Predicting Medication Prescription Rankings with Medication Relation Network

Completed Research Paper

Yuanyuan Gao
University of Utah
1635 Campus Center Dr., BU C 2
Salt Lake City, UT 84112
yuanyuan.gao@eccles.utah.edu

Paul Hu
University of Utah
1655 Campus Center Dr.
Salt Lake City, UT 84112
paul.hu@eccles.utah.edu

Abstract

Medication prescription rankings and demands prediction could benefit both medication consumers and pharmaceutical companies from various aspects. Our study predicts the medication prescription rankings focusing on patients' medication switch and combination behavior, which is an innovative genre of medication knowledge that could be learned from unstructured patient generated contents. We first construct two supervised machine learning systems for medication references identification and medication relations classification from unstructured patient's reviews. We further map the medication switch and combination relations into directed and undirected networks respectively. An adjusted transition in and out (ATIO) system is proposed for medication prescription rankings prediction. The proposed system demonstrates the highest positive correlation with actual medication prescription amounts comparing to other network-based measures. In order to predict the prescription demand changes, we compare four predictive regression models. The model incorporated the network-based measure from ATIO system as one of the predictors achieve the lowest mean squared errors.

Keywords: Data mining, Prescription rankings prediction, Medication relation network, Text mining, Medication switch, Medication combination

Introduction

Medication prescription forecasting allows predicting future medication demands. It is one important process of medication planning by influencing future supply of medication and related products (Sekhri et al. 2006). Accurate medication prescription predictions benefit not only pharmaceutical companies but also general users for whom such information is critical but difficult to gain. To evaluate the competition and the most profitable products by obtaining medication prescription rankings and changes over time, pharmaceutical companies usually invest a huge amount of money to hire analysts manually acquiring such information from institutions or medication users. For medication users, the medication rankings could shed light on the effectiveness of alternative medications. With this information, patients could identify the most widely used medications. To predict the medication prescription, prior research mainly focuses on using historical prescription information (El-Iskandarani et al. 2013 and Ghousi et al. 2012) and patient information (Rose et al. 1985) without considering the influence of relations among medications. Our study, instead, forecasts prescription rankings and demands from a novel aspect by concentrating on two types of medication relations: medication switch and combination, and exploring their predictive utilities for forecasting prescription rankings and future demands.

Medication switch refers to the changing of one brand-name product to another, a brand-name product to a generic medication, or a generic medication to the same product produced by different manufacturers (Furberg et al. 2010). This is usually determined by both patients and their physicians. Based on user generated content, patients could require a medication switch due to severe side effects or inefficacy. Then physicians would change their prescriptions according to the feedback. The medication combination refers to two or more active pharmaceutical ingredients (APIs) combined in a single dosage form (Adams et al. 2015). Changes of medication prescription amounts could be partially result from the altering of patients' medication combining and switching behavior. If more patients switch from other brands to this focal one than patients switch away, this focal medication would receive more prescription demands and better prescription ranking. In addition, high consumption of one medication could serve as an important signal for the increase of prescription demands of its combination medications.

Nevertheless, complete real life medication transition and combination data is difficult to collect. The direct method to retrieve medication relations is to query patients for their medication prescription history from a general population, which is very expensive. On the other hand, with the number of health care related social networks increasing rapidly these years, people are more likely to read, communicate or report their health condition, disease or medication usage experience on different websites such as personal blogs, health related forums, health product reviews, personal microblogs or tweets (Ventola 2014). The user generated data, alternatively, which contain abundant detailed medication relation information, could be analyzed manually or automatically for medication relations extraction. As demonstrated in Figure 1, besides the focal medication experience, they also discussed additional medication involvements. Under the discussions from WebMD[1] of focal medication "Xanax", a few sentences of reviews could be extracted "The combo of Lexapro and Xanax has…" and "I tried valium, buspar and klonopin and …, after trying Xanax …". The two sentences represent two types of medication relations. The first sentence demonstrates a medication combination relation which indicates the medication "Xanax" and medication "Lexapro" are taking together by the same patient in a single dosage form. The second sentence displays a transition relation which demonstrates the sequence of medications taken by patients. In this case, patient was first on medication "Valium", "Buspar" and "Klonopin", but due to the ineffectiveness, he or she switched from these three medications to "Xanax". The switch from one medication to another and combination relations revealed in patients' review could reflect their honest medication switching and combination behavior in real life.

Comment 1: The constant anxiety was destroying my life. The combo of Lexapro and Xanax has made me feel better than I have in years. I've used it for four years and will continue using it.	Comment 2: This is exactly what I needed. After being discharged from the army I began having severe anxiety and frequent panic attacks. I tried valium, busapr and klonopin and none of them worked. After trying Xanax I finally felt normal.

Figure 1. Medication Transition and Combination Example

In order to predict prescription rankings and demands with medication relations, our study, therefore, extracts medication transition and combination relations from patients' generated contents that could reflect their actual common switching and combining actions in general. Two supervised machine learning systems are proposed for medication identification and relationship extraction from user generated content. The medication relations are further mapped into medication relation network (MRN). We then propose a novel adjusted transition in and out (ATIO) system with MRN-based measures for prescription rankings prediction. The machine learning systems and proposed transition system are built and evaluated with real-world medication reviews and prescription data.

Our study is different from the prior social network based medication learning studies from several aspects. First, comparing to the Adverse Drug Reactions (ADRs) extraction and detection studies (Leaman et al. 2010 and Liu et. al 2014) that focus on the adverse effects of one focal medication, our study not only emphasize on the focal medication that is mainly discussed, but also highlight other medications present together with the focal one. Second, in contrast to medication comparison study (Adusumalli et al. 2015) that compares the effectiveness of medications treating similar symptom using pre-existing domain

[1] WebMD is an online publisher of news and information pertaining to human health and well-being. (http://www.webmd.com)

knowledge and user generated ratings, our study analyzes entire medication transition and combination relations based on the review and discussion contents. Third, even though the Drug-Drug Interactions (DDIs) detection studies (Van Puijenbroek et al. 1999 and Caster et al. 2010) analyze more than one medications at once, the purpose of these studies is to discover the interactive effects of the medications so as to complement current pharmacovigilance system. Our study instead retrieves medications occurring in the same discussion content, and further analyzes if they are switching from or taking together with each other. By discovering the transition and combination patterns, the main aim of our study is to forecast the medication prescription rankings and demand changes.

The remainder of the paper is organized as follows. We first review related studies and highlight key differences between our study and representative previous research. We then formally propose two supervised machine learning systems for medication references (all medication names or references mentioned in the reviews) recognition and relationship extraction. Next, we propose ATIO system and use the system for medication prescription rankings prediction. Furthermore, we use medication reviews to build and evaluate two machine learning systems and the proposed ranking system, with several prevalent methods as benchmarks. Medication prescription demand changes are predicted as well with four regression models. We conclude the paper by discussing important contributions, implications, and limitations.

Literature Review

Several streams of research are relevant to our study, including user generated contents mining for medication related knowledge learning, product network extraction from user generated contents and medication demands prediction. In this section, we review representative studies in each stream and highlight the gaps that motivate our research.

Medication Related Knowledge Learning from User Generated Contents

Prior studies take the benefit of large volume user generated data on social media and make it a useful resource for ADRs detection (Leaman et al. 2010 and Liu et. al 2014), DDIs identification (Van Puijenbroek et al. 1999 and Caster et al. 2010) and medication comparison (Adusumalli et al. 2015).

ADRs refer to injuries caused by taking medication (Edwards and Aronson 2000). They are caused by single medication dose or the combination of two or more medications. Comparing to other areas, due to the severity of possible ADRs, ADRs detection and analysis has drawn great attention from both academia and industry. Social media that both health-related such as, MedHelp (Liu et. al 2014 and Yang et al. 2014), DailyStrength (Leaman et al. 2010 and Yeleswarapu et al. 2014), and general social media such as Twitter (Ginn et al. 2014 and Bian et al. 2012) are used for ADRs learning. In terms of extraction method, both supervised and un-supervised machine learning systems are adopted for concept extraction, drug classification and ADRs classification. Lexicon-based, pattern-based, association rule and parsing tree are the most commonly employed un-supervised learning methods. Leaman et al. (2010) and Nikfarjam and Gonzalez (2011) both employ lexicon-based system development and pattern-matching to identify patterns. Association rule mining is utilized to identify ADR pairs by Yang et al. (2012). Yates and Goharian (2013) first develop about 7 patterns of ADR relationship and adopt these patterns to extract ADRs from estimated 125 manually annotated comments. Parsing tree is used to extract dependency relations between words. Liu and Chen (2013) apply parsing tree and discover the shortest dependency path between drug and its adverse reaction, therefore identify the drug-reaction pairs relation. Supervised learning methods on the other hand, are built depending on data annotations. Bian et al. (2012) construct two Support Vector Machine (SVM) classifiers. The two classifiers are used for classifying if patients use the medication and if the review post contains an ADR respectively. Besides SVM, maximum entropy classifier is trained on annotated 600 tweets and another 285 tweets are used for testing by Jiang and Zheng (2013). Yang et al. (2013) combine supervised and unsupervised approaches. For supervised learning, they train both SVM and naïve Bayes classification models with sentiment features as predictors.

DDIs refer to the activity of a drug is affected by a substance (another drug) when both of them are administered together (National Prescribing Service 2009[2]). Since DDIs and ADRs detection are both important for drug safety surveillance, DDIs detection from unstructured user generated contents has been widely studied as well. Supervised and un-supervised high dimensional models such as multivariate logistic regression, the shrinkage measure, multiplicative model and association rule learning are employed for DDIs identification. Van Puijenbroek et al. (1999) propose multivariate logistic regression for retrospective detection of DDIs with FDA Substance Registration System (SRS) database. They discover the potential association between delay of withdrawal bleeding during concomitant use of oral contraceptives (OCs) and antifungal Iitraconazole (I). Similarly, Bayesian logistic regression is utilized by Caster et al. (2010) to address the "masking effect", which is the increase in background could not attenuate values of true association. Norén et al. (2008) propose the Shrinkage measure for recognizing drug interactions and improving the performance comparing to logistic regression, which sometimes miss important DDIs. Thakrar et al. (2007) compare the model performances of multiplicative model and additive model using existing known drug interactions. In terms of detecting performance, additive model has better sensitivity than multiplicative model. Unsupervised learning technique such as association rule learning has been used to draw inferences of hidden relations as well. Harpaz et al. (2010) employ Apriori algorithm based association rule to mine a sample of FDA Adverse Event Reporting System (AERS). At least two medications and one adverse effect as a rule are reported as a detection result.

Due to limited contribution to pharmacovigilance, few studies have extracted multiple medications and their relations from unstructured user generated contents. Adusumalli et al. (2015) compare the effectiveness of medications treating similar symptom by numeric user generated rating from WebMD. The categorization of medication comparison pairs is depending on medication targets learned from the prior domain knowledge. The comparison of effectiveness then further be evaluated by counted results from prior literature. Even though a small portion of pairs (77 out of the 427) are discovered in prior studies, approximately two thirds of the trends are supported.

Comparative, Competitive Opinions and Cooperative Intelligence Learning from User Generated Contents

Limited studies have concentrated on multiple medications comparison on the basis of unstructured text contents. In terms of business realm, although majority of studies focus on investigating opinions towards a single entity (e.g. a product), there are a few studies available aiming at discovering competitive and comparative opinions, as well as cooperative intelligence among various products mentioned together. We elaborate the representative studies in detail.

On the cooperative intelligence front, Dhar et al. (2014) define an economic network based on customer's co-purchasing behaviour. The aggregated network measures such as PageRank and Cluster are employed as features for prediction models, which are used for future focal product demands prediction.

In order to learn product competitive and comparative opinions, Netzer et al. (2012) apply text mining methods to build graphs based on co-occurrence of product name entities from sedan and diabetic drug forum messages. They link sedan product name entities with undirected link and construct undirected graph. The graph is further employed for sedan market structure and competition analysis. Based on the co-occurrence of products from same sentence, Xu et al. (2011) propose a sequence labelling conditional random field (CRF)-based model to extract comparative relations between products from each sentence of online customer reviews. Zhang et al. (2013) group product pairs with co-occurrence at message level and propose a polarity classifier giving sentiment scores for each product pair. The higher sentiment score for one product is comparatively better than the other one from the same product pair. They then construct a directed product network based on the comparative relations. PageRank and other network-based measures are adopted for product rankings prediction and market structure mapping.

[2] National Prescribing Service 2009 is available at http://www.nps.org.au/media-centre/media-releases.

Pharmaceutical Medication Demands Prediction

To predict medication prescription changes in the future, understanding the factors that influence the medication demands is essential. Similar to product demands, the demands for pharmaceutical medication are affected by numerous aspects. Manufacture's marketing strategy, such as Direct-to-Consumer advertising (Donohue et al. 2007), is associated with significant growth in sales for different classes of medication. Other factors such as physician and patient incentives (Dickstein 2011), seasonal and epidemic disease, market share of competitive products and active ingredient ratio (Galarraga et al. 2007) are considered as main predictors for forecasting medication demands.

Forecasting revenue potential and driving a number of tactical and strategic decisions are valuable for pharmaceutical companies, but it is extremely difficult to make accurate predictions. In order to forecast the demands of medication, the most straightforward approach is to collect report data from World Health Organization (WHO). According to forecasts of global and regional demand report, the future demands of medications are determined on the basis of important variables such as total number of people receiving the therapy and the number of person-years of treatment. Then they utilize three forecasting models linear regression, country target model and Clinton health access initiative to predict demands of therapy and antiretroviral medicines in year 2012-2015. On the other hand, historical demand data could be used for forecasting the future medication consumption. El-Iskandarani et al. (2013) propose a temporal pattern matching algorithm. By discovering the most similar consumption patterns of medication in the past, they could predict the patterns in the future. Ghousi et al. (2012) compare different machine learning predictive algorithms such as regression, artificial neural network (ANN) and decision tree with historical report of medication distribution and customers' demography information. However, historical data could offer few guidelines for forecasting the future of a new medication or a medication in a new therapeutic category. Instead of depending on historical prescription data, Rose et al. (1985) forecast the future medication prescription demands by estimating the number of future patients by incorporating potential new patient population and original continuing patients.

Research Gap Summary

As we discussed in the introduction, discovering the pattern of medication transition and combination is important for competition recognition and demands forecasting. Comparing to the current studies which mainly focus on extracting ADRs and DDIs from social networks that contribute to pharmacovigilance, we explore a new realm of study that pay attention to obtain multiple medications, categorize their relations and map the relations to a network for prescription prediction. In contrast to medication effectiveness comparison learning (Adusumalli et al. 2015), we do not pre-group certain medications together as pairs using domain knowledge and only compare pairs. On the other hand, besides numeric general ratings, we extract and learn unstructured review text with detailed medication combination and transition patterns. The key differences between our study and prior social network medication learning studies are summarized in Table 1.

Study	**Focus Level**	**Analysis Level**	**Purpose**	**Implication**
ADRs detection	Focal mediation only	User generated message	ADRs extraction and analysis	Pharmacovigilance
DDIs detection	Multiple medications	User generated message	DDIs extraction and analysis	Pharmacovigilance
Effectiveness comparison	Multiple medications	Multiple user generated messages and ratings	Compare the effectiveness of medication pairs with same target	Medication selection
Current study	Multiple medications	User generated message	Medication relations extraction and categorization	Competition recognition and demand prediction

Table 1. A Brief Summary of Medication Learning Related Studies

Furthermore, prior product network studies emphasize on extracting main comparative, competitive or cooperative relations among product entities. They utilize customer co-purchasing records (Dhar et al 2014) for mining cooperative relations and co-occurrence with sentiment analysis for identifying comparative entities (Netzer et al. 2012, Xu et al. 2011 and Zhang et al. 2013). Our study, not only detects medication co-purchasing relation, which is considered as medication combination, but also further discovers the latent medication switch sequences and transition patterns. Medication switch is not exactly equivalent to medication comparison. Patients do not necessarily compare switch from medication with focal medication or switch to medication in their posts. When users make comparison between two products, these two products might not be consumed sequentially and users might use both of them at the same time. In addition, prior studies either map a directed or undirected entity network for product relation analysis. Our study constructs both undirected and directed networks based on medication combination and transition relations respectively. The major differences between our study and prior product comparison studies are summarized in Table 2.

Study	Focal Problem	Analysis Level	Network Perspective
Dhar et al 2014	Entities co-purchasing	Purchasing records	Directed links and nodes
Netzer et al. 2012	Entities comparison	Each message from posts	Undirected links and nodes
Xu et al. 2011	Entities comparison	Each sentence from messages	Directed links
Zhang et al. 2013	Entities comparison	Each message from posts	Directed links and nodes
Current Study	Entities transition and combination	Each message from posts	Undirected links, directed links and nodes

Table 2. A Brief Summary of Product Comparison Related Studies

Last but not the least, majority of the medication demand forecasting studies rely on historical demand and patient data for certain medication and employ machine learning algorithms such as regression, ANN and decision tree for future demand prediction. Besides historical demands, factors such as manufacture's marketing strategy (Donohue et al. 2007), physician and patient incentives (Dickstein 2011), seasonal and epidemic disease, market share of competitive products and active ingredient ratio (Galarraga et al. 2007) that essentially influence the demand are difficult to access in real world. Our study, without fully depending on detailed historical prescription and patient data, obtain predictors from unstructured user generated discussion. Through learning medication transition and combination relations, the market structures for competitive and cooperative products could be estimated and utilized for prescription demands and rankings prediction.

Research Design

In this section, we first propose two supervised SVM systems for medication reference identification and medication relation classification. The medication relations are further transformed to the medication relation network (MRN). A novel ranking system ATIO using MRN-based measures is proposed for medication prescription rankings prediction.

Medication Reference Identification and Relation Classification

To explore the medication transition and combination patterns, we identify medication references from unstructured patients' review, and then categorize medication references into different classes. First of all, the phrases that refer to the medication that is currently being reviewed and mainly discussed are categorized into CurrentMed (CM) class. For the review sentence "I really do not like to take Prednisone unless I absolutely have to", "Prednisone" is regarded as CM. There are different reasons why users refer to other medications in their generated reviews. Users could mention the medication in their generated

reviews and stated they previously took it before starting on the CM. We then categorize this type of medication references into SwitchFromMed (SFM) class. For the review sentence "I didn't walk for 6 months, even with Methotrexate, I finally found a smart Dr. who prescribed me Prednisone", "Methotrexate" is considered as SFM of the CM "prednisone". Furthermore, users might also mention the medications they plan to take in the future or already started to take after they stopping on the CM. This type of medication references is categorized into SwitchToMed (STM) class. For example, for review sentence "Only took for one week, then back to Ibuprophen", "Ibuprophen" is the STM of the CM. In addition, some medications are mentioned by users in their reviews together with the CM to treat the side effect caused by CM, relieve other symptoms or use as a supplement to enhance the treatment power of CM. This type of medications belongs to TakeCombineMed (TCM) class. Take a review sentence "I also take Asacol and Remicade" as an example, "Asacol" and "Remicade" are both TCMs for the CM. If the medication references that do not fall into any categories above, we categorize this type of medications as Other Medication (OM). The medication transition pattern, therefore, could be recognized as starting from SFM to CM, then ending at STM. The medications belong to CM and TCM classes are medication combination.

In order to identify different medication references and classify the relation types for reviewed medications, we applied two stage system on the extracted reviews. Identification of medication references is the first stage and classification of the medication references into the above mentioned five classes is the second stage.

Extracting all the medication references from the raw text is challenging. Two approaches are employed to identify the medication references. One basic approach is to take advantage of the available tools such as MedEX (Xu et al. 2010) to recognize the medication names from the raw reviews. Besides the MedEx, we further develop our medication identification system. Our system treats each word in the reviews as medication reference candidate. Then we employ SVM binary classifier to identify if these words are medication references or not. The features used for classification model building include the word feature, Part-Of-Speech (POS) tags feature, the prefix. The full feature list is presented in Table 3.

Feature Name	**Description**
Word Type	Current word
Part-Of-Speech	The POS of the current word
Morphology	The suffix and prefix of the current word
Lemma	The lemma of current word
MedEX Tag	The semantic tag assigned by the MedEX

Table 3. Features for Medication Reference Identification Model

A medication relation classification system is employed to recognize the relations between medication references by classifying them into five classes: STM, SFM, TCM, CM and OM. After identifying all the medication references in the reviews, these references then be categorized with relation classification system. The relation classification system consists of five binary classification models with each binary classifier identifying one medication relation class. These classifiers are trained on the manually annotated data with medication class labels. For each binary classifier, we regard all the medication references that belong to the certain medication relation class as positive training examples, and the rest of references are considered as negative training examples. We apply linear SVM classifier as well to construct the five classifiers. The features for the five classifiers are presented in Table 4.

Feature Name	**Description**
BOW	Bag of words as features; window size is 5.
POS	POS of surrounding words, window size is 2
Bigram	Bigram features, window size is 3

Dependency	The dependency relations with current word

Table 4. Features for Medication Reference Classification Model

MRN

In order to rank medication prescription demands with the extracted medication relations, we define two types of network: directed and undirected to map the relations among medications. The three major relations of medications: switching from the focal medication to another, switching from other medication to the focal medication, and taking other medication combine with the focal medication could be illustrated in directed and undirected networks.

Switch from One Medication to Another Medication

We demonstrate the "switch from x to y" relation of medications in a directed graph **M**, with each node representing one distinct medication and one directed edge corresponding to the switching relation between two medications. For example, we extract the relation between medications "Risperdal" and "Geodon" from one patient's review indicating the patient was first prescribed "Risperdal" and then changed to "Geodon". We show this switching relation in a directed graph with "Risperdal" and "Geodon" as two nodes and a directed edge from "Risperdal" pointing to "Geodon". On the other hand, if patient mentioned he/she is taking "Xanax" after stopping on "Pristiq", the directed edge would start from "Pristiq" pointing to "Xanax".

Figure 2. Switch Medications from One to Another

Combine One Medication with Another Medication

We also demonstrate the "taking x together with y" relation of medications in an undirected graph **N**, with each node representing one distinct medication and one undirected edge corresponding to the combination relation between two medications. For example, we extract the relation between medications "Cymbalta" and "Geodon" from one patient's review indicating the patient was taking prescriptions "Cymbalta" and "Geodon" at the same time. We display this combination relation in an undirected graph with "Cymbalta" and "Geodon" as two nodes and an undirected edge connecting these nodes shown in Figure 3.

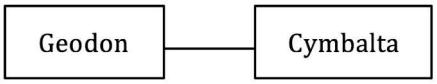

Figure 3. Combined Medications

Convert Undirected Network to Directed Network

We then convert all undirected edges from **N** to directed edges in order to combine all medications in one directed network **S**. We transform undirected edges to directed edges by replacing undirected edges with bidirectional links (Brams et al. 2006). For example, we substitute the undirected edge between nodes "Geodon" and "Cymbalta" by two directed edges both linking from "Geodon" pointing to "Cymbalta" and linking from "Cymbalta" pointing to "Geodon" shown in Figure 4.

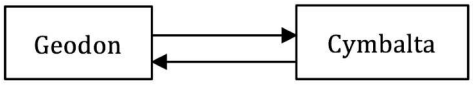

Figure 4. Convert Undirected Graph to Directed Graph

Rank Medication Prescription Demands

To rank medication prescription demands based on medication transition and combination network, we first transform the network to a transition frequency matrix, a transition probability matrix and a Markov transition matrix. We then further propose an ATIO system. The prescription demand rankings are based on the difference of transition in and out for each medication.

Construct a Matrix with MRN

We draw the medications network **S** by combining network **M** and **N** using the medication relations extracted from the reviews. The network then are transformed into a transition frequency matrix $\boldsymbol{A}$ with each element in the matrix defined as A_{ij}. A_{ij} is the frequency of directed edge from node j to i $(i, j \in U)$. U is the collection of all extracted medication references. For example, two patients mentioned switching medication "Xanax" to medication "Pristiq" in their reviews and one patient mentioned taking "Pristiq" after stopping on "Xanax". Therefore, the elements corresponding to the relation between medication "Xanax" and medication "Pristiq" in the matrix $\boldsymbol{A}$ are: $A_{(xanax)(pristiq)} = 1$ and $A_{(pristiq)(xanax)} = 2$. Based on the frequency matrix $\boldsymbol{A}$, we then compute the transition probability for each pair of medications using Equation (1) (Shin et al. 2014).

$$T_{ij} = \frac{A_{ij}+\epsilon}{A_{ij}+A_{ji}+2\epsilon} \qquad (1)$$

A_{ij} is the frequency of transitions from medication j to medication i, while A_{ji} is the frequency of transitions from medication i to medication j. Therefore, T_{ij} is the probability of transition from medication j to medication i among the total transitions between the medication pair j and i. We set $\epsilon > 0$ to ensure that $T_{ij} = \frac{1}{2}$ even when $A_{ij} = A_{ji} = 0$.

Image medications "Xanax", "Zoloft" and "Wellbutrin" having the transition relations shown in Figure 5.

Figure 5. Medication Transition Example

When we use the above equation to compute transition probability and assume $\epsilon = 1$, we would gain $T_{(Xanax)(Zoloft)} = \frac{1}{2}$, $T_{(Zoloft)(Xanax)} = \frac{1}{2}$, $T_{(Zoloft)(Wellbutrin)} = \frac{1}{2}$ and $T_{(Wellbutrin)(Zoloft)} = \frac{1}{2}$. We observe that "Zoloft" could switch to two available medications in Figure 5, while there is only one available medication "Xanax" and "Wellbutrin" could switch to. However, the transition probability for the three medications turn to be equal based on our transition probability (1). In order to distinguish the weight of directed edges for medications like "Xanax" and "Zoloft", we introduced k_j to adjust the weight. k_j is the number of out-degree edges of node j. Out-degree edges refer to the edges that transit away from the focal node. In addition, T_{ij} is the probability of node j pointing to node i. Parameter γ_i allows the sum of total transition probability from j to all other medications to be 1 . σ_{ij} is the element in adjacency matrix of the network $\boldsymbol{A}$ with $\sigma_{ij} = 1$ if and only if the edge $(i, j) \in U$. Hence, the Markov transition probability matrix $\boldsymbol{P}$ could be presented in equation (2):

$$P_{ij} = \gamma_i \cdot T_{ij} \cdot \sigma_{ij} \times k_j^{-1} \qquad (2)$$

By multiplying T_{ij} by k_j^{-1}, and assuming $\epsilon = 1$, we will adjust the value of transition probability by the number of out-degree edges of node j. Then we will gain $T_{(Xanax)(Zoloft)} \times k_{zoloft}^{-1} = \frac{1}{4}$, $T_{(Zoloft)(Xanax)} \times k_{Xanax}^{-1} = \frac{1}{2}$, $T_{(Zoloft)(Wellbutrin)} \times k_{Wellbutrin}^{-1} = \frac{1}{2}$ and $T_{(Wellbutrin)(Zoloft)} \times k_{zoloft}^{-1} = \frac{1}{4}$. The transition

probabilities from "Zoloft" to other nodes would be reduced comparing to the probabilities without adjusted parameter k_j^{-1}.

ATIO System

Park et al. (2005) propose a win-lose system to rank the U.S. football teams with a contest network they constructed. We adapt their system and construct an adjusted transition in and out system to calculate the medication transitions.

For a certain medication i, direct medication prescription transition probability towards medication i from medication j is P_{ij}. If there are in total n different medications switching to medication i, the total direct transition probability towards medication i could be calculated using TI_{ij}.

$$TI_{ij} = \sum_j^n P_{ij} \quad (3)$$

For a certain medication i, indirect medication prescription transition towards medication i from medication k could be demonstrated in Figure 6. Medication k indirectly convert to medication i by directly switching to medication j. The distance from medication k to medication i is 2, which is the number of edges from node k to node i.

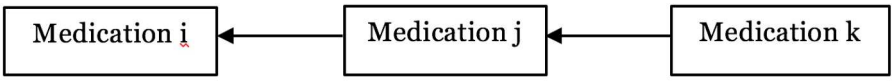

Figure 6. Indirect Transition Example

We define the indirect transition probability from node k to node i using W_{ik}.

$$W_{ik} = \alpha_j \times (P_{ij} \times P_{jk}) \quad (4)$$

α_j is the single free parameter that weighted on indirect prescription transition. If there are in total m different medications switching to medication i through medication j, the total indirect transition probability of medications that towards medication i with distance 2 could be calculated using TW_{ik}.

$$TW_{ik} = \alpha_j \times \sum_j^n \sum_k^m (P_{ij} \times P_{jk}) \quad (5)$$

For further remote indirect medication transition probability with longer distance towards medication i, we could demonstrate them using equation (6) to (7).

$$TW_{ih} = \alpha_j \times \alpha_k \times \sum_{jkh} (P_{ij} \times P_{jk} \times P_{kh}) \quad (6)$$

...

$$TW_{iy} = \prod_j^y \alpha_j \times \sum_{jkh\ldots xy} (P_{ij} \times P_{jk} \times P_{kh} \times \ldots \times P_{xy}) \quad (7)$$

Therefore, total probability of medication prescription transition towards medication i is shown as W_i. It is the sum of all probabilities transition towards medication i.

$$W_i = TI_{ij} + TW_{ik} + TW_{ih} + \cdots + TW_{iy} = \sum_j^n P_{ij} + \alpha_j \times \sum_j^n \sum_k^m (P_{ij} \times P_{jk}) + \alpha_j \times \alpha_k \times \sum_{jkh} (P_{ij} \times P_{jk} \times P_{kh}) + \prod_j^y \alpha_j \times \sum_{jkh\ldots xy} (P_{ij} \times P_{jk} \times P_{kh} \times \ldots \times P_{xy}) = \sum_j^n (1 + \alpha_j \times W_j) \times P_{ij} \quad (8)$$

Similarly, Total probability of medication prescription transition away from medication i is shown as L_i. β is the single free parameter that weighted on indirect prescription transition similar to α.

$$L_i = \sum_j^n P_{ji} + \beta_j \times \sum_j^n \sum_k^m (P_{ji} \times P_{kj}) + \beta_j \times \beta_k \times \sum_{jkh} (P_{ji} \times P_{kj} \times P_{hk}) + \prod_j^y \beta_j \times \sum_{jkh\ldots xy} (P_{ji} \times P_{kj} \times P_{hk} \times \ldots \times P_{yx}) = \sum_j^n (1 + \beta_j \times L_j) \times P_{ji} \quad (9)$$

The transition probability towards medication i is calculated as W_i and the transition probability away from medication i is presented as L_i. Therefore, the increasing number of prescription demands for one medication should be the difference between the number of prescriptions transiting into and the number of prescriptions transiting away from medication i. The growing demand of medication i could be estimated as G_i in equation (10).

$$G_i = \sum_j^n P_{ij} \times F_{ij} + \sum_j^n \sum_k^m \alpha_j \times (P_{ij} \times P_{jk}) \times F_{jk} + \sum_{jkh} \alpha_j \times \alpha_k \times (P_{ij} \times P_{jk} \times P_{kh}) \times F_{kh} + \cdots + \sum_{jkh\ldots xy} \prod_j^y \alpha_j \times (P_{ij} \times P_{jk} \times P_{kh} \times \ldots \times P_{xy}) \times F_{xy} - L_i \times K_i \quad (10)$$

K_i is the total number of direct transitions away from medication i in the relation network. F_{ij} is the total number of direct transitions between medication i and medication j. Then the new estimated demand of prescription N_i' for medication i would be represented as the sum of the total number of user generated reviews N_i for medication i and G_i shown in equation (11)

$$N_i' = N_i + G_i \quad (11)$$

The estimation of medication prescription demand rankings would be based on the amount of N_i' for each medication i ($i \in U$) . We further employ PageRank (Page et al. 1999), weighted PageRank (Xing et al. 2004) and other measures for medication prescription rankings estimation.

Obtain the Parameters α and β

Parameters α and β are free parameters that weighted on indirect prescription transition. More specifically, α is the weight on indirect prescription transition towards medication and β is the weight on indirected prescription transition away from medication. Park et al. (2005) compute the indirect transition weight using the mean and mean square of total number of all edges in the network. We on the other hand, intend to distinguish the weight of different nodes in the network by their centralities. Our method assumes medications transit to a medication with higher centrality would have larger possibility transiting to more new medications. Adjusted eigenvector centrality is applied and adapted to compute the parameters α and β, or approximate importance, of each node in the network. Specifically, we allow $\alpha_i = A_{ij} \times \alpha_j$ and $\beta_i = A_{ji} \times \beta_j$. α_i is the weight on indirect prescription transition towards medication i and β_i is the weight on indirect prescription transition away from medication i.

Empirical Study

In this section, we elaborate how we acquired and annotated patient generated reviews for medication reference identification and medication relation classification models building and evaluation. Reviews of medication from one specific category were extracted additionally and utilized for MRN mapping and ATIO system learning. The rating prediction performance was evaluated with real medication prescription data. In the end, four predictive regression models were constructed for prescription demands prediction.

Data Acquisition

In order to develop and evaluate supervised SVM models for medication reference identification and relation classification, labelled data is necessary for model building. We therefore firstly collected medication review webpages and annotated them. We applied a crawler (Httrack[3]) to crawl the medication review webpages from WebMD website. We then utilized Stanford coreNLP tool[4] to conduct the tokenization and sentence splitting for each online medication review. Brat[5] was employed to label the text annotations. Applying the brat for structured annotation, non-structured free form text would be transformed into a fixed form. Then the fixed form could be processed and interpreted by a computer program.

Totally we crawled and received 75,312 medication user generated reviews of more than 50 medications from year 2007 to year 2015. Reviews have been all extracted from raw html. We contained five different fields including the condition the medication was utilized to treat, the date when the review was posted,

[3] Httrack is an offline browser utility. It allows users to download a World Wide Web site from the Internet to a local directory, building recursively all directories, getting HTML, images, and other files from the server.(https://www.httrack.com)

[4] Stanford CoreNLP provides a set of natural language analysis tools. They could analyze the structure of the sentences, give the base forms of words, part of speech, entity recognition and so on. (http://stanfordnlp.github.io/CoreNLP/)

[5] Brat is a web-based tool for adding notes to existing text documents. (http://brat.nlplab.org)

patient's information such as age, gender, treatment time (how long they have been on this medication), the ratings according to the effectiveness, ease of usage and satisfaction and the review contents of each medication review the patient wrote for the medication. Amongst these 75,312 reviews, we selected 200 reviews from various medications and trained two of our team members to manually annotate the reviews. There are five different types of medication relations we tagged that discussed above and mainly reflected in the reviews. The medication that mainly discussed by patients in the current review is tagged as CM. The relation types of other medication references are tagged as SFM, STM, TCM and OM. The rest reviews were used for example demonstration of medication relations after the models were built. Due to the page limit, we did not include the demonstration section in this study.

To build the prescription ranking system, we need to construct a transition and combination network with a category of medications that frequently transit to or combine with each other. Therefore, we selected 27 psychiatric medications that targeting at 5 main mental illness categories as a demonstration. In total we crawled additional 36,245 reviews for these 27 psychiatric medications. To evaluate the ranking system, actual prescription demand rankings and prescription quantities for psychiatric medications were gathered from IMS Health[6]. We obtained psychiatric medications prescription rankings and quantities from year 2009, year 2011 and year 2013.

Evaluation of Medication Recognition and Relation Classification

We evaluated our systems on the manually annotated dataset. To fully utilize our human annotated reviews, we employed 10-fold cross validation (Geisser 1993) for both medication reference recognition and relation classification tasks. To assess the performance, we employed precision, recall and F-measure (Van Rijsbergen 1979) as our evaluation metrics. Our medication identification model employed the word-type, POS, morphology, lemma, MedEX tag as features. The results are presented in Table 5. We also employed and reported the MedEx as a benchmark technique on our dataset for identification task. MedEx is the rule based medication extraction system. Comparing to this rule based model, our system outperformed the benchmark method in terms of precision, recall and F-measure. Comparing the extraction results fetched from the two systems, we discovered the main issue of MedEX is that they failed to discover the pronouns, if the pronouns refer to medications. In the meantime, our system did fairly well on the prediction of pronouns.

Method	Precision	Recall	F-Measure
MedEX	65.7	24.0	65.7
OurSystem	84.3	73.6	78.6

Table 5. Medication Reference Recognition Results

The medication relation classification results are presented in Table 6. Since there are no other systems that have explored this task. We do not have any available benchmarks to compare with but ourselves. Therefore, we evaluated our system with various feature collections. The first collection included basic features (BOW features) and then the rich features included bigram features, POS tagging and dependency. Comparing the results from two collections of feature, rich features did not improve the system performance much. One possible reason is data sparsity problem. The number of instances in the training sample is limited. Hence, more features might lead to more severe sparsity problem.

	CM	TCM	SFM	Average
Basic Features	77/81/78	51/28/36	13/4/6	73/64/68
Rich Features	78/81/79	52/31/39	20/4/6	73/64/68

[6] IMS Health is a company that provides information, services and technology for the healthcare industry. It is the largest vendor of U.S. physician prescribing data. (http://www.imshealth.com)

Table 6. Medication Classification Results, Results are shown as Precision/Recall/F measure. CM, TCM and SFM refers to current medication, take combine medication and switch from medication[7]

Evaluation of Medication Prescription Rankings Prediction

In order to extract the MRN for prescription rankings prediction, the prior trained medication reference identification model was utilized to recognize psychiatric medication references from each sentence of 36,245 reviews. The trained medication relation classification model was then applied to classify the relations among the 27 medications.

The medication reviews written ranging from year 2007 to year 2008 were crawled and the relations among medications were mapped to predict the ranks of prescriptions in year 2009. Similarly, reviews written ranging from year 2007 to year 2010 were employed to predict the ranks in 2011 and reviews from year 2007 to year 2012 were engaged to predict the prescription ranks in 2013.

We further constructed three MRNs on the basis of psychiatric medication relations extracted from reviews before year 2009, 2011 and 2013. Three frequency matrixes and three Markov transition probabilities matrixes for different years according to the networks were then developed. We further computed and ranked the transition differences of each medication employing ATIO system for year 2009, 2011 and 2013. In addition, network-based measures such as PageRank, weighted PageRank, the authority score (instead of hub score, due to the transition nature of the medication) from Hyperlink-Induced Topic Search (HITS) (Kleinberg 1999) were constructed for rankings prediction. We then transformed our directed MRN to undirected network by treating directed edges as undirected edges. Betweenness centrality (Freeman 1977) was computed for the undirected network. Higher centrality measurements indicate higher importance of the position of the medication. The volume of the medication discussions, known as the number of reviews for each analyzed medication, could serve as an important measurement for rankings comparison. Larger volumes of medication reviews demonstrate the general larger number of discussions associated with the medication, which could also represent the prescription amounts. The above mentioned measurements including PageRank, weighted PageRank, the authority score, Betweenness centrality and number of reviews are served as benchmark methods in our study.

We therefore, compared the predicted rankings performance with true rankings of the psychiatric medications in year 2009, 2011 and 2013 with Kendall's rank correlation tau (Kendall 1938) and Spearman's rank correlation rho (Spearman 1904) in Table 7.

Method	Direction Setting	Kendall's rank correlation tau (Year 2009/2011/2013)	Spearman's rank correlation rho (Year 2009/2011/2013)
ATIO	Directed Graph	-0.2267/-0.3406**/-0.336**	-0.3615*/-0.5304**/-0.4911**
PageRank	Directed Graph	-0.0867/-0.3478**/-0.3281**	-0.1946/-0.4817**/-0.4486 **
Weighted PageRank	Directed Graph	-0.14/-0.3406**/-0.3123**	-0.2408/-0.5374**/-0.4773**
HITS Authority	Directed Graph	-0.1177/-0.362**/-0.3447**	-0.2198/-0.4972**/-0.4658**
Betweenness Centrality	Undirected Graph	-0.1/-0.2899**/-0.2727**	-0.1762/-0.3965*/-0.3844*
NumReview (Volume)	No Graph	-0.2133/-0.3768***/-0.4071***	-0.3338/-0.533***/-0.5326***

[7] Due to the limited number of STM, we only presented the evaluation results for the three medication classes.

Table 7. Univariate Rank Correlation with Prescription Rank in 2009, 2011 and 2013 (* p<0.1, ** p<0.05 and *** p<0.01)

Due to the ranking nature of the prescriptions, a smaller rank value implies larger prescription amount. If the correlation measure is negative, the indicated predictor has a positive influence on the actual prescription quantities.

As we can observe from the correlation results, measurements calculated by PageRank, weighted PageRank and HITS authority demonstrate significant negative correlation with Kendall's τ from -0.3123 to -0.362 and Spearman's ρ from -0.4486 to -0.5326 for prescription rankings of year 2011 and year 2013. Betweenness centrality measures that computed on the undirected network also indicate a strong negative correlation with prescription rankings of year 2011 and year 2013. The volume of the reviews plays an important role in predicting the medication prescription rankings with Kendall's τ from -0.2133, -0.3768 to -0.4071 and Spearman's ρ from -0.3338, -0.533 to -0.5326 for year 2009, 2011 and 2013 respectively. With moderate number of reviews for medication before year 2009, PageRank, weighted PageRank, HITS authority, centrality measurement and volume of reviews fail to demonstrate their significant correlations with actual prescription rankings. Our proposed ATIO system presents the strongest negative correlation with prescription rankings for year 2009 with Kendall's τ equals to -0.2267 and Spearman's ρ equals to -0.3615.

For year 2013 and 2011, non-graph benchmark method number of reviews achieves close and lower rank correlation than our proposed method. It is reasonable that the volume number of reviews indicates the medication prescription amounts. Increasing prescriptions from physicians could result in increasing number of patients, which would reflect in increasing number of patients leaving reviews on various social medias. There are two possible reasons that number of reviews receives lower Kendall's τ and Spearman's ρ than our method. Firstly, the experimental results are based on the reviews crawled from single social media WebMD. Inadequate review variety would negatively affect the performance of our proposed method, since single review source might be biased towards certain medication than others. Secondly, since we used reviews to predict aggregated yearly prescription rankings in our experiment, yearly rankings with aggregated large amount of reviews would whittle down the advantages of our method. With moderate amount of reviews, our proposed method performs best comparing to all benchmark methods. For example, to predict prescription rankings for year 2009, the medication network structure was based on reviews only from 2007 and 2008 and our method outperformed all other methods for prescription rankings prediction in 2009.

Overall, our proposed system achieves the strongest negative correlation with prescription rankings comparing to other network- based measures such as PageRank, weighted PageRank, HITS authority and Betweenness centrality for all three years. Moreover, with moderate number of reviews for medication before year 2009, ATIO method also demonstrates the significant negative correlation with the actual prescription rankings.

Predictive Modeling and Results

To further validate the predictive power of the MRN based ranking measures on top of existing medication review metrics in the literature, we build a set of prediction linear regression models. To demonstrate the predictive power, we apply variables from year 2007 to year 2008, from year 2007 to year 2010 and from year 2007 to year 2012 to predict the percentage of medication prescription changes for year 2009, year 2011 and 2013 respectively. We compare the model performances in terms of Mean Squared Error (MSE), R^2 and adjusted R^2.

$$\log(\text{Prescription}_t) = \alpha_{t-1} + \beta_1 \text{NumReview}_{t-1} + \beta_2 \text{EaseofUse}_{t-1} + \beta_3 \text{Effective}_{t-1} + \beta_4 \text{Satisfy}_{t-1} + \varepsilon_t \quad (1)$$

$$\log(\text{Prescription}_t) = \alpha_{t-1} + \beta_1 \text{NumReview}_{t-1} + \beta_2 \text{EaseofUse}_{t-1} + \beta_3 \text{Effective}_{t-1} + \beta_4 \text{Satisfy}_{t-1} + \beta_5 \log(\text{NumTransition})_{t-1} + \varepsilon_t \quad (2)$$

$$\log(\text{Prescription}_t) = \alpha_{t-1} + \beta_1 \text{NumReview}_{t-1} + \beta_2 \text{EaseofUse}_{t-1} + \beta_3 \text{Effective}_{t-1} + \beta_4 \text{Satisfy}_{t-1} + \beta_5 \log(\text{NumTransition})_{t-1} + PageRank_{t-1} + WeightedPageRank_{t-1} + \varepsilon_t \quad (3)$$

$$\log(\text{Prescription}_t) = \alpha_{t-1} + \beta_1 \text{NumReview}_{t-1} + \beta_2 \text{EaseofUse}_{t-1} + \beta_3 \text{Effective}_{t-1} + \beta_4 \text{Satisfy}_{t-1} + \beta_5 \log(\text{NumTransition})_{t-1} + PageRank_{t-1} + WeightedPageRank_{t-1} + \log(G)_{t-1} + \varepsilon_t \quad (4)$$

Model	Predictors	MSE	R^2	Adjusted R^2	F – stat Signifance
1	NumReview, EaseofUse, Effective, Satisfy	0.3114	0.4359	0.4022	<0.001
2	NumReview, EaseofUse, Effective, Satisfy, NumTransition	0.2994	0.4577	0.4166	<0.001
3	NumReview, EaseofUse, Effective, Satisfy, NumTransition, PageRank, WeightedPageRank	0.2749	0.4851	0.4279	<0.001
4	NumReview, EaseofUse, Effective, Satisfy, NumTransition, PageRank, WeightedPageRank, TransitionProbability(G)	0.2633	0.5068	0.4431	<0.001

Table 8. Predictive Modeling Results

The variables (predictors) included in Model 1 are number of reviews, ease of use, effective and satisfaction level of each medication. Number of reviews refers to the volume of reviews for each medication. Ease of use, effective and satisfaction level refer to the numeric ratings that are given by each customer on top of their review evaluating the medication from these three aspects, which are variables directly extracted from medication reviews. Each patient gave the ratings to the medication from three aspects and the variables used in Model 1 are the average rating scores for each medication. In terms of Model 2, we contained an additional predictor $log\ (NumTransition)$. NumTransition is the frequency of transitions from other medications to the focal medication, which is automatically counted after the extraction of medication relations by text mining techniques. For example, after preprocessing the reviews and conducting the medication identification and classification models, two patients in their reviews mentioned they switch their prescriptions from Valium to Xanax. Hence, the volume of NumTransition variable for medication Xanax is 2 according to the medication transition results. Furthermore, the predictors PageRank and weighted PageRank incorporated in the Model 3 are variables that computed based on the MRN. In the last model, G is the estimated transition differences between certain medication's transition in and transition out computed using ATIO system.

We could observe that, when compared with the baseline model 1, the introduction of the NumTransition measure (model 2) resulted from the medication reference extraction and relation classification leads to a decrease in MSE and increase in both R^2 and adjusted R^2. The additional inclusion of MRN-based measures PageRank, weighted PageRank (model 3) further results in a decrease in MSE and improvements in both R^2 and adjusted R^2. By introducing the predictor that is from ATIO system $log\ (TransitionProbability(G))$ in the last model (model 4), we achieve the best model comparing to the prior models with least MSE and best R^2 and adjusted R^2. We also conducted the overall F-test for four regression models with p value <0.001. Therefore, we could conclude that all the models provide better fits than the intercept-only model.

Model	Variables with Significant Coefficients
1	NumReview***
2	NumReview ***, Satisfy*
3	NumReview*, WeightedPageRank*
4	NumReview*, Satisfy**, NumTransition*, WeightedPageRank*, TransitionProbability(G)*

Table 9. Variables with Significant Coefficients (* p<0.1, ** p<0.05 and *** p<0.01)

In Table 9, we present the variables (predictors) with significant coefficients. Among all other predictors such as effectiveness, ease of use or satisfaction level scores patients provided, number of reviews (volume) is the only significant variable that affect the prediction of medication prescription amounts. Comparing to model 1, model 2 includes $log\ (NumTransition)$ as additional predictor and achieves lower MSE and both better R^2 and adjusted R^2. However, based on the results presented in table 9,

$log\ (NumTransition)$ is not significantly affecting the prescription prediction results. Satisfy (satisfaction level score) in model 2, on the other hand, turns to be a significant predictor. Model 3 includes extra network structure based predictors such as PageRank and weighted PageRank. Besides significant predictor NumReview, weighted PageRank plays a significant role in prediction as well. Last but not the least, by further inserting the predictor $log(G)$ that is from ATIO system in our model 4, not only $log(G)$ is significant, but predictors such as satisfy, NumTransition and weighted PageRank turn significant as well.

Conclusion

Medication prescription demands and rankings prediction are challenging tasks that rely highly on historical prescription demands and user data. By integrating both text mining techniques and network analysis, our study identifies the important predictive power of medication relations, that have not been analyzed before, for prescription changes and rankings forecasting. Medication switch and combination are common actions among patients and frequent medication switching and combination behavior could serve as an influential indicator for prescription demand changes and competition detection. However, the medication switching sequences and combination records are difficult to access without patients' full prescription history. To predict the medication prescription rankings with medication relation network, our study first aims at discovering the medication transition and combination patterns from unstructured patient generated contents. We construct two supervised SVM systems for identifying medication references and categorizing transition and combination relations from reviews. We further draw directed network based on medication transition relations and undirected network for combination relations. After transforming the network to transition probability matrix, we propose ATIO system for medication prescription rankings prediction. The proposed method suggests significant strongest negative correlation with the actual rankings and positive influence on the actual prescription demands. Other transition and combination network-based measures such as PageRank, weighted PageRank and HITS authority also demonstrate significant negative correlations with the actual rankings. Furthermore, we employ predictive regressions for percentage of prescription changes prediction. The predictive model incorporating the measure from ATIO system as one of the predictors achieves the least MSE and the highest R^2 and adjusted R^2.

Besides the academic value shown here, the transition and combination network as a construct also provides important practical implications for pharmaceutical companies and medication consumers. Besides employing the network for rankings and demands prediction, it is easier to recognize the most competitive brands for one medication by inspecting the transition network. Through examining the transition probability of the focal medication to its STMs, companies could consider the STMs with the highest transition probability as the most comparative product. This STM product would threat the demands of the focal medication the most. After detecting the competitions, pharmaceutical companies could adopt respective strategies such as better marketing promotions or change of the formula so as to gain back the customers. On the other hand, if pharmaceutical companies identify the TCMs for their focal medications, they could design better marketing strategies by promoting the TCM and focal medication as a bundle. Similarly, by inspecting the MRN, medication consumers could identify the alternatives of their medication. Customers could recognize and learn STMs with highest transition probability and treat them as comparative candidates for switching if they are not completely satisfied with their current focal medication.

Furthermore, in addition to clinical predictions, the transition and combination network can be applied to support important predictions and analysis in other domains. The network could serve as a good decision-support tool for business executives. Customers who purchase and consume merchandises such as beauty, pet or baby products would normally need to re-fill their orders in a time interval. Some of them are willing to try different products fulfilling the same need. For example, after customer finished one shampoo brand, he or she might be willing to try another brand. In addition, shampoo and conditioner are always purchased together. These transition and combination behavior for consumable merchandises are very common and are also reflected in the user generated reviews. By constructing transition and combination network using customers' reviews, business executives could not only recognize their competitions, but also use the network structure as an important feature to predict the merchandise sales and rankings. Existing marketing research seems to focus on brand switching and often use different

models such as Bayesian learning model (Erdem and Keane 1996). We could learn from these models and compare the performance of our model with the current models in our future study.

However, our study also suffers from a few limitations. We highlight several important ones and potential future work in the following. First of all, we only annotated 200 medication reviews from WebMD for medication reference identification and relation classification models building and evaluation. The data sparsity problem would affect our model performance and the further network learning. In addition, WebMD is the only forum we extracted reviews and discussions from, multiple data sources such as AskPatient.com and MedHelp would increase the review volume and also the generality of patients. It is therefore important to extend the variety of data source and increase number of annotated reviews. Secondly, we were able to predict rankings for year 2009, 2011 and 2013 since IMS health only reveal yearly rankings and prescription amounts for psychiatric medications. Monthly rankings and prescription quantities in the future would motivate accurate and dynamic prediction results. Thirdly, we did not focus on establishing causal relationships in the current study. Due to data limitations, we have little information about the important factors such as marketing strategies or patient and physician incentives. The inaccessible of these missing variables prevents us from building causality-based models. Future work in this direction could involve medication transition and combination metrics for differentiating medications and focus on causality-based demand system building. We could further extract the time point of direct switching behavior to answer questions, such as when patients would transit and how long patients could stay on certain medication, because our direct transition is counted from text mining results of review contents. For example, for the time duration of direct transition, we could track certain patient's other medication reviews and check if they stopped the other medication again and how long is the time duration. However, indirect transition is estimated based on probability derivate from the network structure and the duration of each transition is difficult to track. Our method therefore is limited in recognizing the time span of staying on the transited medications.

References

Adams, M. P., and Carol Urban PhD, R. N. 2015. *Pharmacology: Connections to Nursing Practice*. Prentice Hall.

Adusumalli, S., Lee, H., Hoi, Q., Koo, S. L., Tan, I. B., and Ng, P. C. 2015. "Assessment of Web-Based Consumer Reviews as a Resource for Drug Performance," *Journal of medical Internet research* (17:8).

Bian, J., Topaloglu, U., and Yu, F. 2012. "Towards Large-scale Twitter Mining for Drug-Related Adverse Events," in *Proceedings of the 2012 international workshop on Smart health and wellbeing*, ACM, October, pp. 25-32.

Brams, S.J., Mutlu, H. and Ramirez, S.L. 2006. "Influence in Terrorist Networks: From Undirected to Directed graphs," *Studies in Conflict and Terrorism* (29:7), pp. 703-718.

Caster, O., Norén, G. N., Madigan, D., and Bate, A. 2010. "Large-Scale Regression-Based Pattern Discovery: The Example of Screening the WHO Global Drug Safety Database," *Statistical Analysis and Data Mining* (3:4), pp. 197-208.

Dhar, V., Geva, T., Oestreicher-Singer, G. and Sundararajan, A. 2014. "Prediction in Economic Networks," *Information Systems Research* (25:2), pp. 264-284.

Dickstein, M., 2011. "Physician vs. Patient Incentives in Prescription Drug Choice."

Donohue, J. M., Cevasco, M., and Rosenthal, M. B. 2007. "A Decade of Direct-To-Consumer Advertising of Prescription Drugs," *New England Journal of Medicine* (357:7), pp. 673-681.

Edwards, I. R., and Aronson, J. K. 2000. "Adverse Drug Reactions: Definitions, Diagnosis, and Management," *The Lancet* (356:9237), pp. 1255-1259.

El-Iskandarani, M. A., Darwish, S. M., and Hefnawy, M. A. 2013. "Drug Consumption Prediction through Temporal Pattern Matching," in *Proceedings of the World Congress on Engineering* (3).

Erdem, T., & Keane, M. P. (1996). "Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets," *Marketing science* (15:1), pp. 1-20.

Freeman, L.C., 1977. "A Set of Measures of Centrality Based on Betweenness," *Sociometry*, pp. 35-41.

Furberg, C. D., Furberg, B. D., and Sasich, L. D. 2010. *Knowing Your Medications*. Potata Inc.

Galarraga, O., O'Brien, M.E., Gutiérrez, J.P., Renaud-Thery, F., Nguimfack, B.D., Beusenberg, M., Waldman, K., Soni, A., Bertozzi, S.M. and Greener, R. 2007. "Forecast of Demand for Antiretroviral Drugs in Low and Middle-income Countries: 2007–2008," *Aids* (21), pp. S97-S103.

Geisser, S., 1993. *Predictive inference* (Vol. 55). CRC press.

Ghousi, R., Mehrani, S., Momeni, M., and Anjomshoaa, S. 2012 "Application of Data Mining Techniques in Drug Consumption Forecasting to Help Pharmaceutical Industry Production Planning," in *Proceedings of the 2012 International Conference on Industrial Engineering and Operations Management*, pp. 1162-1167

Ginn, R., Pimpalkhute, P., Nikfarjam, A., Patki, A., O'Connor, K., Sarker, A., Smith, K. and Gonzalez, G. 2014. "Mining Twitter for Adverse Drug Reaction Mentions: A Corpus and Classification Benchmark," in *Proceedings of the fourth workshop on building and evaluating resources for health and biomedical text processing*.

Harpaz, R., Haerian, K., Chase, H. S., and Friedman, C. 2010. "Statistical Mining of Potential Drug Interaction Adverse Effects in FDA's Spontaneous Reporting System," in *AMIA Annual Symposium Proceedings* (2010), pp. 281-285.

Jiang, K., and Zheng, Y. 2013. "Mining Twitter Data for Potential Drug Effects," in *Advanced data mining and applications*, Springer Berlin Heidelberg, pp. 434-443.

Kendall, M.G., 1938. "A New Measure of Rank Correlation," *Biometirika* (30:1/2), pp. 81-93.

Kleinberg, J.M., 1999. "Hubs, Authorities, and Communities," *ACM Computing Surveys (CSUR)* (31:4es), pp. 5.

Leaman, R., Wojtulewicz, L., Sullivan, R., Skariah, A., Yang, J. and Gonzalez, G. 2010. "Towards Internet-Age Pharmacovigilance: Extracting Adverse Drug Reactions from User Posts to Health-Related Social Networks," in *Proceedings of the 2010 workshop on biomedical natural language processing*, Association for Computational Linguistics, pp. 117-125.

Liu, X., and Chen, H. 2013. "AZDrugMiner: An Information Extraction System for Mining Patient-Reported Adverse Drug Events in Online Patient Forums," in *Smart Health*, Springer Berlin Heidelberg, pp. 134-150.

Liu, X., Liu, J., and Chen, H. 2014. "Identifying Adverse Drug Events from Health Social Media: A Case Study on Heart Disease Discussion Forums," in *Smart Health*, Springer International Publishing, pp. 25-36.

Page, L., Brin, S., Motwani, R. and Winograd, T., 1999. "The PageRank Citation Ranking: Bringing Order to the Web."

Nikfarjam, A., and Gonzalez, G. H. 2011. "Pattern Mining for Extraction of Mentions of Adverse Drug Reactions from User Comments," in *Proceedings of the American medical informatics association (AMIA) annual symposium* (2011), pp. 1019–1026.

Netzer, O., Feldman, R., Goldenberg, J. and Fresko, M. 2012. "Mine Your Own Business: Market-Structure Surveillance Through Text Mining," *Marketing Science* (31:3), pp. 521-543.

Norén, G. N., Sundberg, R., Bate, A., and Edwards, I. R. 2008. "A Statistical Methodology for Drug–Drug Interaction Surveillance," *Statistics in medicine* (27:16), pp. 3057-3070.

Park, J., and Newman, M. E. 2005. "A Network-Based Ranking System for US College Football," *Journal of Statistical Mechanics: Theory and Experiment* (2005:10), pp. P10014.

Rose, M., and Merk, H. 1985. "Forecasting the Demand for a New Drug in a New Market: When History Does Not Repeat Itself," in *Proceedings of the 3rd International Conference of the System Dynamics Society*, pp. 719-727

Sekhri, N., Chisholm, R., Longhi, A., Evans, P., Rilling, M., Wilson, E. and Madrid, Y., 2006. "Principles for Forecasting Demand for Global Health Products," *Global health forecasting working group, background paper. The Center for Global Development.*

Shin, S., Ahnert, S.E. and Park, J., 2014. "Ranking Competitors Using Degree-Neutralized Random Walks," *PloS one* (9:12), pp. e113685.

Spearman, C., 1904. "The Proof and Measurement of Association Between Two Things," *The American Journal of Psychology* (15:1), pp. 72-101.

Thakrar, B. T., Grundschober, S. B., and Doessegger, L. 2007. "Detecting Signals of Drug–Drug Interactions in A Spontaneous Reports Database," *British journal of clinical pharmacology* (64:4), pp. 489-495.

Van Puijenbroek, E. P., Egberts, A. C., Meyboom, R. H., and Leufkens, H. G. 1999. "Signaling Possible Drug–Drug Interactions in A Spontaneous Reporting System: Delay of Withdrawal Bleeding During Concomitant Use of Oral Contraceptives and Itraconazole," *British journal of clinical pharmacology* (47:6), pp. 689-693.

Van Rijsbergen, C. J. 1979. *Information Retrieval* (2nd ed.). Butterworth.

Ventola, C.L., 2014. "Social media and health care professionals: benefits, risks, and best practices," *Pharmacy and Therapeutics* (39:7), pp. 491.

Xing, W. and Ghorbani, A., 2004. "Weighted PageRank algorithm," in *Proceedings of Second Annual Conference on on Communication Networks and Services Research*, IEEE, May, pp. 174-194.

Xu, K., Liao, S.S., Li, J. and Song, Y. 2011. "Mining Comparative Opinions from Customer Reviews for Competitive Intelligence," *Decision support systems* (50:4), pp. 743-754.

Xu, H., Stenner, S.P., Doan, S., Johnson, K.B., Waitman, L.R. and Denny, J.C. 2010. "MedEx: A Medication Information Extraction System for Clinical Narratives," *Journal of the American Medical Informatics Association* (17:1), pp. 19-24.

Yang, M., Wang, X., and Kiang, M. Y. 2013. "Identification of Consumer Adverse Drug Reaction Messages on Social Media," in *PACIS*, pp. 193.

Yang, C. C., Yang, H., and Jiang, L. 2014. "Postmarketing Drug Safety Surveillance Using Publicly Available Health-Consumer-Contributed Content in Social Media," *ACM Transactions on Management Information Systems (TMIS)* (5:1), pp. 2.

Yates, A., and Goharian, N. 2013. "ADRTrace: Detecting Expected and Unexpected Adverse Drug Reactions from User Reviews on Social Media Sites," in *Advances in Information Retrieval*, Springer Berlin Heidelberg, pp. 816-819.

Yeleswarapu, S., Rao, A., Joseph, T., Saipradeep, V. G., and Srinivasan, R. 2014. "A Pipeline to Extract Drug-Adverse Event Pairs from Multiple Data Sources," *BMC medical informatics and decision making* (14:1), pp. 1.

Zhang, Z., Guo, C. and Goes, P. 2013. "Product Comparison Networks for Competitive Analysis of Online Word-of-Mouth," *ACM Transactions on Management Information Systems (TMIS)* (3:4), p. 20.