A Temporal Interestingness Measure for Drug Interaction Signal Detection in Post-marketing Surveillance

Yanqing Ji, Hao Ying, *Fellow, IEEE,* John Tran, Peter Dews, Ayman Mansour, R. Michael Massanari

***Abstract*—Drug-drug interactions (DDIs) can result in serious consequences, including death. Existing methods for identifying potential DDIs in post-marketing surveillance primarily rely on the FDA's (Food and Drug Administration) spontaneous reporting system. However, this system suffers from severe underreporting, which makes it difficult to timely collect enough valid cases for statistical analysis. In this paper, we study how to signal potential DDIs using patient electronic health data. Specifically, we focus on discovery of potential DDIs by analyzing the temporal relationships between the concurrent use of two drugs of interest and the occurrences of various symptoms using novel temporal association mining techniques we developed. A new interestingness measure called *functional temporal interest* was proposed to assess the degrees of temporal association between two drugs of interest and each symptom. The measure was employed to screen potential DDIs from 21,405 electronic patient cases retrieved from the Veterans Affairs Medical Center in Detroit, Michigan. The preliminary results indicate the usefulness of our method in finding potential DDIs for further analysis (e.g., epidemiology study) and investigation (e.g., case review) by drug safety professionals.**

I. Introduction

Drug-drug interactions (DDIs) represent a significant public health issue [1]. They can complicate a patient's medical condition, increase morbidity, and even result in death. A prospective analysis of 18,820 patients by Pirmohamed et al. indicated that DDIs contributed to 1% of all hospital admissions [2]. In another study, up to 2.8% of admissions were found to be caused by DDIs [3]. Lepori et al. found that, in a Swiss hospital, 21% of all drug-related hospital admissions were caused by DDIs, which contributed to 1.3% of all admissions [4].

DDIs are often not recognized in pre-marketing clinical trials because the size and duration of these trials are necessarily limited and the concurrent use of medications is well controlled in the trials [5]. When two or more drugs are used together in a real-world clinical setting, they may cause various DDIs. At present, thousands of drugs are on the US market and evidence on adverse drug effects including DDIs is generated primarily by the FDA's post-marketing surveillance system MedWatch. MedWatch is a ***passive*** system that depends on voluntary, spontaneous reports of suspected adverse effects to be filed at the discretion of healthcare professionals, drug manufactures, and consumers. Because of the voluntary nature of the reporting process, there is a serious underreporting [6]. Consequently, the accumulation of sufficient cases to enable a critical analysis is very slow, which delays the identification and withdrawal of problematic drugs from the market or labeling them with black box warnings. These delays have resulted in unnecessary mortality, morbidity, and costs of healthcare.

In the literature, there only exist several studies that attempted to use data mining methods to detect potential DDIs. For instance, Almenoff et al. studied the value of using disproportionality reporting to identify known DDIs that occur in the spontaneous reporting database at higher than expected frequencies [5]. They were able to discover the well-known DDIs between beta blockers and the calcium channel blocker verapamil. They concluded that their method was a promising tool, but the interpretation of the results had to be very cautious since many factors might affect the reporting rate of an adverse effect. Statistical or probability methodologies have been used by a couple of other studies to discover potential DDIs [7, 8]. All these studies were based on spontaneous reports. Thus, the performance of these techniques could be highly situation-dependent due to the weaknesses and potential biases inherent in spontaneous reporting [9].

In recent years, electronic patient records become more easily accessible in various healthcare organizations. They include huge amount of "event" data such as diagnoses and dispensing of drugs. By analyzing the temporal relationships among these event data, it is possible to find the adverse effects that might be caused by the interaction of two or more drugs. Thus, these data provide a new valuable source of information from which potential DDIs could be generated more effectively and much earlier.

In this paper, we study how to mine potential DDIs from electronic patient data using temporal association mining techniques. Specifically, we introduce *degree* of temporal association between two events whose value is within [0, 1] where 0 indicates no temporal association and 1 represents full temporal association. The value is determined by a function defined on $[0, T]$ where T represents the length of a time window beyond which the two events will have no temporal association in a sequence. We then introduced a new interestingness measure called *functional temporal interest*

Y. Ji is with the Department of Electrical and Computer Engineering, Gonzaga University, Spokane, WA 99258 USA (corresponding author; phone:509-313-3529; fax:509-313-5871; email: ji@gonzaga.edu).

H. Ying is with the Department of Electrical and Computer Engineering, Wayne State University, Detroit, MI 48202, USA (email: hao.ying@wayne.edu).

J. Tran is with the Frontier Behavioral Health, Spokane, WA 99202, USA (email: jtran@smhca.org).

P. Dews is with the Department of Medicine, St. Mary Mercy Hospital (Trinity Health), Livonia, MI 48154, USA (email:dewsp@trinity-health.org).

A. Mansour is with Department of Electrical and Computer Engineering, Tafila Technical University, Tafila 66110, Jordan (email: mansour@ttu.edu.jo).

R. Michael Massanari is with the Research for the Critical Junctures Institute, Bellingham, WA 98225, USA (email: Michael.Massanari@wwu.edu).

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

where the contribution of each sequence to the measure depends on the degree of temporal association between the two events (or event sets) of interest within that sequence. This measure was employed to mine *functional temporal association rules* from relational electronic databases. The effectiveness of our data mining strategy was evaluated using electronic patient data retrieved from the Veterans Affairs Medical Center in Detroit.

II. Background on Temporal Association Mining

Temporal association mining is a natural extension of the traditional association rule mining, where an association rule is an implication expression in the form of $X \rightarrow Y$. X and Y represent two event sets and they are disjoint (i.e., $X \cap Y = \emptyset$), meaning that they share no common events. An association rule indicates that the presence of X implies the presence of Y. If X and Y have temporal relationship, a temporal constraint can be applied to the association rule. Such an association rule, represented as $X \xrightarrow{T} Y$, is called temporal association rule, where $\xrightarrow{T}$ denotes that Y occurs after X within a time window of length T in the same event sequence.

Rather than simply mining the co-occurrence between X and Y, temporal association mining allows to investigate contextual and temporal relationships, some of which may indicate a cause-effect association since the concept of causality is linked to time dependencies. That is, the causes must precede their effects. Many researchers have developed various measures and algorithms to mine different types of temporal data, especially in medical domain where finding the potential causal factors for particular medical conditions is a fundamental objective [10-12]. For instance, Jin et al. attempted to mine unexpected temporal associations with applications in signaling potential adverse drug reactions caused by a single drug using administrative health databases [11]. Noren et al. proposed another temporal association mining method which contrasts the observed-to-expected ratio in a time period after X to the observed-to-expected ratio in a control period before X [10]. More recently, Concaro et al. extended traditional temporal association mining by handling both point-like events and interval-like events (e.g., drug consumption) [12].

The above approaches suffer from the sharp boundary problem. On the one hand, the events near the time boundaries are either ignored or overemphasized. As indicated by the *sequence1* in Figure 1, if we assume the time window T is equal to 60 days and two events Y_1 and Y_2 occur 59 days and 61 days, respectively, after the occurrence of X, then only event Y_1 is considered as having temporal association with X using their approaches. But the two events Y_1 and Y_2 should have similar levels of association with X since the two numbers, 59 and 61, are close. On the other hand, two events are considered to contribute equally to an interestingness measure in the above approaches as long as they occur within the time window T. The two events Y_3 and Y_4 in Figure 1 are considered as having the same (full) temporal association with X, even though Y_3 occurs much earlier than Y_4 after X. That is, the length of the time duration between two events has no effect on the interestingness measure. This does not reflect reality. For example, if two adverse effects A and B occur 2

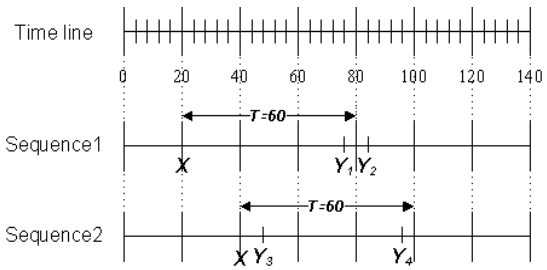

Figure. 1. Sharp boundary problem: 1) in Sequence1, only Y_1 is considered as having temporal association with X, even though Y_2 is close to Y_1; 2) in Sequence2, Y_3 and Y_4 are considered as having the same degree of temporal association with X, even though Y_3 occurs much shorter after X than Y_4.

days and 59 days, respectively, after a patient's exposure to two drugs, then, intuitively adverse effect A is more likely to be causally associated with the potential interaction of the two drugs and thus should contribute more to the interestingness measure than adverse effect B. Besides the sharp boundary problem, to the best of our knowledge, none of the existing temporal association mining methods has been applied to the DDI problem which represents our interest in this paper.

III. A Novel Functional Temporal Interest Measure

To overcome the sharp boundary problem of traditional temporal association mining, we introduce a category of functional temporal association rules, denoted by $X \xrightarrow{\varphi(t)|t\in[0,T]} Y$, meaning that Y occurs after X within a time window of length T and the degree of temporal association between X and Y is determined by a function $\varphi(t)$: $[0,T] \rightarrow [0,1]$ where $t \in [0,T]$ and represents the time duration between X and Y. Both $\varphi(t)$ and T are application-dependent and should be chosen by users or domain experts. $\varphi(t)$ defines the likelihood that Y is caused by X along time after the occurrence of X.

Next, we will define a new interestingness measure based on the concept of degree of temporal association. Our new measure extends an existing objective measure called interest factor defined as follows:

$$I(X,Y) = \frac{supp(X,Y)}{supp(X) \times supp(Y)} \quad (1)$$

where $supp$ represents the support of an event set and is defined as the proportion of sequences in which an event set or a pattern occurs at least once, among all the event sequences. The above measure compares the frequency of a pattern against a baseline frequency obtained under the statistical independence assumption. The measure indicates an association if its value is larger than 1. Next, we present how to extend this measure by incorporating degree of temporal association.

We define the support of a functional temporal association rule, $supp\left(X \xrightarrow{\varphi(t)|t\in[0,T]} Y\right)$, as the accumulated degrees of temporal association (with respect to $< X, Y >$) over all sequences divided by the total number of sequences. That is,

$$supp\left(X \xrightarrow{\varphi(t)|t\in[0,T]} Y\right) = \sum_{i=1}^{N} \varphi(t_i) \Big/ N \quad (2)$$

where $\varphi(t)$ defines the degree of temporal association over $[0, T]$ and t_i represents the time duration between X and Y in the i^{th} sequence. N is the total number of sequences. In general, $\varphi(t)$ takes a value between 0 and 1 when Y occurs within a *T-sided* window after X. $\varphi(t)$ is 0 for a sequence in the following two situations; 1) the sequence does not contain both X and Y; 2) the sequence contains both X and Y, but the occurrence of Y is not within the *T-sided* window after X. If $\varphi(t)$ is 1 for all the sequences within which Y occurs within a *T-sided* window after X, then the above definition becomes the support of a normal temporal association rule. In this sense, our definition is a generalization of the traditional definition of support for a temporal association rule.

Let us examine how the above definition of support can be applied to analyzing the temporal association between exposure to two drugs and a symptom, which is not trivial. First, the two drugs do not simply coexist within a patient record which is considered as an event sequence. The two drugs themselves, as the antecedent of a temporal association rule, have a temporal relationship. That is, the consumptions of the two drugs much overlap in time. Otherwise, no adverse effect will be expected to be caused by their potential interactions. In this context, drug consumptions are considered duration-like events and their relationships must be examined. Second, a proper function $\varphi_i(t)$ must be defined. Note that we only consider interactions of two drugs in this paper since, for the interactions of more than two drugs, it is very difficult to get the evidence for evaluating the results and obtain an appropriate interpretation from clinical practice.

In the following description, we utilize D and S to represent a drug and a symptom, respectively. Equation (2) can be transformed to an equation as follows:

$$supp\left(\left\{D \overset{\Delta}{\rightarrow} D'\right\} \xrightarrow{\varphi(t)|t\in[0,T]} \{S\}\right) = \sum_{i=1}^{N} \varphi(t_i) \Big/ N_D \qquad (3)$$

where D and D' represent two drugs that may interact and Δ is a temporal operator. In this particular application, Δ represents *overlap*. t_i represents the time duration between exposure to the two drugs and the occurrence of the symptom in the i^{th} sequence. N_D is the total number of patients who have taken drug D. The function $\varphi(t)$ is defined below:

$$\varphi(t) = \begin{cases} -\frac{t}{T} + 1 & 0 < t < T \\ 0 & otherwise \end{cases} \qquad (4)$$

where T is a time window beyond which the two drugs will be considered having no temporal association with the symptom. The above definition of $\varphi(t)$ indicates that if the adverse effect occurs within a shorter time after taking the two drugs, it is more likely to be caused by the drug interaction.

Based on (1) and (3), we define a new interestingness measure called *functional temporal interest* (*FTI*) as follows:

$$FTI = \frac{supp\left(\left\{D \overset{\Delta}{\rightarrow} D'\right\} \xrightarrow{\varphi(t)|t\in[0,T]} \{S\}\right)}{supp\left(\left\{D \overset{\Delta}{\rightarrow} D'\right\} \rightarrow\right) \times supp\left(\left\{D \overset{\Delta}{\rightarrow} *\right\} \xrightarrow{\varphi(t)|t\in[0,T]} \{S\}\right)} \qquad (5)$$

where $supp\left(\left\{D \overset{\Delta}{\rightarrow} D'\right\} \rightarrow\right)$ represents the proportion of cases in which D overlaps D'. $supp\left(\left\{D \overset{\Delta}{\rightarrow} *\right\} \xrightarrow{\varphi(t)|t\in T} \{S\}\right)$ is the support for D with any drug other than D' relative to symptom S. The symbol * represents any drug other than the two drugs of interest. Both $supp\left(\left\{D \overset{\Delta}{\rightarrow} D'\right\} \rightarrow\right)$ and $supp\left(\left\{D \overset{\Delta}{\rightarrow} *\right\} \xrightarrow{\varphi(t)|t\in T} \{S\}\right)$ differ from the corresponding factors in the original definition of interest where the former instead would have included the proportion of cases in which D and D' coexist and the later would have included the support for the symptom S. That is, both terms only involve a subset of cases that their original definitions would have covered. The new definitions emphasize the *interaction* of two drugs – the time duration of two drugs must overlap. Moreover, they can reduce the search space since the calculation of both terms would be limited to the cases that contain at least one drug.

Given two drugs of interest, their FTI values relative to each symptom can be computed using (5). The higher the FTI value, the more likely it is for a particular symptom to be associated with the interaction of the two drugs.

IV. Experiments

A. Experiment Data

To evaluate our new interestingness measure , we retrieved electronic patient data from Veterans Affairs Medical Center in Detroit, Michigan after the IRB (Institutional Review Board) protocol was approved. The de-identified electronic data included drug prescriptions from year 2005 to 2010. "Event" data such as dispensing of drug, office visits, and certain laboratory tests were retrieved for all the patients. For each event certain details were obtained. For example, the data for dispensing of drug include the name of the drug, quantity of the drug dispensed, dose of the drug, drug start date, drug schedule, and the number of refills.

The retrieved data included 21,405 patients. 20,507 (95.8%) were male, and 898 (4.2%) were female. Their average age was 66.7. The drugs benazepril and losartan were selected to test the proposed data mining framework since their interactions are known and our physicians are familiar with these drugs. The total number of patients who took one of the two drugs benazepril and losartan at least once is 1203, and 2086, respectively. The total number of distinctive ICD-9 (International Classification of Diseases, 9th Revision) codes was 4,472 in the data. Note that, in electronic health databases, symptoms are coded using ICD-9 codes.

B. Preliminary Results

The *FTI* values were calculated for all the functional temporal association rules which were formed by the two drugs of interest and each of the 4,472 distinctive ICD-9 codes. Since different ICD-9 codes may represent the same (or similar) diagnoses, we clustered the data mining results into a manageable number of categories based on the Clinical Classifications Software (CCS) for ICD-9-CM Fact Sheet [13]. The CCS was developed by the Agency for Healthcare Research and Quality (AHRQ). It groups over 14,000 ICD-9 codes into 285 mutually exclusive and clinically meaningful

TABLE I CLINICAL CLASSIFICATIONS CATEGORIES FOUND TO INDICATE POSSIBLE OR VERY LIKELY DDIS ASSOCIATED WITH DRUGS BENAZEPRIL AND LOSARTAN AMONG THE TOP 20 CLASSIFICATION CATEGORIES RANKED ACCORDING TO THE ACCUMULATIVE *FTI* VALUES

Rank	Clinical classification category description	Accumulative FTI value	DDI?
2	Mood disorders	19.96	Possible
5	Delirium/dementia/amnestic/other cognitive	12.11	Possible
13	Acute renal failure	4.18	Very likely
16	Chronic renal failure	2.68	Very likely

categories. We accumulated the *FTI* values of the ICD-9 codes that belonged to the same CCS category and then ranked the resulting categories according to the accumulated *FTI* values. The top 20 categories were evaluated by the physicians on our project team. The physicians used one of the four linguistic terms (i.e., "unlikely", "possible", "probably" and "very likely") to describe the potential causal association between two drugs and a symptom. While 2 of the top 20 categories were found to be "possible" DDIs, 2 of them were "very likely" DDIs associated with drugs benazepril and losartan. These categories as well as their corresponding accumulated *FTI* values and rankings are shown in Table I

To establish the value of our new *FTI* measure, we compared the ranks generated by our measure and three other measures for the four CCS categories which are considered to be "very likely" or "possible" DDIs. These three measures are temporal interest (*TI*), traditional interest (*I*) measure without considering temporal information, and the standard definition of support (*supp*). The definition of the *TI* measure is similar to our *FTI* measure except that it does not incorporate the concept of degree of association which is defined by a function. That is, the temporal association between the two drugs of interest and a symptom within a patient case is either 1 or 0, depending on whether the symptom occurs within the time window *T* after the two drugs taken together. The *I* measure is defined by (1). Table II presents the ranks for the four CCS categories generated by these four different interestingness measures. It indicates that the *FTI,* and *TI* measures performed much better when compared to the traditional *I* and *supp* measures. This implies that incorporating temporal information into a measure can greatly improve its performance. Our *FTI* measure obtained higher ranks than the *TI* measure for all the four CCS categories. The top 20 CCS categories ranked by the *TI* measure were also evaluated by the physicians on our project team. Only two categories (i.e. "delirium/dementia/amnestic/ other cognitive" and "acute renal failure") were found to have "possible" or "very likely" association with the two drugs of interest. All these two categories were also among the top 20 CCS categories identified by the *FTI* measure as shown in Table II. Thus, our measure generally preforms better than the *TI* measure since it has a better capability to capture the temporal information.

TABLE II RANKS FOR FOUR CCS CATEGORIES USING FOUR DIFFERENT MEASURES

Clinical classification category description	Ranks for three different measures			
	FTI	*TI*	*I*	*supp*
Mood disorders	2	13	65	41
Delirium/dementia/amnestic/other cognitive	5	9	97	100
Acute renal failure	13	18	99	76
Chronic renal failure	16	22	42	28

V. CONCLUSION

We introduced functional temporal association rules where the strength of temporal association between two events within a sequence is defined by a function. We have developed a new interestingness measure that incorporates the degrees of temporal associations among all sequences into one single value. Our preliminary results indicated that four known DDIs were ranked high among all the 285 clinically meaningful categories, each of which represents a potential DDI caused by drugs benazepril and losartan.

REFERENCES

[1] Hines LE, Murphy JE, Grizzle AJ, Malone DC. Critical issues associated with drug-drug interactions: highlights of a multistakeholder conference. American journal of health-system pharmacy : AJHP : official journal of the American Society of Health-System Pharmacists. 2011;68:941-6.

[2] Pirmohamed M, James S, Meakin S, Green C, Scott AK, Walley TJ, et al. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. BMJ (Clinical Research Ed). 2004;329:15-9.

[3] Jankel CA, Fitterman LK. Epidemiology of drug-drug interactions as a cause of hospital admissions. Drug Safety: An International Journal Of Medical Toxicology And Drug Experience. 1993;9:51-9.

[4] Troiano L, Scibelli G, Birtolo C. A Fast Algorithm for Mining Rare Itemsets. Proceedings of the 2009 Ninth International Conference on Intelligent Systems Design and Applications: IEEE Computer Society; 2009. p. 1149-55.

[5] Almenoff JS, DuMouchel W, Kindman LA, Yang X, Fram D. Disproportionality analysis using empirical Bayes data mining: a tool for the evaluation of drug interactions in the post-marketing setting. Pharmacoepidemiology And Drug Safety. 2003;12:517-21.

[6] Hazell L, Shakir SA. Under-reporting of adverse drug reactions : a systematic review. Drug safety : an international journal of medical toxicology and drug experience. 2006;29:385-96.

[7] Noren GN, Sundberg R, Bate A, Edwards IR. A statistical methodology for drug-drug interaction surveillance. Statistics in medicine. 2008;27:3057-70.

[8] Tatonetti NP, Fernald GH, Altman RB. A novel signal detection algorithm for identifying hidden drug-drug interactions in adverse event reports. Journal of the American Medical Informatics Association : JAMIA. 2012;19:79-85.

[9] Hauben M. Early postmarketing drug safety surveillance: data mining points to consider. The Annals of pharmacotherapy. 2004;38:1625-30.

[10] Norén GN, Hopstadius J, Bate A, Star K, Edwards IR. Temporal pattern discovery in longitudinal electronic patient records Data Mining and Knowledge Discovery. 2010;20:361-87.

[11] Jin H, Chen J, He H, Williams G, Kelman C, O'Keefe C. Mining unexpected temporal associations: applications in detecting adverse drug reactions. IEEE Transactions on Information Technology in Biomedicine. 2008;12:488-500.

[12] Concaro S, Sacchi L, Cerra C, Fratino P, Bellazzi R. Mining health care administrative data with temporal association rules on hybrid events. Methods of information in medicine. 2011;50:166-79.

[13] AHRQ. Clinical Classifications Software (CCS) for ICD-9-CM Fact Sheet. Healthcare Cost and Utilization Project (HCUP)2007.