

# SEARCHING ADVERSE DRUG REACTION SIGNAL PAIRS USING MINING ASSOCIATION

P.AMBIKA<sup>1</sup>

<sup>1</sup>Anna University-(Chennai), Holycross Engineering College, Vagaikulam, Tuticorin Dist, Tamil Nadu, India

## Abstract

In this paper, propose an innovative data mining framework and apply it to mine potential causal associations in electronic patient data sets where the drug-related events of interest occur infrequently. Specifically, created a novel interestingness measure, exclusive causal-leverage, based on a computational, fuzzy recognition-primed decision (RPD) model that we previously developed. On the basis of this new measure, a data mining algorithm was developed to mine the causal relationship between drugs and their associated adverse drug reactions (ADRs). The exclusive causal-leverage was employed to rank the potential causal associations between each of the selected drugs. Algorithm could effectively make known ADRs rank high among all the symptoms in the database.

**Keywords:** *Adverse drug reactions, Recognition-primed decision, causal-leverage.*

## 1. Introduction

Finding causal associations between two events or sets of events with relatively low frequency is very useful for various real-world applications. For example, a drug used at an appropriate dose may cause one or more adverse drug reactions (ADRs), although the probability is low. Discovering this kind of causal relationships can help us prevent or correct negative outcomes caused by its antecedents. However, mining these relationships is challenging due to the difficulty of capturing causality among events and the infrequent nature of the events of interest in these applications.

In this paper, try to employ a knowledge-based approach to capture the degree of causality of an event pair within each sequence. since the determination of causality is often ultimately application or domain dependent. Then develop an interestingness measure that incorporates the causalities across all the sequences in a database. This study was motivated by the need of discovering ADR signals in post marketing surveillance, even though the

proposed framework can be applied to many different applications.

## 2. Modules

Data load and Pre-processing, Fuzzy RDP model, Pair Generation and Causal leverage are the modules which are as follows

### 2.1 Fuzzy RDP Model

The fuzzy RPD model was preliminarily validated in our previous study by using it to calculate the extent of causality between cisapride and some of its adverse effects. We used the real patients to create simulated patient cases, all of which containing drug-symptom pairs of interest with various degrees of causality. The model's validity was then established by comparing the decisions made by the model and those by two independent experienced physicians for the 100 simulated patients. The levels of agreements were measured by the weighted Kappa statistic, which is a measure of agreement between two raters after chance agreement is controlled.

### 2.2. Pair Generation and Evaluation

This the process for pair generation and evaluation. In this algorithm drug-symptom pairs that can be easily generated. The pairs are drug-drug pairs, symptom-symptom pairs or combinations of multiple drugs and symptoms. Thus, this algorithm generates a much fewer number of candidate rules, which implies much less complexity.

### 2.3 causal-leverage

This algorithm shows how to compute the causal-leverage value of a general pair between event X and Y. Both X and Y could be either drug event or symptom event. First, the drug or symptom hash table is searched in order to get the support count for event Y. Then, for each PID that supports the pair, a process called cue abstraction is used to extract a set cue values V from the related patient case. Specifically, a list of drug start dates and a list of symptom dates are retrieved from the Patient Drug Table and the Patient Symptom Table, respectively. Finally rank all the pairs in a decreasing order according to their exclusive causal-leverage values after all these values are computed.

## 3. Tables, Figures and Equations

### 3.1 Tables and Figures

Cues	Cue value set 1	Cue value set 2	Cue value set 3	Cue value set 4
Temporal association	Likely	Likely	Possible	Unlikely
Rechallenge	Likely	Possible	Unlikely	Unlikely
Dechallenge	Likely	Likely	Possible	Unlikely
Other explanations	No	No	No	Yes
<b>Causality Category</b>	Very likely	Probable	Possible	Unlikely

As shown, an experience consists of four components—cues, goals, actions, and expectancies. Cues represent the higher level information (synthesized from elementary or environmental data) that a decision maker must pay attention to. Expectancies describe what will happen next as the current situation continues to evolve in a changing context. Goals represent an end state that the decision maker is trying to achieve. Actions represent a set of potential decisions that the decision maker can take in the current situation. Cues are used to match the current situation with prior experiences and determine which experience can be reused to solve a new problem. This sample experience has four cues: temporal association, dechallenge, rechallenge, and other explanation. The first three cues have been explained in Introduction. Other explanations denote alternative explanations by concurrent disease or other drugs.

The type of cue could be quantitative, nominal or fuzzy in the proposed computational fuzzy RPD model. For instance, the cue “temporal association” may have fuzzy values (e.g., unlikely, possible, likely). The weights for these cues are design parameters and are assigned by domain experts. Table 1 shows how the four cues are related to degree of causality of a signal pair in the four experiences.

These mappings were given by the physicians in our research project. For instance, when cues temporal association, rechallenge and dechallenge have fuzzy cue values possible, unlikely, and possible, respectively, and there is no other explanations, this cue value pattern represents a possible causal association between the drug of interest and the suspected ADR from the perspective of a physician.

### 3.2 Equations

The similarity between two sets of cue values V and V 0 is named as global similarity  $SG_{\delta V ; V 0}$ . It is defined as the weighted sum of all the local similarities with respect to each pair of cue values. That is, where  $w_i \geq 0$ ;  $\sum w_i = 1$  is the weight for cue i, which represents the relative significance of the cue and is assigned by the user.

The above global similarity is used to find the most matching experience whose associated causality category can characterize the causal association of a pair of interest in a particular event sequence. We use this approach to obtain a similarity value between the current pair and each of the experiences. After that, these similarity values are normalized so that their sum is equal to 1. These normalized values are then used to represent the membership values of corresponding categories for the pair of interest in a particular sequence. In general, if there exist m experiences that classifies the causality between X and Y into m distinctive categories, the degree of causality is defined as where  $\mu_i$  is the membership of the ith causality category for the pair, and  $w_i$  represent the corresponding weight when converging all the causality categories into one.

The selection of  $w_i$  is based on two considerations. First, causality categories representing stronger causal associations should have higher weights. That is, if we assume the causality levels represented by m experience are in a decreasing order, then  $w_m > w_{m-1} > w_2 > w_1$  must be satisfied, where  $w_m$  and  $w_1$  correspond to the highest and lowest causality levels, respectively. Second, the range

of  $C\langle X;Y \rangle$  should be  $[0, 1]$ . That is,  $C\langle X;Y \rangle$  should be 0 for the extreme situation  $\frac{1}{4} f_0; \dots; 0; 1g$  where the evidence in a sequence strongly shows “unlikely” association of the pair. If all the evidence in the patient supports “very likely” association (i.e.,  $\frac{1}{4} f_1; 0 \dots 0g$ ),  $C\langle X;Y \rangle$  should be 1. Otherwise,  $C\langle X;Y \rangle$  is between 0 and 1

S.M. Wolfe, and D.H. Bor, “Timing of New Black Box Warnings and Withdrawals for Prescription Medications,” J. Am. Medical Assoc., vol. 287, pp. 2215-2220, 2002.

#### 4. Conclusions

However, mining these associations is very difficult especially when events of interest occur infrequently. We have developed a new interestingness measure, exclusive causal-leverage, based on an experience-based fuzzy RPD model. This measure can be JI ET AL.: A METHOD FOR MINING INFREQUENT CAUSAL ASSOCIATIONS AND ITS APPLICATION IN FINDING ADVERSE DRUG REACTION... 731 . A data mining algorithm was developed to search a real electronic patient database for potential ADR signals. Experimental results showed that our algorithm could effectively make known ADRs rank high among all the symptoms in the database.

#### ACKNOWLEDGMENTS

The authors would like to thank Ellison Floyd and Richard E. Miller for retrieving the patient information from the VAMC Detroit’s databases. This work was supported in part by National Institutes of Health under grant 1 R21

#### References

- [1] J. Talbot and P. Waller, Stephens’ Detection of New Adverse Drug Reactions, fifth ed. John Wiley & Sons, 2004.
- [2] L.T. Kohn, J.M. Corrigan, and M.S. Donaldson, To Error is Human: Building a Safer Health System. Nat’l Academy Press, 2000.
- [3] J. Lazarou, B.H. Pomeranz, and P.N. Corey, “Incidence of Adverse Drug Reactions in Hospitalized Patients: A Meta-Analysis of Prospective Studies,” J. Am. Medical Assoc., vol. 279, pp. 1200-1205, Apr. 1998.
- [4] K. Hartmann, A.K. Doser, and M. Kuhn, “Postmarketing Safety Information: How Useful are Spontaneous Reports?,” Pharmacoepidemiology and Drug Safety, vol. 8, pp. 65-71, 1999.
- [5] L. Hazell and S.A.W. Shakir, “Under-Reporting of Adverse Drug Reactions - A Systematic Review,” Drug Safety, vol. 29, pp. 385-396, 2006.
- [6] K.E. Lasser, P.D. Allen, S.J. Woolhandler, D.U. Himmelstein,