2011 International Conference on Information and Intelligent Computing
IPCSIT vol.18 (2011) © (2011) IACSIT Press, Singapore

N-Unexpected Temporal Association Rule for Diagnosing Adverse Drug Reaction from Health Database

K.Shanmugapriya[1], D.Shanmugapriya[2], H.Summia Parveen[3] and V.Niranjani[4]

[1,2,3]Department of Information Technology, Sri Eshwar College of Engineering
[4]Department of Computer Science & Engg, Sri Eshwar College of Engineering
[1]shanmugapriya.kumaresan@gmail.com
[2]shanmugapriyait2006@gmail.com
[3]summiaparveen@yahoo.in
[4]niranjani0408@gmail.com

Abstract. In medical field, drugs have led to major improvement in treatment and control of diseases. But they continue to produce adverse effects on human body. Drugs interact negatively with other drugs and produce distressful effects. Adverse drug reaction is the major cause for death and hospitalization all over the world. As adverse effects are unexpected, pre-market drug testing is not efficient to detect adverse effects of drug interactions. The traditional post-market drug testing relies on spontaneous case reports of the affected individuals, which suffer from delay in reporting. Clinical development programs cannot identify all the potential adverse effects of drugs. So sponsors and regulatory organization conduct post marketing surveillance program to identify new adverse effects and evaluate them for their potential impact on the public health. Current post-market adverse effect detecting techniques signal only pairwise unexpected temporal association rule. This paper focuses on developing data mining techniques for detecting adverse drug reactions related with drug-drug interactions. Occurrence of adverse drug reaction is unexpected and infrequent in health domain. To address this feature n-wise unexpected temporal association rule is provided.

Keywords: Adverse Drug Reaction (ADR), temporal association rule, unexpected pattern, health domain, data mining.

1. Introduction

Drugs are never completely safe. In addition to their desired effects, they may also cause side effects. Those side effects that are unexpected and harmful are referred to as adverse drug reaction [1].WHO defines ADR as "a response to a medicine which is noxious and unintended and which occurs at doses normally used in man for prophylaxis, diagnosis or therapy of diseases or for modification of physiological function [2]. The American Food and drug administration defines a series of adverse events as one when the patient outcome is death, life threatening, hospitalization, damage or disruption in the body function [2].

Study of ADR is concern of the field, known as pharmacovigilance[2]. ADR related hospitalization and death is increasing every day [3]. Whenever two or more drugs are being taken, there is a chance that there will be an interaction among the drugs. The interaction may increase or decrease the effectiveness of the drugs or the side effects of the drugs. Therefore, people who take several drugs are at the greatest risk for interactions.

This paper deals with detecting ADR related with drug to drug interactions. The n-wise Unexpected Temporal Association Rule is used to detect combination of drugs that should not be consumed at a time by the individual patient.

2. Related Work

This paper is concerned with the problem of mining drug to drug interactions that may cause adverse effects. In [4], Spontaneous case reports were used to mine the ADR. In [5], [6] UTAR was used to mine drug specific ADR signals from administrative health databases. [6], [7] explored pairwise UTAR.

It has been noted in [8] that objective measures are not sufficient for determining the interestingness of the discovered patterns. Subjective measures are needed. Two main subjective measures are unexpectedness and actionability. The post market ADR detection technique like Bayesian data mining technique [10] perform on spontaneous case reports which describes the suspected causality between drugs and adverse effect for one patient. Spontaneous case reports suffered from underreporting and delay. This work aims to highlight the problem of mining n-wise UTAR where antecedents and consequents occur together unexpectedly.

3. Background Study

3.1. Association Rule

Let I={*x*1, *x*2, *x*3….xm} be a set of items. Let the set X be a subset of I. k=|X| is called k-item set. Let database D be a set of transactions where each transaction T is a set of items such that T is a subset of I. A transaction T is said to support X if and only if X is a subset of T.

Association rule is an implication of the form X=>Y, meaning that the presence of the set X implies the presence of another set Y, where X is a subset of I, Y is a subset of I and X. This rule X=>Y holds in the transaction set D with confidence of c if c% transaction in D that contain X also contain Y. This rule X=>Y has support in the transaction set D if s% of transaction in D contains X.

3.2. Unexpected Pattern

Patterns are unexpected if they are previously unknown to the user [8]. Unexpected patterns are, by definition, interesting. Unexpected patterns include:

1. Unexpected consequent
2. Unexpected reason
3. Totally unexpected pattern

3.3. Temporal association rule

Temporal association rule [6], [11] is denoted as A→C with respect to T, where A is the antecedent, C is the consequent and T is the time window of length T. A and C are constrained within T.

3.4. Unexpected temporal association rule

UTAR is denoted by A→C with respect to T, where the consequent C occurs unexpectedly within a T size period after the antecedent A. The support of the UTAR, supp (A→C with respect to T), is the proportion of T-constrained subsequences that unexpectedly contain A followed by C among all of the T-constrained subsequences. Its confidence is given by conf (A→C with respect to T) = supp (A→C with respect to T)/supp (A→with respect to T), where supp (A→with respect to T) is the proportion of T constrained subsequences that unexpectedly contain A.

3.5. Drug Interactions

Drug interactions may make the drug less effective, cause unexpected side effects, or increase the action of a particular drug. Some drug interactions can even be harmful causing death or disability of the organs of human body. Drug interactions fall in 3 broad categories.

1. Drug-drug interaction
2. Drug-food interaction
3. Drug-condition interaction

4. Proposed System

4.1. Algorithm

1. Initialize Antecedent list A_i (i=1,2,…,n) and Consequent list C(d1,d2…dm),time periods study period T_s,

control period T_c, hazard period T_h ,output setA ,setC, number k.

2. Initialize each A along with its ingredient as

$$\sum_{A=0}^{n} \left(\sum_{l=0}^{m} \right)$$

3. For each Ai generate the ingredient list I_i (i=1, 2,…,n and side effect list C_j(j=1,2,..,m).
4. Calculate leverage l as
   I(A(I1,I2,...,In))=supp(A(I1,I2,...,In)→C(d1,d2,..,d m)) with respect to T,where supp denotes minimum support.
5. Rank the leverage and store top k consequents in setC.
6. For each Consequent listed in set C find the case sequence and non case sequence.
7. Calculate the residual leverage rl as
rl(A(I1,I2,…,In)→C(d1,d2,...,dm))= supp((A(I1,I2,...,In)→C(d1,d2,...,dm))-supp (A(I1,I2,...,In))*supp(→C(d1,d2,…,dm)) with respect to T.
8. Rank top k rl and store in seta.
9. Compare setC with seta.
10. If A (I1, I2,…,In) in set C is also present in setA then output the top k ingredients.
11. Compare the ranked ingredients with patient data and then rank.
12. Output the top k adverse drug combinations.

4.2. System architecture

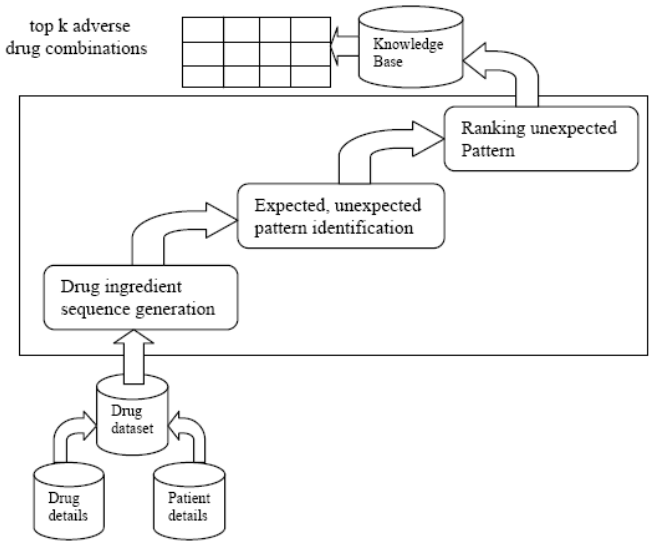

Figure 1. Adverse Drug Combination Detecting System

Fig 1. shows the architecture of the system that is used to detect adverse drug combination. The Drug sequence and ingredient combination is generated from the drug dataset. From those combinations expected and unexpected combinations are identified according to their minimum support values. Unexpected events are then ranked. Adverse effect causing ingredients containing drugs are selected .The drugs containing those ingredients are listed as the possible combination of drugs that might cause adverse effect.

4.3. Comparison and ranking

Comparison is made with the sets A and C. The drug (ingredients) that fall in both the sets are listed and ranked. This signal can help the medical practitioners to avoid prescribing that drug combination to patients.

For example consider the output of the first module as the side effects s1, s2, s3, s4, s5 in a ranked order .This output is fed as input for the second module. In the second module for each side effect some case based exclusion are made. The residual leverage is calculated by finding the support values with respect to side effect and the drug. The output produced by this module will be the top k drug combinations associated with the given side effects. The output of the second and first module are compared and ranked. These ranked adverse drug combinations undergo further ranking with comparison to patient record.

5. Conclusion and Future Work

ADR signals associated with drug interactions can be used to prevent the harm to patients at an early stage. In order to detect the unexpected drug combination n-wise UTAR is provided. The interesting measure "unexpectedness" is calculated based on the given algorithm. The medical authorities can use this solution to detect the combination of drugs that should not be prescribed together at a time for a patient .This solution may reduce the death and hospitalization that are caused due to ADR associated with drug to drug interactions .This paper describes only drug to drug interactions. This work can be extended to include multievent syndromes and drug-food interactions.

6. References

[1] D.W. Bates et al., "Incidence of Adverse Drug Events and Potential Adverse Drug Events: Implications for Prevention," The J. Am. Medical Assoc., vol. 274, no. 1, pp. 29-34, July 1995.

[2] The ICH Expert Working Group, "Post-Approval Safety DataManagement: Definitions and Standards for Expedited Reporting,"ICH Harmonised Tripartite Guideline, http://www.fda.gov/cber/gdlns/ichexrep.htm, Nov. 2003.

[3] C.L. Burgess, C.D. Holman, and A.G. Satti, "Adverse Drug Reactions in Older Australians, 1981-2002," The Medical J. Australia, vol. 182, no. 6, pp. 267-270, Mar. 2005.

[4] M. Hauben and X. Zhou, "Quantitative Methods in Pharmacovigilance: Focus on Signal Detection," Drug Safety, vol. 26, no. 3, pp. 159-186, 2003.

[5] C.-H. Lee, M.-S. Chen and C.-R. Lin, "Progressive Partition Miner: An Efficient Algorithm for Mining General Temporal Association Rules," IEEE Trans. Knowledge and Data Eng., vol. 15, no. 4, pp. 1004-1017, July/Aug. 2003.

[6] Huidong (Warren) Jin, Member, IEEE, Jie Chen, Member, IEEE, Hongxing He,Chris Kelman, Member, IEEE, Damien McAullay, and Christine M. O'Keefe "Signaling Potential Adverse Drug Reactions from Administrative Health Databases" IEEE Transactions on knowledge and data engineering, vol. 22, no. 6,Jjune 2010.

[7] H. Jin, J. Chen, H. He, G.J. Williams, C. Kelman, and C.M. O'Keefe, "Mining Unexpected Temporal Associations: Applications in Detecting Adverse Drug Reactions," IEEE Trans. Information Technology in Biomedicine, vol. 12, no. 4, pp. 488-500, July 2008.

[8] B. Liu, W. Hsu, L.-F. Mun, and H.-Y. Lee, "Finding Interesting Patterns Using User Expectations," IEEE Trans. Knowledge and Data Eng., vol. 11, no. 6, pp. 817-832, Nov. 1999.

[9] V. Curcin, M. Ghanem, M. Molokhia, Y. Guo, and J.Darlington, "Mining Adverse Drug Reactions with E-science Workflows," Proc. IEEE Cairo Int'l Biomedical Eng. Conf. (CIBEC '08), pp. 1-5, Dec. 2008.

[10] D.M. Fram, J.S. Almenoff, and W. DuMouchel, "Empirical Bayesian Data Mining for Discovering Patterns in Post-Marketing Drug Safety," Proc. ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining (KDD '03), pp. 359-368, 2003.