learning systems aimed to identify ADRs, WDs, and drug indications from consumer health posts in social media. Third, the dataset can be used for training machine learning-based classifiers aiming to distinguish ADRs from other semantic types, such as drugs indications. Forth, the dataset has important implications for developing and testing automatic system aimed to measure effectiveness and ineffectiveness associated with psychiatric medications. Fifth, it can be used for developing systems targeting automatic mapping between layperson expression of health information to UMLS and SNOMED.

Currently, there is one open source corpus that has been developed using consumer health posts, which is called CADEC ("Health Information Technology for Economic and Clinical Health Act," ; Karimi et al., 2015). This corpus consists of 1,231 comments for two categories of drugs, Diclofenac and Lipitor. The identified entities include ADRs, symptoms, disease, and mentions of drugs, which were mapped to SNOMED-CT, MEDRA, and AMT.

Our corpus is different from the CDADE from several important aspects, including type of drugs studied, methodology of development, identified entities, and normalization process. Because of the nature of the psychiatric medications and the patient population of this study, the ADRs events, specifically functional problems are not expected to be well-covered by the

CADEC corpus. Moreover, the corpus created in this study includes the drug effectiveness/ineffectiveness and qualifiers representing severity and persistency of ADRs and WDs that are not covered by CADEC. Finally, all the identified entities were mapped to UMLS, which was not reported by CADEC.

5.5.2 Testing the association between physiological, psychological, cognitive, and functional problems with attitude

In this study, 698 ADRs and 218 WDs were identified from which the association between 164 ADRs and attitude were tested. 128 ADRs were grouped using ADRs specified in the ASEC questionnaire, and 36 ADRs grouped into psychological and physiological ADRs specified in the literature. Accordingly, we did not measure the association between 534 ADRs (specified in the dataset) with attitude. The ADRs, such as “detailed recall of dream” and “suicidal feeling”, “malaise” may have severe negative effect on patients’ attitudes toward antidepressants and consequently adherence behavior. Future studies can investigate the association between the ADRs and patients’ attitudes.

5.5.3 Developing self-report scales

The dataset that was produced using the analytical framework and the ADRs/WDs identification phase can be used for designing patient-driven self-report scale for measuring patients’ attitudes towards antidepressants.

4.5.4 Implications of the analytical framework

The analytical framework developed in this study has significant implications for data analysis from other healthcare forums collecting patients’ experiences with pharmacological treatments and also personal health records that include patients’ experiences with medications.

5.5.5 Implication of the findings related to the factors associated with attitude

Findings of this study have significant implications for developing clinical interventions aiming to improve patients' attitudes and adherence towards medications. In this study, we showed that lack of knowledge about ADRs and WDs is significantly associated with patients' negative attitudes towards antidepressants. Therefore, clinicians should design and implement effective communication mechanisms to inform patients the potential ADRs and WDs that they may experience during the process of treatment and discontinuation.

In addition, we found that patients' withdrawal experience is associated with patients' negative attitudes towards antidepressants. The finding implies that clinicians should consider interventions to reduce the perceived distress that patients may receive during the discontinuation process of antidepressants.

We also found that drug effectiveness is the most significant factor associated with attitude, implying that patients at initial phase of treatment may discontinue antidepressants abruptly if they do not find it effective in resolving depression symptoms. Regarding the fact that antidepressants full effects are not seen for typically four to six weeks, clinician should consider interventions to track antidepressants' effectiveness and inform patients for treatment mechanism of the drugs.

5.6 Limitations of this Study

1) *Lack of information on drug-drug interactions, drug-herb interaction, and drug overdose:*

Identifying ADRs for consumer health posts and evaluating their associations with patients' attitude toward drugs and antidepressants has some limitations. The focus of patients in the review posts was mostly on a specific 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 treatments. Moreover, some of the ADRs for psychiatric medications, such as suicidal ideation or emergency visits, can happen due to patient's overdose, which were not available in these review posts.

2) *Uncertainty of data in social media:* Although patients' self-report experience is a reliable source for evaluating pharmacological effects of medications, the risk of inaccurate and false information still exists.

3) *Concern about dominating dissatisfied patients in drug reviews forum:* There is a concern that patients with negative perception for medication have more willingness to report their experiences with medications. However, the statistical analysis of the sample size in this study showed that almost 50% of patients were satisfied with psychiatric medications as they rated them as 4 and 5. While only 35% of patients rated the drugs as 1 and 2. In addition, the data set showed that almost half of the reviewers used the antidepressants more than a year, suggesting that they were not the most dissatisfied patients with antidepressants.

4) *Concern about misinterpretation of ADRs and WDs:* one problem with data analysis for consumer health posts for psychiatric medications is the possibility that users misinterpret the symptoms of their mental disorder with ADRs. Although this study recorded only ADRs that patients certainly linked them to the psychiatric medications, we cannot exclude patients' misinterpretation of depression symptoms with ADRs.

5) *Concern about representativeness of the sample:* As healthcare forum data are self-selected, there is the risk that the dataset is not representative of patient population. However, the distribution of sample size in this study for both gender, female and male, is consistent with those identified in conventional prevalence studies.

The review posts in an online healthcare forum may also not be a representative source for all demographic groups. Some minorities may not have the same access level and skills for adoption and usage of technology for reporting their experiences in online healthcare forums.

6) *Insufficiency of information to estimate prevalence of ADRs/WDs:* This study could not estimate the prevalence of ADRs and WDs, as the healthcare forum contained no prompt for patients to disclose the particular effects. While some patients seemed to list all ADRs and WDs they experienced, others mentioned few ADRs. In addition, some patients reported difficulty in medication discontinuation, while they did not report any specific withdrawal symptoms. However, the significance of this study compared to the conventional methods for collecting patients' experiences with medications is that the drug reviews in healthcare forums are spontaneous, open-ended, and uncensored format, and they were not collected for a specific project.

7) *Sample size:* The sample size of this study is limited to 892 posts for four psychiatric medications. While this sample size is a good representative of posts available in “askapatient.com” for the four psychiatric medications, it may not be a balanced representative of other consumer posts in this forum or other healthcare forums. Additionally, it is possible that a specific group of patients report an unbalanced sample of experiences with drugs in the forum.

8) *Limitation for coverage of medications:* Our corpus covers sentence classification and entities identifications for two classes of psychiatric medications, SSRI and SNRI. While limiting dataset to specific set of drugs enabled us to have a better understanding of the conceptual models associated with layperson and professional expressions of medical entities, it may not include the rare ADRs related to other classes of psychiatric medications, such as TCA.

9) Terminology Mapping challenges

We managed to improve accuracy of inferring the intended meaning of colloquial expression of ADRs by use of a preliminary mapping table based on clinical trials and qualitative studies, contextual cues in patients reviews, and discussion meetings. However, we had some challenges for selecting proper UMLS/SNOMED-CT concepts for ADR expressions. Throughout the corpus, we could not map 130 expressions (out of 6534) to either UMLS or SNOMED concept. For example, the ADR of “hardly feel human anymore”, could not be mapped to any concepts due to uncertainty of the underlying concepts associated with it. Indeed, it is unclear what the patient meant with this expression: is it about the patient feelings emotionally detached, having a problem in performing daily activities, or is it about feeling detached from his/her mind or body (de-realization)?

Another challenge is that layperson expressions of ADRs are fuzzier than the corresponding UMLS/SNOMED concepts. For synonym concepts, the layperson ADRs expressions are more likely to be “narrower-than” or “broader-than” their closest UMLS concept. For example, “not being able to express sadness” or “could not cry in funeral ceremony” were all mapped to “blunted affects”. This happens particularly for psychological systems and functional problems.

There were also some cases that, while expression of an ADR is clear and can be translated to equivalent medical concept, no UMLS concept is available for that expression. For example, brain shivers (brain zap) (J. Aronson, 2005) is a medical concept that has been reported in a few medical research as an ADR/WD of psychiatric medication, however, no medical concept is available for it in UMLS. Eighty-seven (out of 130) expressions without any UMLS concepts were related to this medical concept.

Limiting drug reviews to a set of drugs prescribed for a specific disease or mental disorders, helps to have a better understanding of patients' colloquial expressions of ADRs by comparing the expressions with clinical trial studies and qualitative studies addressed the ADRs for the drugs.

10) 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 it to a wrong label. In addition, the span of the identified entities may include less or more information than necessary. These issues may affect the performance of machine learning systems trained based on this corpus to identify drug effectiveness, ADRs, and drug indications in consumer health posts.

11) Possibility of bias in findings: as the statistics of data source in this study showed, 76% of the participants in the healthcare forum are female. So, findings of this study may be associated with gender bias.

Conclusion

In this study, using a mixed method data analysis, we showed that consumer health posts in social media provide a unique insight into issues such as patients perspective, withdrawal experience, tolerability for adverse drug reactions, and overall attitudes towards medications. However, drug review posts is not a rich source for patients' perceptions of communication with healthcare providers, affordability of the drugs, or social support for the patients. The most important use of the drug reviews is providing safety detection signals for ADRs and WDs associated with medications. The drug reviews in healthcare forums are specifically important for psychiatric medications because of the limitations of clinical trials in detecting ADRs associated with this drugs, such as emotional indifference and suicidal ideation.

Using methods of framework analysis and terminology mapping, we provided structured data for testing hypotheses concerning patients' attitude toward antidepressants. The result showed that drug effectiveness and perceived distress received from ADRs are the most significant factors affecting patients' attitude toward antidepressants. In addition, among three sets of psychological and physiological, and cognitive ADRs associated with antidepressants, we found that physiological and cognitive ADRs are significantly related to patients' negative attitude, while psychological ADRs are partially associated with attitude towards antidepressants.

This work has important implications for generating and testing medical hypotheses concerning patients experiences with psychiatric medications specified in this study, such as withdrawal experience and also measuring association between ADRs and attitude. In addition, it can be used for improving performance of text-mining algorithms aiming to automatically detect healthcare information from consumer health posts and personal health records.

Future Work

In this study, we developed an analytical framework that can be used for identifying underlying factors associated with patient attitudes towards antidepressants. The analytical framework was created and tested using drug reviews in a healthcare forum called "askapatients.com". Future work may use the framework in other healthcare forums or personal health records to identify factors associated with patients' withdrawal and perceptions towards medications. In addition, in this study, the phase of sentence classification, entities identification, and entities normalization created a corpus that can be used for developing automatic systems aimed to identify health-related information from healthcare forums and map them to standard medical vocabularies. We are in the process of developing classifier algorithms to differentiate drug reviews contains withdrawal experience (both intentional and unintentional) from drug reviews without report of

withdrawal. Finally, we are going to compare the identified ADRs and WDs in this study with SIDER Side Effect Resource, in order to identify new ADRs/WDs that were not reported by clinical trials.

REFERENCES

- Acosta, F., Rodríguez, L., & Cabrera, B. (2013). Beliefs about depression and its treatments: Associated variables and the influence of beliefs on adherence to treatment. *Revista de Psiquiatría y Salud Mental (English Edition)*, 6(2), 86-92.
- Aikens, J. E., & Klinkman, M. S. (2012). Changes in patients' beliefs about their antidepressant during the acute phase of depression treatment. *General hospital psychiatry*, 34(3), 221-226.
- Aikens, J. E., Nease, D. E., & Klinkman, M. S. (2008). Explaining patients' beliefs about the necessity and harmfulness of antidepressants. *The Annals of Family Medicine*, 6(1), 23-29.
- Aikens, J. E., Nease, D. E., Nau, D. P., Klinkman, M. S., & Schwenk, T. L. (2005). Adherence to maintenance-phase antidepressant medication as a function of patient beliefs about medication. *The Annals of Family Medicine*, 3(1), 23-30.
- Ajzen, I. (1985). From intentions to actions: A theory of planned behavior *Action control* (pp. 11-39): Springer.
- Akerblad, A., Bengtsson, F., Holgersson, M., von Knorring, L., & Ekselius, L. (2008). Identification of primary care patients at risk of nonadherence to antidepressant treatment. *Patient Prefer Adherence*, 2, 379-386.
- Al-Harbi, K. S. (2012). Treatment-resistant depression: therapeutic trends, challenges, and future directions. *Patient Prefer Adherence*, 6, 369.
- Alekhyia, P., Sriharsha, M., Ramudu, R. V., Shivanandh, B., Darsini, T. P., Siva, K., . . . Reddy, Y. H. (2015). Adherence to Antidepressant Therapy: Sociodemographic Factor Wise Distribution. *Int J Pharm Clin Res*, 7, 180-184.
- Aronson, A. R. (2006). Metamap: Mapping text to the umls metathesaurus. *Bethesda, MD: NLM, NIH, DHHS*, 1-26.
- Aronson, J. (2005). Bottled lightning. *BMJ*, 331(7520), 824.
- Artstein, R., & Poesio, M. (2008). Inter-coder agreement for computational linguistics. *Computational Linguistics*, 34(4), 555-596.
- Benton, A., Hill, S., Ungar, L., Chung, A., Leonard, C., Freeman, C., & Holmes, J. H. (2011). A system for de-identifying medical message board text. *BMC bioinformatics*, 12(3), 1.
- 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.
- Bergman, M. M. (1998). A theoretical note on the differences between attitudes, opinions, and values. *Swiss Political Science Review*, 4(2), 81-93.
- Bet, P. M., Hugtenburg, J. G., Penninx, B. W., & Hoogendijk, W. J. (2013). Side effects of antidepressants during long-term use in a naturalistic setting. *European Neuropsychopharmacology*, 23(11), 1443-1451.
- Bhaskaran, K., & Smeeth, L. (2014). What is the difference between missing completely at random and missing at random? *International journal of epidemiology*, 43(4), 1336-1339.
- Blenkinsopp, A., Wilkie, P., Wang, M., & Routledge, P. (2007). Patient reporting of suspected adverse drug reactions: a review of published literature and international experience. *British journal of clinical pharmacology*, 63(2), 148-156.

- Bradley, K. L., McGrath, P. J., Brannen, C. L., & Bagnell, A. L. (2010). Adolescents' attitudes and opinions about depression treatment. *Community mental health journal*, 46(3), 242-251.
- Brown, C., Battista, D. R., Bruhlman, R., Sereika, S. S., Thase, M. E., & Dunbar-Jacob, J. (2005). Beliefs about antidepressant medications in primary care patients: relationship to self-reported adherence. *Med Care*, 43(12), 1203-1207.
- Bulloch, A. G., & Patten, S. B. (2010). Non-adherence with psychotropic medications in the general population. *Social psychiatry and psychiatric epidemiology*, 45(1), 47-56.
- Cash, T. F., & Brown, M. A. (2000). Attitudes about antidepressants: influence of information about weight-related side effects. *Perceptual and motor skills*.
- Castaldelli-Maia, J. M., Scomparini, L. B., de Andrade, A. G., Bhugra, D., Alves, T. C. d. T. F., & D'elia, G. (2011). Perceptions of and attitudes toward antidepressants: stigma attached to their use-a review. *The Journal of nervous and mental disease*, 199(11), 866-871.
- Chakraborty, K., Avasthi, A., Kumar, S., & Grover, S. (2009). Attitudes and beliefs of patients of first episode depression towards antidepressants and their adherence to treatment. *Social psychiatry and psychiatric epidemiology*, 44(6), 482-488.
- 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.
- Charan, J., & Biswas, T. (2013). How to Calculate Sample Size for Different Study Designs in Medical Research? *Indian journal of psychological medicine*, 35(2), 121-126.
doi:10.4103/0253-7176.116232
- Chisholm, D. (2015). Investing in mental health. *East Mediterr Health J*, 21(7), 531-534.
- Christensen, A. J. (2004). *Patient adherence to medical treatment regimens: Bridging the gap between behavioral science and biomedicine*: Yale University Press.
- Cohen, S., Kühn, K., Sträter, B., Scherbaum, N., & Weig, W. (2010). Adverse side-effect on sexual function caused by psychotropic drugs and psychotropic substances. *Der Nervenarzt*, 81(9), 1129-1137; quiz 1138.
- Culang - Reinlieb, M. E., Sneed, J. R., Keilp, J. G., & Roose, S. P. (2012). Change in cognitive functioning in depressed older adults following treatment with sertraline or nortriptyline. *International journal of geriatric psychiatry*, 27(8), 777-784.
- Day, J. C., Bentall, R. P., Roberts, C., Randall, F., Rogers, A., Cattell, D., . . . Power, C. (2005). Attitudes toward antipsychotic medication: the impact of clinical variables and relationships with health professionals. *Archives of general psychiatry*, 62(7), 717-724.
- De las Cuevas, C., Peñate, W., & Sanz, E. J. (2014). Risk factors for non-adherence to antidepressant treatment in patients with mood disorders. *Eur J Clin Pharmacol*, 70(1), 89-98.
- De las Cuevas, C., & Sanz, E. J. (2007). Attitudes toward psychiatric drug treatment: the experience of being treated. *Eur J Clin Pharmacol*, 63(11), 1063-1067.
- Demyttenaere, K., Adelin, A., Patrick, M., Walthère, D., & Michèle, S. (2008). Six-month compliance with antidepressant medication in the treatment of major depressive disorder. *International clinical psychopharmacology*, 23(1), 36-42.
- Demyttenaere, K., Bruffaerts, R., Albert, A., Mesters, P., Dewé, W., Debruyckere, K., & Sangeleer, M. (2004). Development of an antidepressant compliance questionnaire. *Acta Psychiatrica Scandinavica*, 110(3), 201-207.

- Dougherty, L. R., Klein, D. N., Olino, T. M., Dyson, M., & Rose, S. (2009). Increased waking salivary cortisol and depression risk in preschoolers: the role of maternal history of melancholic depression and early child temperament. *Journal of Child Psychology and Psychiatry*, 50(12), 1495-1503.
- Edwards, B. J., Bunta, A. D., Lane, J., Odvina, C., Rao, D. S., Raisch, D. W., . . . Garg, V. (2013). Bisphosphonates and nonhealing femoral fractures: analysis of the FDA Adverse Event Reporting System (FAERS) and international safety efforts: a systematic review from the Research on Adverse Drug Events And Reports (RADAR) project. *The Journal of bone and joint surgery. American volume*, 95(4), 297.
- Fawzi, W., Mohsen, M. Y. A., Hashem, A. H., Moussa, S., Coker, E., & Wilson, K. C. (2012). Beliefs about medications predict adherence to antidepressants in older adults. *International Psychogeriatrics*, 24(01), 159-169.
- Ferguson, J. M. (2001). SSRI antidepressant medications: adverse effects and tolerability. *Prim Care Companion J Clin Psychiatry*, 3(1), 22-27.
- Fortney, J. C., Pyne, J. M., Edlund, M. J., Stecker, T., Mittal, D., Robinson, D. E., & Henderson, K. L. (2011). Reasons for antidepressant nonadherence among veterans treated in primary care clinics. *The Journal of clinical psychiatry*, 72(6), 827-834.
- Fox, S., & Jones, S. (2012). The social life of health information. Pew Internet & American Life Project 2009.
- Fung, K. W., Jao, C. S., & Demner-Fushman, D. (2013). Extracting drug indication information from structured product labels using natural language processing. *Journal of the American Medical Informatics Association*, 20(3), 482-488.
- Gabriel, A., & Violato, C. (2010). Knowledge of and attitudes towards depression and adherence to treatment: The Antidepressant Adherence Scale (AAS). *Journal of affective disorders*, 126(3), 388-394.
- Gale, N. K., Heath, G., Cameron, E., Rashid, S., & Redwood, S. (2013). Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Medical Research Methodology*, 13, 117-117. doi:10.1186/1471-2288-13-117
- Ganguly, D., Roy, D., Mitra, M., & Jones, G. J. (2015). *Word embedding based generalized language model for information retrieval*. Paper presented at the Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval.
- Gardner, J. S. (1996). The European Agency for the Evaluation of Medicines and European regulation of pharmaceuticals. *European Law Journal*, 2(1), 48-82.
- Giannangelo, K., Bowman, S. E., Dougherty, M., & Fenton, S. H. (2005). ICF: representing the patient beyond a medical classification of diagnoses. *ICF: Representing the Patient beyond a Medical Classification of Diagnoses/AHIMA, American Health Information Management Association*.
- 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.
- Goldstein, D. J., Lu, Y., Detke, M. J., Wiltse, C., Mallinckrodt, C., & Demitrack, M. A. (2004). Duloxetine in the treatment of depression: a double-blind placebo-controlled comparison with paroxetine. *Journal of clinical psychopharmacology*, 24(4), 389-399.

- Gorman, J. M., Korotzer, A., & Su, G. (2002). Efficacy comparison of escitalopram and citalopram in the treatment of major depressive disorder: pooled analysis of placebo-controlled trials. *CNS spectrums*, 7(S1), 40-44.
- Green, E. C., & Murphy, E. (2014). Health belief model. *The Wiley Blackwell encyclopedia of health, illness, behavior, and society*.
- Greenberg, P. E., Fournier, A.-A., Sisitsky, T., Pike, C. T., & Kessler, R. C. (2015). The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *The Journal of clinical psychiatry*, 76(2), 155-162.
- Grenard, J. L., Munjas, B. A., Adams, J. L., Suttorp, M., Maglione, M., McGlynn, E. A., & Gellad, W. F. (2011). Depression and medication adherence in the treatment of chronic diseases in the United States: a meta-analysis. *J Gen Intern Med*, 26(10), 1175-1182.
- Griffiths, K. M., Cleave, A. L., & Banfield, M. (2009). Systematic review on Internet Support Groups (ISGs) and depression (1): Do ISGs reduce depressive symptoms? *J Med Internet Res*, 11(3).
- Grover, S., Chakrabarti, S., Sharma, A., & Tyagi, S. (2014). Attitudes toward psychotropic medications among patients with chronic psychiatric disorders and their family caregivers. *Journal of neurosciences in rural practice*, 5(4), 374.
- Gurulingappa, H., Klinger, R., Hofmann-Apitius, M., & Fluck, J. (2010). *An empirical evaluation of resources for the identification of diseases and adverse effects in biomedical literature*. Paper presented at the 2nd Workshop on Building and evaluating resources for biomedical text mining (7th edition of the Language Resources and Evaluation Conference).
- Gurulingappa, H., Mateen-Rajput, A., & Toldo, L. (2012). Extraction of potential adverse drug events from medical case reports. *J Biomed Semantics*, 3(1), 15.
- Gurulingappa, H., Rajput, A. M., Roberts, A., Fluck, J., Hofmann-Apitius, M., & Toldo, L. (2012). Development of a benchmark corpus to support the automatic extraction of drug-related adverse effects from medical case reports. *Journal of Biomedical Informatics*, 45(5), 885-892.
- Hallgren, K. A. (2012). Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutorials in quantitative methods for psychology*, 8(1), 23-34.
- Hansson, M., Chotai, J., & Bodlund, O. (2010). Patients' beliefs about the cause of their depression. *Journal of affective disorders*, 124(1), 54-59.
- Harpaz, R., Callahan, A., Tamang, S., Low, Y., Odgers, D., Finlayson, S., . . . Shah, N. H. (2014). Text mining for adverse drug events: the promise, challenges, and state of the art. *Drug safety*, 37(10), 777-790.
- Haslam, C., Atkinson, S., Brown, S., & Haslam, R. (2005). Anxiety and depression in the workplace: effects on the individual and organisation (a focus group investigation). *Journal of affective disorders*, 88(2), 209-215.
- Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication methods and measures*, 1(1), 77-89.
- Health Information Technology for Economic and Clinical Health Act. Retrieved from http://en.wikipedia.org/wiki/Health_Information_Technology_for_Economic_and_Clinical_Health_Act
- Hoencamp, E., Stevens, A., & Haffmans, J. (2002). Patients' attitudes toward antidepressants. *Psychiatr Serv*, 53(9), 1180-a-1181.

- Hofer, A., Kemmler, G., Eder, U., Edlinger, M., Hummer, M., & Fleischhacker, W. W. (2004). Quality of life in schizophrenia: the impact of psychopathology, attitude toward medication, and side effects. *The Journal of clinical psychiatry*, 65(7), 932-939.
- Hogan, T. P., Awad, A., & Eastwood, R. (1983). A self-report scale predictive of drug compliance in schizophrenics: reliability and discriminative validity. *Psychological medicine*, 13(01), 177-183.
- Holma, I. A., Holma, K. M., Melartin, T. K., & Isometsä, E. T. (2010). Treatment attitudes and adherence of psychiatric patients with major depressive disorder: a five-year prospective study. *Journal of affective disorders*, 127(1), 102-112.
- Horne, R., Weinman, J., & Hankins, M. (1999). The beliefs about medicines questionnaire: the development and evaluation of a new method for assessing the cognitive representation of medication. *Psychology and health*, 14(1), 1-24.
- Hung, Y., YlaCHAO, W., & JAUa, P. (2014). The Components of Adherence for Depression Patients. *中華心理衛生學刊*, 27(2), 201-221.
- Hunot, V. M., Horne, R., Leese, M. N., & Churchill, R. C. (2007). A cohort study of adherence to antidepressants in primary care: the influence of antidepressant concerns and treatment preferences. *Primary care companion to the Journal of clinical psychiatry*, 9(2), 91.
- Huynh, T., He, Y., Willis, A., & Rüger, S. (2016). *Adverse drug reaction classification with deep neural networks*.
- Jacob, S. A., Ab Fatah Ab Rahman, M. A., & Hassali, A. (2015). Attitudes and beliefs of patients with chronic depression toward antidepressants and depression. *Neuropsychiatric disease and treatment*, 11, 1339.
- Jin, J., Sklar, G. E., Oh, V. M. S., & Li, S. C. (2008). Factors affecting therapeutic compliance: A review from the patient's perspective. *Therapeutics and clinical risk management*, 4(1), 269.
- Jorm, A. F., Christensen, H., & Griffiths, K. M. (2005). Belief in the harmfulness of antidepressants: results from a national survey of the Australian public. *Journal of affective disorders*, 88(1), 47-53.
- Kane, J. M., Kishimoto, T., & Correll, C. U. (2013). Non - adherence to medication in patients with psychotic disorders: epidemiology, contributing factors and management strategies. *World Psychiatry*, 12(3), 216-226.
- 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.
- Karter, J. (2015). Percentage of Americans on Antidepressants Nearly Doubles. Retrieved from <https://www.madinamerica.com/2015/11/percentage-of-americans-on-antidepressants-nearly-doubles/>
- Kennedy, J., Tuleu, I., & Mackay, K. (2008). Unfilled prescriptions of medicare beneficiaries: prevalence, reasons, and types of medicines prescribed. *Journal of Managed Care Pharmacy*, 14(6), 553-560.
- Keselman, A., Smith, C. A., Divita, G., Kim, H., Browne, A. C., Leroy, G., & Zeng-Treitler, Q. (2008). Consumer health concepts that do not map to the UMLS: where do they fit? *Journal of the American Medical Informatics Association*, 15(4), 496-505.
- Kessing, L. V., Hansen, H. V., Demyttenaere, K., & Bech, P. (2005). Depressive and bipolar disorders: patients' attitudes and beliefs towards depression and antidepressants. *Psychological medicine*, 35(08), 1205-1213.

- Kikuchi, T., Uchida, H., Suzuki, T., Watanabe, K., & Kashima, H. (2011). Patients' attitudes toward side effects of antidepressants: an Internet survey. *European archives of psychiatry and clinical neuroscience*, 261(2), 103-109.
- Kuhn, M., Letunic, I., Jensen, L. J., & Bork, P. (2015). The SIDER database of drugs and side effects. *Nucleic acids research*, 44(D1), D1075-D1079.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, 159-174.
- 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.
- Levine, J., & Schooler, N. (1986). SAFTEE: a technique for the systematic assessment of side effects in clinical trials. *Psychopharmacology bulletin*, 22(2), 343.
- Lin, E. H., Von Korff, M., Katon, W., Bush, T., Simon, G. E., Walker, E., & Robinson, P. (1995). The role of the primary care physician in patients' adherence to antidepressant therapy. *Med Care*, 67-74.
- Lingjaerde, O., Ahlfors, U., Bech, P., Dencker, S., & Elgen, K. (1987). The UKU side effect rating scale: A new comprehensive rating scale for psychotropic drugs and a cross - sectional study of side effects in neuroleptic - treated patients. *Acta Psychiatrica Scandinavica*, 76(s334), 1-100.
- 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.
- López-Torres, J., Párraga, I., Del Campo, J. M., & Villena, A. (2013). Follow up of patients who start treatment with antidepressants: treatment satisfaction, treatment compliance, efficacy and safety. *BMC Psychiatry*, 13(1), 1.
- Malpass, A., Shaw, A., Sharp, D., Walter, F., Feder, G., Ridd, M., & Kessler, D. (2009). "Medication career" or "Moral career"? The two sides of managing antidepressants: A meta-ethnography of patients' experience of antidepressants. *Soc Sci Med*, 68(1), 154-168.
- Mao, J. J., Chung, A., Benton, A., Hill, S., Ungar, L., Leonard, C. E., . . . Holmes, J. H. (2013). Online discussion of drug side effects and discontinuation among breast cancer survivors. *Pharmacoepidemiology and drug safety*, 22(3), 256-262.
- Masand, P. S. (2003). Tolerability and adherence issues in antidepressant therapy. *Clinical therapeutics*, 25(8), 2289-2304.
- Masand, P. S., Roca, M., Turner, M. S., & Kane, J. M. (2009). Partial adherence to antipsychotic medication impacts the course of illness in patients with schizophrenia: a review. *Prim Care Companion J Clin Psychiatry*, 11(4), 147-154.
- McHugh, R. K., Whitton, S. W., Peckham, A. D., Welge, J. A., & Otto, M. W. (2013). Patient preference for psychological vs. pharmacological treatment of psychiatric disorders: a meta-analytic review. *The Journal of clinical psychiatry*, 74(6), 595.
- Medeffect Canada. MedEffect Canada [Internet]. Retrieved from <http://www.hc-sc.gc.ca/dhp-mps/medeff/index-eng.php>
- 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.
- Moncrieff, J., Cohen, D., & Mason, J. (2009). The subjective experience of taking antipsychotic medication: a content analysis of Internet data. *Acta Psychiatrica Scandinavica*, 120(2), 102-111.
- Murata, A., Kanbayashi, T., Shimizu, T., & Miura, M. (2012). Risk factors for drug nonadherence in antidepressant-treated patients and implications of pharmacist adherence instructions for adherence improvement. *Patient Prefer Adherence*, 6, 863-869.
- Murray, C. J., & Lopez, A. D. (1997). Alternative projections of mortality and disability by cause 1990–2020: Global Burden of Disease Study. *The Lancet*, 349(9064), 1498-1504.
- NCCMH. (2010). *Depression: the treatment and management of depression in adults (updated edition)*.
- Nemeroff, C. B., Heim, C. M., Thase, M. E., Klein, D. N., Rush, A. J., Schatzberg, A. F., . . . Dunner, D. L. (2003). Differential responses to psychotherapy versus pharmacotherapy in patients with chronic forms of major depression and childhood trauma. *Proceedings of the National Academy of Sciences*, 100(24), 14293-14296.
- Ng, C. H., Smith, D. J., King, J., Ong, S., & Schweitzer, I. (2012). Medication attitudes and beliefs in patients with psychotic and affective disorders on maintenance treatment. *Human Psychopharmacology: Clinical and Experimental*, 27(1), 57-62.
- 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*, oeu041.
- Nunkoosing, K. (2005). The problems with interviews. *Qual Health Res*, 15(5), 698-706.
- Olfson, M., & Marcus, S. C. (2009). National patterns in antidepressant medication treatment. *Archives of general psychiatry*, 66(8), 848-856.
- Orgeta, V., Qazi, A., Spector, A., & Orrell, M. (2015). Psychological treatments for depression and anxiety in dementia and mild cognitive impairment: systematic review and meta-analysis. *The British Journal of Psychiatry*, 207(4), 293-298.
- Price, J., Cole, V., & Goodwin, G. M. (2009). Emotional side-effects of selective serotonin reuptake inhibitors: qualitative study. *The British Journal of Psychiatry*, 195(3), 211-217.
- Prins, M. A., Verhaak, P. F., Bensing, J. M., & van der Meer, K. (2008). Health beliefs and perceived need for mental health care of anxiety and depression—The patients' perspective explored. *Clin Psychol Rev*, 28(6), 1038-1058.
- Python. (2017). Python Random Module. Retrieved from <https://www.programiz.com/python-programming/modules/random>
- Reilly JL, L. R., Thase ME, et al. (2011). Effects of sertraline on neurocognition in outpatients with major depressive disorder. *Neuropsychopharmacology*, 36(suppl 1), S117–S118.
- Richardson, M., McCabe, R., & Priebe, S. (2013). Are attitudes towards medication adherence associated with medication adherence behaviours among patients with psychosis? A systematic review and meta analysis. *Social psychiatry and psychiatric epidemiology*, 48(4), 649-657.

- Rivero-Santana, A., Perestelo-Perez, L., Pérez-Ramos, J., Serrano-Aguilar, P., & De las Cuevas, C. (2013). Sociodemographic and clinical predictors of compliance with antidepressants for depressive disorders: systematic review of observational studies. *Patient Preference Adherence*, 7, 151-169.
- Russell, J., & Kazantzis, N. (2008). Medication beliefs and adherence to antidepressants in primary care. *The New Zealand Medical Journal (Online)*, 121(1286).
- Sajatovic, M., Velligan, D. I., Weiden, P. J., Valenstein, M. A., & Ogedegbe, G. (2010). Measurement of psychiatric treatment adherence. *Journal of psychosomatic research*, 69(6), 591-599.
- Salem, H., Nagpal, C., Pigott, T., & Lucio Teixeira, A. (2017). Revisiting Antipsychotic-induced Akathisia: Current Issues and Prospective Challenges. *Current neuropharmacology*, 15(5), 789-798.
- Sansone, R. A., & Sansone, L. A. (2010). SSRI-Induced Indifference. *Psychiatry (1550-5952)*, 7(10).
- Sansone, R. A., & Sansone, L. A. (2012). Antidepressant adherence: are patients taking their medications? *Innovations in clinical neuroscience*, 9.
- Santarsieri, D., & Schwartz, T. L. (2015). Antidepressant efficacy and side-effect burden: a quick guide for clinicians. *Drugs in context*, 4.
- Sarker, A., Ginn, R., Nikfarjam, A., O'Connor, K., Smith, K., Jayaraman, S., . . . Gonzalez, G. (2015). Utilizing social media data for pharmacovigilance: A review. *Journal of Biomedical Informatics*, 54, 202-212.
- 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.
- Sayyah, M., Eslami, K., AlaiShehni, S., & Kouti, L. (2016). Cognitive Function before and during Treatment with Selective Serotonin Reuptake Inhibitors in Patients with Depression or Obsessive-Compulsive Disorder. *Psychiatry Journal*, 2016.
- Scheppers, E., Van Dongen, E., Dekker, J., Geertzen, J., & Dekker, J. (2006). Potential barriers to the use of health services among ethnic minorities: a review. *Family practice*, 23(3), 325-348.
- Smith, B. L. (2012). Inappropriate prescribing. *Monitor on Psychology*, 43(6), 36.
- Smith, D., Dempster, C., Glanville, J., Freemantle, N., & Anderson, I. (2002). Efficacy and tolerability of venlafaxine compared with selective serotonin reuptake inhibitors and other antidepressants: a meta-analysis. *The British Journal of Psychiatry*, 180(5), 396-404.
- Taylor, S. A., & Cronin Jr, J. J. (1994). Modeling patient satisfaction and service quality. *Marketing Health Services*, 14(1), 34.
- Thase, M. E., Entsuah, A. R., & Rudolph, R. L. (2001). Remission rates during treatment with venlafaxine or selective serotonin reuptake inhibitors. *The British Journal of Psychiatry*, 178(3), 234-241.
- Townsend, L., Floersch, J., & Findling, R. L. (2009). Adolescent attitudes toward psychiatric medication: the utility of the Drug Attitude Inventory. *Journal of Child Psychology and Psychiatry*, 50(12), 1523-1531.
- Uher, R., Farmer, A., Henigsberg, N., Rietschel, M., Mors, O., Maier, W., . . . Placentino, A. (2009). Adverse reactions to antidepressants. *The British Journal of Psychiatry*, 195(3), 202-210.

- Verdoux, H., Lengronne, J., Liraud, F., Gonzales, B., Assens, F., Abalan, F., & Os, J. v. (2000). Medication adherence in psychosis: predictors and impact on outcome. A 2 - year follow - up of first - admitted subjects. *Acta Psychiatrica Scandinavica*, 102(3), 203-210.
- Vergouwen, A. C., Burger, H., Verheij, T. J., & Koerselman, F. (2009). Improving patient's beliefs about antidepressants in primary care: a cluster-randomized controlled trial of the effect of a depression care program. *Prim Care Companion J Clin Psychiatry*, 11(2), 48-52.
- Vives, M., Mateu, C., Vicens, C., Molina, R., Puebla-Guedea, M., & Gili, M. (2014). Therapeutic adherence in primary care depressed patients: A longitudinal study. *Actas Esp Psiquiatr*, 42(3), 91-98.
- Weiden, P., Rapkin, B., Mott, T., Zygmunt, A., Goldman, D., Horvitz-Lennon, M., & Frances, A. (1994). Rating of medication influences (ROMI) scale in schizophrenia. *Schizophrenia Bulletin*, 20(2), 297-310.
- Williams, B. (1994). Patient satisfaction: a valid concept? *Soc Sci Med*, 38(4), 509-516.
- Wishart, D. S., Knox, C., Guo, A. C., Shrivastava, S., Hassanali, M., Stothard, P., . . . Woolsey, J. (2006). DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic acids research*, 34(suppl_1), D668-D672.
- Wolters, H. A., Kneegting, H., van den Bosch, R. J., & Wiersma, D. (2009). Effects and side effects of antipsychotic treatment in schizophrenia: pros and cons of available self-rating scales. *Schizophrenia research*, 112(1), 114-118.
- Woolley, S. B., Fredman, L., Goethe, J. W., Lincoln, A. K., & Heeren, T. (2010). Hospital patients' perceptions during treatment and early discontinuation of serotonin selective reuptake inhibitor antidepressants. *Journal of clinical psychopharmacology*, 30(6), 716-719.
- World Health Organization. (2003). Investing in mental health.
- 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.

Appendix A:
Self-report questionnaire designed/used for measuring ADRs associated with antidepressants

The Antidepressant Side-Effect Checklist (ASEC) was constructed as a self-report scale to measure twenty-one physiological and subjective Antidepressants' adverse reactions. For each item, a patient can rate the severity of the specified adverse effect on a four-point scale (0 absent; 1 mild; 2 moderate; 3 severe). The patients can also specify whether a symptom (if present) is likely to be linked to antidepressants adverse effects (yes or no). The total score can vary between "0", if patients marked all symptoms as absent and "63", if patient marked all symptoms as severe and specified that they are highly related to antidepressants.

The Liverpool University Neuroleptic Side Effect Rating Scale (LUNSERS) is a fifty-one items that was primarily constructed to measure adverse effects of antipsychotic drugs. Although this scale was not specifically designed for antidepressants' adverse effects, it covers some of the psychological, emotional, and cognitive adverse reactions of the medications. The adverse effects in LUNSERS scale are classified in eight classes including extra-pyramidal, psychic, anticholinergic, autonomic, allergic reactions, hormonal, Miscellaneous (such as weight gain), and red hearings. The aim of incorporating read hearing category is to detect over-reporting cases. For each item, a participant can rate the severity of the specified adverse reaction on a four-point scale (0: not at all; 1: Very Little; 2: A Little; 3: Quite a Lot ; 4: Very Much"). The total score of LUNSERS can vary between "0" if patients market all specified ADR as "0" and "204" if marked all specified "ADRs" as "4".

Studies that were also interested in measuring the impact of the antidepressants' adverse effects on quality of life employed the WHOQOL-BREF. The WHOQOL-BREF assesses the physical, psychological, social, and environmental aspects of quality of life. In this

questionnaire, patients can rate level of satisfaction, frequency, or magnitude of each item on a five- point scale.

Ng et al. (2012) used the LUNSERS, WHOQOL-BREF, and also BAS and SAS scales to assess association between attitude toward antidepressants and ADRs among patients with psychotic and affective disorders such as schizophrenia and depression. Their findings indicated that there is a negative correlation between LUNSERS and DAI/BEMID scores (scales for measuring attitude) and attitude toward antidepressants, implying that the increased reported side effects are associated with poorer attitudes towards antidepressants. But they did not answer the question of whether there is an association between scores of WHOQOL-BREF, SAS, and BAS and patients' attitudes toward antidepressants.

To my knowledge, there have been no studies that applied ASEC to assess the association between patients' attitudes and adverse effects of antidepressant. However, De las Cuevas et al. (2014) used this scale to find out the association between adherence to antidepressants and side-effects. Their findings indicated that non-adherent patients reported more frequent and more intense adverse effects than adherent patients. In addition, adherence and non-adherence groups were significantly different in four reported adverse effects including "dry mouth", "diarrhea", "feeling like the room is spinning", and "weight gain".

In contrast to the studies using self-report scales to identify antidepressants adverse effects, Murata et al. (2012) directly asked patients to report their feelings and their uncomfortable experiences (subjective experiences) with antidepressants at the point of care. In addition, the study used tracking of patients' vital and physical signs to find out potential adverse effects. Their findings indicated that sleepiness, malaise, dry mouth, constipation, dysuria,

dizziness, cephalalgia, hidrosis, and anorexia were significant adverse effects that may lead to negative attitude toward antidepressants.

Although Uher et al. (2009) demonstrated that adverse effects to antidepressants can be reliably assessed through self-report scales, the inherent limitations associated with the scales such as not including all potential adverse effects, particularly the rare ones, may reduce the reliability and validity of the scales. For example, the ASEC scale does not cover adverse effects indicating emotional problems, cognitive dysfunctions, or quality of life. This is a major limitation for studies employing this scale as a self-report measure for antidepressants adverse reactions (Bet et al., 2013; Uher et al., 2009). The LUNSERS is also associated with some limitations. First, it is not specifically designed for measuring the antidepressants' adverse effects. Second, items indicating emotional problems and cognitive dysfunctions are very limited and do not reflect the wide range of emotional and cognitive reactions linked to all agents and classes.

Moreover, items in the WHOQOL-BREF assess individual quality of life from a general perspective that may not be directly related to the impact of antidepressants on quality of life. For examples, items such as "level of satisfaction with living place or having enough money to support your needs" do not measure the impacts of antidepressants. Also, other items such as "satisfaction with interpersonal relationships" or "level of performance in daily activities" do not provide specific information about patients' incapacities resulted from antidepressants usage. For example, patients may report difficulty with driving or exercising because of "brain shivers" that was not specified in the WHOQOL-BREF. The limitations with WHOQOL-BREF can lead to an insufficient understanding of the dimensions of antidepressants' adverse effects on a patient's quality of life. These scales may also suffer from patient recall bias, because they do not define a

specific time period during which the adverse effects have been occurred. Hence, patients may just report severe adverse effects they experienced and disregard the mild ones(Bet et al., 2013).

Appendix B: Self-Report Scales Used in Antidepressants Studies for Measuring Attitude

a) Drug Attitude Inventory (DAI)

The DAI consists of seven components that mainly measure attitude and adherence behavior towards prescribed psychotic medications mainly in the maintenance phase of treatment. The DAI assesses patient attitudes based on two major dimensions: Patients' subjective experience and the perceived necessity of medication and their concern towards psychotic medications. Hogan et al. (1983) concluded that patients' attitudes towards psychotic medications are highly dependent on patient's subjective experiences during a course of treatment.

Table Appendix B.1. Drug Attitude Inventory (DAI): summary of components' items

Subjective Experience	Components 1 and 2 include items measuring a patient's subjective experiences of medications. These components reflect the balance of a patient's perceived benefits of medication and perceived distress from adverse effects. Items in these components are formulated to assess impact of medication on cognitive functionalities, Interpersonal relationship, and emotional status.
Necessity and concern	Components 6 and 7 measure a patient's perceived necessity of medications in forestalling relapse and the concern about the potential toxic effects.
Locus of control	Items in Component 4 and 5 express a patient's perception of locus of control in taking medication, whether the locus of control is the physician's authority or a lack of free choice for pharmacological treatment. Items in this component are generally important to determine adherence behavior.
Adherence behavior	Items in component 5 indicate a patient's understanding of adherence behavior. Patients with an erroneous view towards adherence behavior may discontinue medication when they feel better or take more medication when they feel worse.

The ADI has two versions: 30-item and 10-item attitudinal scales that were generally employed by studies focused on attitudes towards antipsychotic medications. However, it was also used occasionally for assessing attitudes toward antidepressant medications.

Ng et al. (2012) used the DAI scale to compare outpatient's attitudes on the maintenance phase of treatment among patients with psychotic and affective disorders such as schizophrenia and depression. Their findings indicated that patients with psychotic disorders do not have more

negative attitudes and beliefs compared to those with affective disorder, implying that the type of mental illness is not a significant predictive factor for attitude. They also concluded that the patients reported side effects have significant correlation with medication attitude and belief but not with education, duration or severity of illness.

Murata et al. (2012) used the DAI tool to investigate different attitudes among both antidepressant-naïve (patients were not treated by antidepressants) and antidepressant-treated patients. They found that the mean DAI -10 total score is significantly lower in antidepressant-treated patients compared to antidepressant-naïve patients. Their findings also suggested that patients' demographic variables, such as age, gender, or education are not significantly associated with attitude score. However, reported medication side effects and type of depression (melancholic, nonmelancholic, bipolar) were significantly correlated with negative attitude.

Townsend, Floersch, and Findling (2009) employed DAI to measure adolescents' attitudes towards psychiatric medications. The adolescents diagnosed with a wide range of psychiatric disorders including major depression. The findings of this study suggested that impact of psychotic medications on emotional status, cognitive functioning, and interpersonal relationship is strongly correlated with adolescents' attitudes toward these medications.

(De las Cuevas & Sanz, 2007) used DAI and to compare stable psychiatric outpatients attitude toward psychiatric medications with public opinion on this subject. The result of this study showed that receiving beneficial aspects of this medications leads to forming positive attitude in psychiatric patients compared with the attitude of general population without experience of this medication. This study also implies the stigma in society against psychiatric medications.

b) Antidepressants Compliance Questionnaire (ADCQ)

The ADCQ (Demyttenaere et al., 2004) is self-report scale that is primarily designed to assess attitudes of psychiatric patients towards antidepressant medications. The ADCQ assesses patients' attitudes towards antidepressants based on four dimensions: perceived patient-clinician relationship, concern about the mechanism of antidepressants on personality, subjective experience of antidepressant usage, and perceived family support. ADCQ has 33 items that were divided into four components. Table 2 presents a summary of the four components.

Table Appendix B.2. Antidepressant Compliance Questionnaire (ADCQ): Summary of components' items

Perceived patient-clinician relationship	Perceived Patient-Clinician Relationship assesses the patient's perception of communication with clinicians. Items in this component mainly indicate three dimensions of patient-physicians interaction including: 1) patient's perception of physician's knowledge; 2) patient's perception of sufficiency of knowledge provided by physicians about the disorder and treatment process; 3) patient's perception of communication effectiveness, such as his perception of clinician's interest in the patient's problem and the level of support received from clinicians.
Perceived Concern (Preserved autonomy)	Preserved autonomy includes items indicating a patient's concern with the antidepressant mechanism in the pre-treatment phase. Items in this component are the additive possibility of antidepressants, control over feelings and thought, altering personality, and immunity to antidepressants.
Subjective Experience (Perceived Benefits from Medications) And Perception of Adherence	Positive Belief about Antidepressants component assesses the <u>patient's perception of antidepressants' helpfulness in improving his coping mechanism, improving emotional status</u> and removing causes (symptoms) of depression. Some of items in this component also indicate patient's understanding of adherence behavior. Patients with erroneous views about antidepressant behavior may think they can take additional dosages or they can skip a dose if they feel better.
Patients' Partner Support	Items in this component express the perceived support that a patient received from his/her partner or family. Whether family members agree with the diagnosis or the pharmacological treatment of depression influences a patient's perception of antidepressants.

Several studies administered the ADCQ to patients with depression to analyze patients' attitude toward antidepressants based on the four dimensions proposed in this scale. The findings of studies (Chakraborty et al., 2009; Hung et al., 2014; Jacob et al., 2015) applied ADCQ showed that the majority of patients with depression reported a positive experience in the patient-physician relationship. Hung et al. (2014) found that women had a more negative experience than men regarding their patient-physician relationship. Regarding the age factor, Chakraborty et al. (2009) and (Hung et al., 2014) did not agree on the correlation between age and the patient-

physician relationship. While Chakraborty et al. (2009) reported older patients had a more negative view of patient-physician relationship, Hung et al. (2014) reported younger patients had lower perception of the patient-physician relationship.

Most of the research administering the ADCQ to patients with depression, they found that a majority of patients believe antidepressants do not preserve their autonomy. Patients believe that by being on antidepressants have less control over their thoughts and feelings (Kessing et al., 2005). They also feel that antidepressants can reduce their alertness (Demyttenaere et al., 2004; Kessing et al., 2005). Patients also have concern about the impact of antidepressants on their personality (Chakraborty et al., 2009; Jacob et al., 2015; Kessing et al., 2005). In addition, They have concern about addictive possibilities (Chakraborty et al., 2009; Jacob et al., 2015; Prins et al., 2008) and immunity to antidepressants (Kessing et al., 2005). Using the ADCQ, Jacob et al. (2015) found that patients have misperception about the cause of depression. Patients may attribute the cause of depression to non-biological factors, such as family issues or a stressful life. Moreover, Chakraborty et al. (2009) found majority of patients on antidepressants reported that their partners or families have a positive attitude towards diagnosis and treatment.

c) Beliefs about Medications Questionnaire (BMQ)

The Beliefs about Medications Questionnaire (BMQ) (Horne et al., 1999) is a Necessity-Concern Framework that assess patient attitude towards pharmacological treatment in general and also for a specific prescribed medication. Items representing patients' necessity and concern are organized in four components: general necessity, general concern, specific necessity, and specific concern.

Table Appendix B.3. Beliefs about Medications Questionnaire (BMQ): a summary of components' items

General necessity	Items in the general necessity components are formulated to assess patient beliefs about the necessity of pharmacological treatment in curing illness. Patients with negative belief may feel that natural remedies are better than medications or that they should not take medications continuously.
General Concern	Items in this component indicate patients' general concerns about pharmacological treatment based on beliefs about harmful aspects such as addiction, harm, and poison caused by medications and concern about overuse of medication by physicians.
Specific necessity	Items in this component express a medication's effectiveness and the patient's attitude about dependency to the prescribed medication as well as their concern about their future healthcare status and their quality of life if they continue using the prescribed medication.
Specific Concern	Items in this category mainly indicate three dimensions of specific concern of patients including: 1) Long-term effects and its impact on quality of life; 2) The mechanism and addictive possibilities of antidepressants; and 3) losing autonomy because of the control of the medication.

Aikens et al. (2005) used this BMQ scale to develop a categorical perspective that is more practical to clinicians. The categories are: skeptical (low necessity, high concerns), ambivalent (high necessity, high concerns), indifferent (low necessity, low concerns), accepting (high necessity, low concerns). Based on this framework adherence is lowest in the accepting group, and lowest in the skeptical group (Aikens et al., 2005; Russell & Kazantzis, 2008).

Brown et al. (2005) used the BMQ to find the relationship between beliefs and adherence to antidepressant medication. The findings of this study suggested that a great number of patients have concern about the long-term effects and addictive possibilities of antidepressants in specific, and about the overuse and harmfulness aspects of medications in general (Brown et al., 2005).

Aikens et al. (2008) employed BMQ to identify factors affecting beliefs about necessity and concern of antidepressants. This study findings shows that perceived necessity is correlated with age, severity of symptoms, anticipated duration of symptoms, and source of depression (attribution of symptoms to chemical imbalance vs. random factors). Perceived concern was higher in patients using antidepressants for the first time (who are new to antidepressants), had a

sufficient knowledge about antidepressants and attribute the source of depression to random factors rather than biological factors. Patient gender, education, age, and type of depression were not significantly associated with necessity and concern of antidepressants ” (Aikens et al., 2008).

d) Rating of Medication Influence Scale (ROMI)

ROMI (Weiden et al., 1994) is a self-report scale that was primarily developed for assessment of attitudinal and behavioral factors affecting outpatient adherence with neuroleptic disorder . This scale assess patients attitude and adherence in seven major domains that we reorganize them in six main components: 1) subjective experience, 2) perceived necessity, 3) perceived psychosocial support, 4) perception about illness, 5) perceived support from healthcare provider, and 6) features of healthcare services.

Table Appendix B.4. Rating of Medication Influence Scale (ROMI): Summary of Components' items

Subjective experience and adverse effects	Two items in the domain of therapeutic regimen features including treatment effectiveness and perceived distress of psychotic medications indicate subjective experience of patients with antidepressants.
Perceived necessity (prevention)	Three items in the domain of patients characteristics represents patients belief about the necessity of the medications, whether patient believes his/her current and future healthcare status depends on medication. Fear of relapse, fear of hospitalization, and perceived recovery from illness are the items.
Perception about illness	Items in the domain of disease feature and psychological factors indicate patient perception about illness. The items assess patient perception about existence and severity of diseases, and whether patient feels stigma/embarrassment against illness. Item of substance abuse in the category patient characteristic might reflect patients attitude about illness and medication.
Perceived psychosocial support (Influence of others)	Items in the domain of psychosocial factors domain evaluate patients' perceived support received from family and also family attitude toward medication.
Perceived physician-patient relationship	Items in the category of patient-physician interaction and also treatment system features assess patients perceived relationship with physicians and whether they felt any pressure/force from healthcare providers.
Features of healthcare services (adherence behavior)	Items in domain of treatment system feature and therapeutic regimens features evaluate indicate feature of healthcare services, whether healthcare services is accessible, affordable, it is structured, and physicians believe to medications.

In studies pertain to adherence and attitudes towards antidepressants, ROMI scale primarily has been applied in research related to adherence (Sajatovic, Velligan, Weiden, Valenstein, & Ogedegbe, 2010; Sansone & Sansone, 2012).

There are other scales, such as *Brief Evaluation of Medication Influences and Beliefs* (BEMIB), *Medication Adherence Rating Scale* (MARS) that includes items for assessing patient attitude, but they are mainly developed for measuring and identifying predictive factors influencing adherence.

Studies used these scales has yielded two De las Cuevas and Sanz (2007) De las Cuevas and Sanz (2007) or multi-factor structure Alekhya et al. (2015) as outcome attitude. Two-factor structure simply classifies individuals' attitudes as either positive or negative, while multi-factor structure, such as multi-point Likert scale provides more information about the effect of different factors on forming patients' attitudes by classifying patients in more than 2 categories. Some studies may treat the scales outcome as a continuous variables to calculate correlation of other continues variable, such as correlation of adverse effects with attitude in a study conducted by Day et al. (2005).

Curriculum Vitae
MARYAM ZOLNOORI

EDUCATION

PhD. University of Wisconsin-Milwaukee, Milwaukee, WI 2015- 2017

Department of Health Sciences, Biomedical and Health Informatics Program

Dissertation: *Utilizing Consumer Health Posts for Pharmacovigilance: Identifying underlying factors associated with patients' adherence towards antidepressants*

GPA: 3.96/4

M.Sc. Indiana University School of Informatics and Computing, Indianapolis, IN 2012-2015
Department of BioHealth Informatics, Health Informatics Program

Thesis: *Building a Framework for a Portal to Improve Access of Cancer Survivors to Health Information*

GPA: 3.97/4

M.Sc. Tarbiat Modares University, Tehran, Iran 2004-2007

Department of Information Technology

Thesis: *Developing a Fuzzy Expert System for Screening and Managing Pediatric Asthma*

GPA: 3.43/4

B.Sc. Tehran University, Tehran, Iran 2001-2004
Department of Management

Final Project: *Strategic Planning: A Framework for Internal Analysis of SMEs*

GPA: 3.7/4

PROFESSIONAL APPOINTMENTS / EXPERIENCE

(See Appendix A for details of all research projects.)

- 5/2016—12/2016 **Research Fellow**, Lister Hill National Center for Biomedical Communications, National Institute of Health
Primary Project: Social Media text mining and statistical analysis for pharmacovigilance
- 1/2015—5/2016 **Research Fellow and Lecturer**, Biomedical Data and Language Processing Center of the Department of Health Sciences at the University of Wisconsin-Milwaukee
- 9/2012—12/2015 **Research Assistant**, Indiana University School of Informatics and Computing, Department of BioHealth Informatics.
- 1/2011—8/2012 **Informatics Specialist**, Immunology & Asthma & Allergy Research

	Institute (IAARI), Tehran, Iran
2009—2012	Research Specialist , Academic Center for Education, Culture, and Research (ACECR) Tehran, Iran
2009—2010	IT Consultant , Negahe Nou Company and Jahan Rayane Company
2007—2009	System Analyst, Director of Analytics Department , E-Sabz Company
2005—2007	Tarbiat Modares University , Research Assistant

SKILLS

Coding systems: The Unified Medical Language System (UMLS)
Controlled Terminologies: SNOMED-CT, RX-NORM, MEDRA, ICD-10, DSM, NANDA, NIC, NOC, ICF
Statistical Software: SAS, R, Matlab, Python/Pandas
Text mining and Machine learning: R, Python/NLTK, Scikit-learn
Database: Mysql, Oracle, Access
Programming: XML, PHP, HTML & CSS, Python

COURSE AUDIT

- a) Natural Language Processing and Machine Learning- School of Informatics and Computing, Indiana University
- b) Algorithm Design and Implantation, School of Computer Sciences, Indiana University

PUBLICATIONS

1. Zolnoori, M., Fung, KW., Patrick, T.B., Fontelo P., Kharrazi, H., Faiola, A., Wu, YS. S., Xu, K., Conway, M., Luo, J. (*In Manuscript*). A systematic approach for developing a corpus of Adverse Drug Events for psychiatric medications, **Submitted to: Journal of Biomedical Informatics**. †
- 4)
2. Zolnoori, M., Fung, KW., Fontelo, P., Patrick, T. B., and Faiola, A., Wu YS. S. (*In Manuscript*). Utilizing social media (consumer health posts) to identify underlying factors associated with negative attitudes towards antidepressants, *Slated Journal: Journal of the American Medical Informatics Association*. †
3. Zolnoori, M., Fung, KW., Fontelo, P. (in Manuscript) Real-world data on SSRI and SNRI antidepressant Adverse effects: *Slated Journal: Psychiatry*
4. Zolnoori, M., Fung KW., Fontelo, P., Patrick, T. B., Faiola, A., Using the framework method for the analysis of patient's drug reviews in social media, *Slated Journal: BMC Medical Research Methodology*

5. Zolnoori, M., Patrick, T. B., Faiola, A., Stoffel, V., Fontelo P., and Fung, KW. (*In Manuscript*) Underlying factors affecting attitude and adherence toward antidepressants: a literature review, *Slated Journal: Journal of Social Psychiatry and Psychiatric Epidemiology*.
6. Zolnoori, M., Fung, KW., Fontelo, P., Patrick, T. B., and Faiola, A. Designing automatic text classification for detecting patients with withdrawal experience, *Slated Journal: Journal of Social Psychiatry and Psychiatric Epidemiology*.
7. Zolnoori, M., Fung, KW., Fontelo, P., Patrick, T. B., and Faiola, A., Wu YS. S. (*In Manuscript*). Utilizing social media (consumer health posts) to identify underlying factors associated with patients withdrawal from antidepressants, *Slated Journal: Journal of the Mental Health*.

Peer-Reviewed Journals

1. Alizadeh, B., Safdari, R., Zolnoori, M., & Bashiri, A. (2015). Developing an Intelligent System for Diagnosis of Asthma Based on Artificial Neural Network. *Acta Informatica Medica*, 23(4), 220. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4584092/>
2. Samad Soltani, T., Langerizadeh, M., & Zolnoori, M. (2015). Data Mining And Analysis: Reporting Results For Patients With Asthma. *Payavard Salamat*, 9(3), 224-234. http://payavard.tums.ac.ir/browse.php?a_id=5729&sid=1&slc_lang=en
3. Darabi, S. A., Teimourpour, B., & Zolnoori, M. (2014). Case-Based-Reasoning System for Feature Selection and Diagnosing Disease; Case Study: Asthma. *Innovative Systems Design and Engineering*, 5(5), 43-59. <http://www.iiste.org/Journals/index.php/ISDE/article/view/12915>
4. Zolnoori, M., Jones, J. F., Moin, M., Heidarnejad, H., Fazlollahi, M. R., & Hosseini, M. (2013). Evaluation of user interface of computer application developed for screening pediatric asthma. In *Universal Access in Human-Computer Interaction. Applications and Services for Quality of Life* (pp. 563-570). Springer Berlin Heidelberg, http://link.springer.com/chapter/10.1007/978-3-642-39194-1_65
5. TahaSamadSoltaniHeris, M., Mahmoodvand, Z., & Zolnoori, M. (2013). Intelligent Diagnosis of Asthma Using Machine Learning Algorithms. *International Research Journal of Applied and Basic Sciences*. http://www.irjabs.com/files_site/paperlist/r_1245_130901231500.pdf
6. Zolnoori, M., Zarandi, M. H. F., Moin, M., Heidarnejad, H., & Kazemnejad, A. (2012). Computer-aided intelligent system for diagnosing pediatric asthma. *Journal of medical systems*, 36(2), 809-822. <http://www.ncbi.nlm.nih.gov/pubmed/20703652>

7. Zolnoori M., Fazel Zarandi M.H., Moin M. (2012). "Fuzzy Rule-Based Expert System for Evaluating Level of Asthma Control, *Journal of Medical Systems*", DOI: 10.1007/s10916-011-9773-3, <http://www.springerlink.com/content/7540476248388854/>
8. Zolnoori M., Fazel Zarandi M. H., Moin M. (2012). "Application of Intelligent Systems in Asthma Disease: Designing a Fuzzy Rule-Based System for Evaluation Level of Asthma Exacerbation" *Journal of Medical Systems*, DOI: 10.1007/s10916-011-9671-8, <http://www.springerlink.com/content/j0xu834740670748/>
9. Fazel Zarandi M.H., Zolnoori M., Moin M., Heidarnejad H. (2010). "Fuzzy Expert System for Diagnosing Asthma", *Journal of Scientia Iranica: Transaction E: Industrial Engineering*, 17 (2), 129-142, http://www.sid.ir/En/VEWSSID/J_pdf/9552010E205.pdf
10. Zolnoori M., Fazel Zarandi M.H., Moin M., Teimorian S. (2012). "Fuzzy Rule-Based Expert System for Assessment Severity of Asthma", *Journal of Medical Systems*, Dec 3., DOI: 10.1007/s10916-010-9631 <http://www.springerlink.com/content/e14g62611t2407h3/>
11. Zolnoori M., Fazel Zarandi M.H., Moin M. (2011). "Fuzzy Rule-Based Expert System for Evaluation Possibility of Fatal Asthma", *Journal of Health Informatics in Developing Countries*, 5(1), 171-184, <http://www.jhdc.org/index.php/jhdc/article/viewArticle/59>

Peer-Reviewed Conference Proceedings—Long Papers

12. Zolnoori, M., Patrick, T. B., Fung, K.W., Fontelo P., Faiola, A., Wu YS. S., Xu, K., Zhu, J., Eldredge, C. E. Development of an Adverse Drug Reaction Corpus from Consumer Health Posts. *Workshop: American Medical Informatics Association, Annual Symposium*, Washington, DC., November, 2017. <https://healthlanguageprocessing.org/smm4hprogram/>
13. Elahi.G., Shahrivari S., Zolnoori M. (December 2011) *Comparison the results of linear (Discriminative Analysis, and Logistic Regression) and non-linear (SVM, Neural Network, and Decision Tree) Techniques in evaluating credit of Banks' Customer (in Persian)*. 5th National Conference on Data Mining, Tehran, Iran. <http://www.iranai.ir/?cat=12>
14. Zolnoori M., Shahrivari S., Faliat V. (February 2011) *Strategic Planning: a Framework for Internal Analysis*, This paper presented at 5th International Conference on Strategic Management, Tehran, Iran, http://www.civilica.com/EnPaper-ICSM05-ICSM05_116.html
15. Zolnoori M., Fazel Zarandi M.H., Moein M. (March 2010) *Intelligent Systems in Medicine: Designing a Fuzzy Expert System for Evaluation Severity of Asthma Exacerbation (in Persian)*, Paper presented at 6th International Conference on information and communication technology, Tehran, Iran http://www.civilica.com/Paper-ICTM06-ICTM06_030.html

16. Zolnoori M., Fazel Zarandi M.H., Moein M. (July 2010) *Designing an Expert System for Evaluating Possibility of Fatal Asthma*, This paper presented at 10th Iranian Conference on Fuzzy Systems, July 2010, Tehran, Iran, <http://ifs10.sbu.ac.ir/Files/Proceeding.pdf>, page 348/493
17. Faaliate V., Zolnoori M. *Is It Possible to Create Software Clusters in Iran? (in Persian)*, (February 2010) 6th International Conference on Information Technology Management, Tehran, Iran, http://www.civilica.com/Paper-ICTM06-ICTM06_052.html
18. Bohrani A., Zolnoori M. (February 2010), *A systematic approach for designing national portal" (in Persian)*, 6th International Conference on Information Technology Management, Tehran, Iran, http://www.civilica.com/Paper-ICTM06-ICTM06_004.html

Peer-Reviewed Conference Proceedings—Short Papers

19. Patrick, T. Zolnoori, M. (2016 October), *Depression Drug Side Effect Beyond Clinical Conditions* (October 2016),” American Public Health Association, Denver
20. Zolnoori, M. Patrick, T. Conway, M. Faiola A., Lue, J. (2016 November), *Evaluating Acceptability and Efficacy of Antidepressant Medications using Patients Comments*, American Medical Informatics Association, Chicago
21. Zolnoori, M. Nambisan, P., Patrick, T. (2015 November), *Seeking support on Facebook, Content Analysis of Depression Group*. Paper presented at American Medical Informatics Association, San Francisco.
22. Jones, J., Zolnoori, M., Binkheder, S., Schilling, K., lenox , M., Pondugala, L.R. (2014, November). *The Extent to which U.S. Hospitals Promote Their Patient Engagement Activities and Outcomes: Preliminary Results of Quantitative Content Analysis Research*. Paper presented at American Medical Informatics Association, Washington DC.
23. Zolnoori, M., Schilling, K., Jones, J. (2014, November). *Patient-Centered Decision Support for Pediatric Asthma Signs and Symptoms: Development of a Web-Based User Interface for Parents*. Paper presented at American Medical Informatics Association, Washington DC.
24. Jones, J., Boston-Clay C., Kilaru A., Zolnoori M., Malika M. (2013, November) *Building a Portal to Health Resources for Cancer Survivors*. Paper presented at American Medical Informatics Association, Washington DC.
25. Zolnoori M., Fazel Zarandi M.H., Moein M., Teimorian S. (2010 November) *Intelligent Systems in Medicine: Designing a Fuzzy Expert System for Evaluation Severity of Asthma Exacerbation" (in Persian)*. Paper presented at 17th International

TEACHING EXPERIENCE

Courses Taught

COURSE#	COURSE TITLE	WHEN	SIZE	LOCATION
HS 311	Ethics and Laws for Healthcare Professional	Spring 2015	38	UWM
HCA 450	Healthcare Quality Management	Fall 2015	40	UWM
HCA 590	Research Design and Methodology	Fall 2015	8	UWM
CHS 100	Seminar in the Health Professions	Fall 2017	15	UWM

Course Designed

HCA.450	Healthcare Quality Management	Fall 2015		UWM
HCA 590-SEC210	How to formulate a research question.	Fall 2015		UWM
HCA 590-SEC206	Accessing data and information for research	Fall 2015		UWM
HCA 590-SEC207	How to write a scientific paper?	Fall 2015		UWM

Teaching Certificate

Spring 2016 Received UWM certificate for excellence in online and blended teaching

Thesis Mentoring

Pondugala, LR. (2014) Evaluation of various hospital websites based on Patient Engagement Framework, Department of Health Informatics, Indiana University Purdue University Indianapolis (IUPUI), US.

Bayani, M. (2010- 2011) diagnosing pediatric asthma using type 2 fuzzy clustering, Department of Industrial Engineering, Amirkabir University Technology, Tehran, Iran

Darabi, SA. (2011-2012) Diagnosing adult asthma using case-based reasoning, Department of Industrial Engineering, Tarbiat Modares University, Tehran, Iran

Teaching Training

Summer 2015 Training Program for Online and Blended Teaching, University of Wisconsin-Milwaukee

Course Reviewer

Spring 2017 reviewer of the Online & Blended Teaching Certificates, UWM 2017

RESEARCH GRANTS AND PATENTS

- 2011 Research grant awarded by the Immunology, Asthma and Allergy Research Center (IAARI) for designing a patient-centered decision support system for screening pediatric asthma, Tehran, Iran
- 2012 Patent awarded by the Iran Intellectual Property Center for CDS system for screening asthmatic patients, Tehran, Iran

AWARDS

- The Chancellor's Graduate Student Award (CGSA)- UWM* , Spring 2015, Fall 2015, Fall 2017
- 2016 Travel Award, UWM (Fall 2016), American Medical Informatics Association (AMIA) Conference, Chicago, Illinois.
- 2015 The Chancellor's Graduate Student Award (CGSA)- UWM*
- 2015 Travel Award, American Medical Informatics Association (AMIA) Conference, San Francisco, California.
- 2014 Travel Award, HIMSS Conference & Exhibition, HIMSS Conference, Orlando, Florida.
- 2014 Travel Award, Indiana University, Fall (2014), American Medical Informatics Association (AMIA) Conference, Washington, D.C.
- 2013 Travel Award, Indiana University, Summer (2013), Human Computer Interaction Conference, Las Vegas, Nevada

**CGSA is designed to give UWM a competitive edge in attracting and retaining high quality, talented graduate students.*

SERVICES

Journal Editor

- Editor of American Medical Informatics Association (AMIA), 2016, 2014
- Editor of mental health informatics, American Public health Association (APHA), 2017
- Editor of Special issues of Journal of Pulmonary & Respiratory Medicine:
<http://www.omicsonline.org/specialissueJPRM.php>
- Reviewer of Information Technology books in ACECR (Academic Center for Education, Culture and Research) Student Book Festivals

5)

Volunteer Experience

- International Graduate Welcome Volunteer (IGWV) Program, Office of International Affairs, INDIANA UNIVERSITY, 2013-2014

- Indy Big Data conference, Indianapolis, February 2014. <http://www.indybigdata.com>
- Middle East cultural festival, Eli Lilly Company, August 2014
- Program Assistant, HIMSS Conference, Orlando, Florida, February 2014
- Project Coordinator, Promoting Programs and Activities of the Department of Health Informatics administration (HIA) at UWM using Social Media Features, Summer 2015
- AMIA 2016, Mentoring high school students

OTHER

Miscellaneous Research Projects

- **Database design:** *A Web-based System for Organizing Literature Review. This project was defined as the final project of database design and management course.*
- **Contextual Inquiry for Developing a Computerized Patient Education System.** *This project was defined for the course of Meaning and Use in HCI.*
- **Calculating Diagnosing Value of Risk Factors and Symptoms in Diagnosing Asthma.** *This project was defined for the course of statistical analysis.*
- **Designing a Feed Forward Neural Network for Prediction of Gold Price.** *This project was defined for the course of Neural Network.*
- **A Prototype Decision Support System for Recognizing Pesticide Type in Agriculture.** *This project was defined for the course of expert system designing and implementation.*
- **Predicting the Inflation rate in Iran using Time Serious Analysis.** *This project was defined for the course of statistical analysis.*
- **An Analysis of Accounting System of Management Department of Tehran University**
- **Comparison and Contrast Business Intelligence Style in Product-based Software Companies and Service- Based Software Companies.**
- **Strategic Planning for BaftLoran Company, Lorestan, Iran.** *This project was defined for the course of strategic planning.*

REFEREES

1) Dr. Timothy B. Patrick,

- Major Advisor
- Department of Health Sciences, University of Wisconsin-Milwaukee
- Phone: (414) 975-1998
- Email: tp5@uwm.edu

2) Dr. Kin Wah Fung

- Mentor at NLM/NIH and Dissertation committee member
- National Library of Medicine

- Phone: (301) 827-5001
- Email: kwfung@lhcnlm.nih.gov

6)

7)

3) Dr. Paul Fontelo

- Dissertation Committee member
- Director of training program at NLM/NIH
- Phone: (301) 827-4964
- Email: pfontelo@mail.nih.gov

4) Dr. Jake Luo,

- Dissertation Committee member
- Department of Health Sciences, University of Wisconsin-Milwaukee
- Phone: (414) 229-7333
- Email: jakeluo@uwm.edu

APPENDIX A

1. Lister Hill National Center for Biomedical Communications/ National Institute of Health (May 2016- December 2016), *Research Fellow*

1.1. *Project:* *Social Media text mining and statistical analysis to identify underlying factors for attitude to antidepressant treatment (in progress).*

- 1.1.1. Designing research methodology including literature review, data collection method, data analysis framework, annotation guideline, and statistical analysis plan.
- 1.1.2. Developing a consumer health vocabulary of medical concepts, such as side-effects and withdrawal symptoms by systematically mapping these concepts to both SNOMED-CT and UMLS concepts
- 1.1.3. Designing a classifier to differentiate symptoms, side-effects, and beneficial effects of antidepressants

8)

2. Biomedical Data and Language Processing Center of the Department of Health Sciences at the University of Wisconsin-Milwaukee, Research Fellow (January 2015, May-2016)

2.1. *Project (1):* *Representing Depression Drug Side Effect beyond clinical Conditions*

- 2.1.1. Mapped drug adverse effects to limitations in daily activities and social participations based on data extracted from drug clinical trials and International Classification of Functioning Disability and Health (ICF) terminology
- 2.1.2. Analyzed the result and prepared the manuscript for American public health association (APHA) conference.

2.2. *Project (2):* *Evaluating Acceptability and Efficacy of Antidepressant Medications based on Patients drug reviews*

- 2.2.1. Compared acceptability (satisfaction) and efficacy (effectiveness) of 13 antidepressant drugs using patients drug reviews in Webmd.com

- 2.2.2. Analyzed the results and prepared the manuscript for American Medical Informatics Association (AMIA)
- 2.3. **Other projects:**
 - 2.3.1. Analyzed the Patients Comments in Depression Groups in Facebook
 - 2.3.2. Health data in Social Media and Exploring Text-Mining Methods for Data Extraction
- 3. **Indiana University, Research Assistant (August 2012- December 2015)**
 - 3.1. **Project (1): *The Extent to which U.S. Hospitals Promote Their Patient Engagement Activities and Outcomes: Preliminary Results of Quantitative Content Analysis Research***
 - 3.1.1. Collaborated with investigators to design research methodology
 - 3.1.2. Conducted literature review to develop framework for content analysis
 - 3.1.3. Worked with PhD students to analyze contents and medical tools provided in hospitals' websites
 - 3.1.4. Conducted statistical analysis to test the hypotheses
 - 3.1.5. Prepared manuscript for American Medical Informatics Association (AMIA)
 - 3.2. **Project (2): *Building a Portal to Health Resources for Cancer Survivors***
 - 3.2.1. Collaborated with investigators to develop a framework for the portal
 - 3.2.2. Conducted a research to select proper content for the website
 - 3.2.3. **Project (3): *Improving Healthcare Systems for Access to Care by Underserved Patients,***
 - 3.2.4. Worked with investigators and team members for conducting patients interviews and analyzing the collected information
- 9)
- 4. **Immunology & Asthma & Allergy Research Center (IAARI), Tehran, Iran, Informatics Specialist, (January 2011- August 2012)**
 - 4.1. **Project (1): *A Clinical Decision Support System (CDSS) for diagnosing pediatric asthma***
 - 4.1.1. Prepared a grant proposal for IAARI to finance designing and implementing a clinical decision support system.
 - 4.1.2. Worked with asthma specialists to generate diagnostic rules
 - 4.1.3. Worked with medical students to provide medical content of the CDSS
 - 4.1.4. Designed the algorithm of diagnosis
 - 4.1.5. Implemented a prototype of the system using Matlab
 - 4.1.6. Worked with programmers and user-interface designers to implement the system using C#
 - 4.1.7. Tested the system using 138 asthmatic and 138 non-asthmatic patients
 - 4.2. **Project (2): *Clinical Decision Support System (CDSS) prototype for measuring level of asthma severity, asthma control, and asthma exacerbation***
 - 4.2.1. Worked with asthma specialists to generate rules related to level of severity, control, and exacerbation
 - 4.2.2. Implemented system prototype using Matlab
 - 4.2.3. Tested the prototype using 85 asthmatic patients
- 5. **Academic Center for Education, Culture, and Research (ACECR) Tehran, Iran, Research Specialist, (2009-2012)**

5.1. List of projects:

- 5.1.1. Studied ICT-based SMEs in Iran and providing the government with appropriate solutions in order to support the enterprises.
- 5.1.2. Strategic planning: a framework for Internal analysis in high tech SMEs
- 5.1.3. Designed an efficient engineering and administrative system for exporting technical services

6. Negah Nou company and Jahan Rayane Company, Tehran, Iran, IT Consultant (2009-2010)

- 6.1. Determining information systems requirements
- 6.2. Designing business plan based on company's clients' needs
- 6.3. Designing system analysis framework for system analysis team

7. E-Sabz Company, Tehran, Iran, Director of Analytics Department, (2007-2009)

- 7.1. *For the following projects, I managed a team of system analysts and programmers. I was also responsible for preparing marketing business plan.*
 - 7.1.1. Content Management System
 - 7.1.2. Customer Relationship Management (CRM) system
 - 7.1.3. Archive Management System
 - 7.1.4. Learning Management System