1. Ayman M Mansour[1], 2. Fayez Khazalah[2], 3. Mohammad A Obeidat[1]

Tafila Technical University (1), Al al-Bayt University (2), Jordan
ORCID. 1. 0000-0001-7086-1613, 2. 0000-0002-7448-156X, 3. 0000-0002-0288-962X

doi:10.15199/48.2022.01.05

Suspected Adverse Drug Reaction Detection using Association Rules Mining and Fuzzy sets

***Abstract**. Finding adverse drug reaction ADRs is vital to sustaining human life. The effect of the drug under several factors such as age, drug quantity, laboratory results, sex and drug duration are necessary to improve the quality of human body treatment has not been used previously. The paper uses real databases collected from US hospitals to validate the developed detection system. This paper presents an intelligent system based on association rule mining rules and fuzzy set theory. The developed system has the potential to determine the relationships between a drug and its adverse reactions. This is done by extracting several rules with high support and confidence. Two physicians review the results of the proposed system to validate the results. The results matches the ADRs defined by medical associations and drug companies.*

Streszczenie.** Znalezienie niepożądanych reakcji na leki ADR ma kluczowe znaczenie dla podtrzymania życia ludzkiego. Wpływ leku na kilka czynników, takich jak wiek, ilość leku, wyniki laboratoryjne, płeć i czas trwania leku są niezbędne do poprawy jakości leczenia ludzkiego ciała, nie były wcześniej stosowane. W artykule wykorzystano rzeczywiste bazy danych zebrane ze szpitali w USA do walidacji opracowanego systemu wykrywania. W artykule przedstawiono inteligentny system oparty na regułach eksploracji reguł asocjacyjnych i teorii zbiorów rozmytych. Opracowany system ma potencjał do określenia zależności między lekiem a jego działaniami niepożądanymi. Odbywa się to poprzez wyodrębnienie kilku reguł z dużym wsparciem i pewnością. Dwóch lekarzy dokonuje przeglądu wyników proponowanego systemu, aby je zweryfikować. Wyniki są zgodne z ADR-ami zdefiniowanymi przez stowarzyszenia medyczne i firmy farmaceutyczne.* (Wykrywanie podejrzeń niepożądanych reakcji na lek za pomocą eksploracji reguł asocjacyjnych i zbiorów rozmytych***)

Keywords: Aassociation rules, Fuzzy sets, Adverse drug reactions, Medical cues.
Słowa kluczowe: niepożadane reakcje na leki, reguły asocjacji, zbiory rozmyte.

Introduction

ADR was defined in [1] as “An appreciably harmful or unpleasant reaction, resulting from an intervention related to the use of a medicinal product; adverse effects usually predict hazard from future administration and warrant prevention, or specific treatment, or alteration of the dosage regimen, or withdrawal of the product”. ADRs are a major public health problem in the United States. Before drugs are marketed, they are extensively tested in animals and in clinical trials in humans. Clinical trials often refer to pre-marketing studies. Clinical trials have been playing a crucial role in evaluating the overall safety and efficacy of new medications before they get into the market. However, due to many reasons [1] the clinical trials are limited in size and duration, and thus are not capable of detecting rare ADRs. Given the limited information available when the drug is marketed, post-marketing surveillance has become increasingly important. Post-marketing surveillance is the process of identifying, reporting, and responding to the issues occurred while taking medication. The post marketing reporting systems is considered the first line reference for the detection of any new and unexpected drug related adverse effects.

Some serious adverse drug reactions are detected after a drug has been on the market for a while. Therefore, when a drug is approved and began to be available in market, huge numbers of patients will be affected by that potential adverse effect until that potential adverse event has been identified. Until now there is no clearly defined process to be followed in order to detect adverse events [1].

ADRs reporting system is voluntary in Jordan. The reporting is done through spontaneous reporting of ADRs. The rate of reporting is very low because of the lack of awareness. All ADRs are reported to the Jordan Pharmacovigilance center (JPVC) at Jordan Food and Drug Administration (JFDA). The following summarize the procedure spontaneous reporting system in Jordan:

A healthcare professional or marketing authorization holder reports a suspected adverse drug reaction related to one or more pharmaceutical products, to The Jordan pharmacovigilance center (JPC). Reports are made in writing (e.g. using report forms), electronically, or by any other approved way. Then, Reports are collected, collated, and validated by the pharmacovigilance center and are usually entered into a database. Serious reactions are handled with the highest priority. The database is used to identify potential signals and analyze data in order to clarify risk factors, apparent changes in reporting profiles etc. Therefore, effective pharmacovigilance systems are needed in order to reduce unnecessary suffering by patients' financial loss of unsafe use of medicines. It is essential that the drug safety monitoring system is supported by all available databases and diagnostic databases.

Data mining algorithms are being used to explore spontaneous reporting databases for adverse reaction signal pairs. To extract association rules in datasets with unbalanced class distribution, [2] proposed an algorithm that uses a set of new interestingness criteria. It then applied it on a real-world health dataset with unbalanced classes to find whether the angioedema (which is an adverse drug reaction) is related to the use of ACE inhibitors. The results showed that the algorithm was able to identify the groups that are most likely to have angioedema due to ACE inhibitors. According to [2], the conventional association rule mining algorithms suffer from several issues. One of these issues is that they generate too many rules that limited memory cannot handle even when using some rule pruning techniques. Another issue is that when the dataset has unbalanced classes, these conventional algorithms are unable to find many of the interesting rules even though the minimum support is set to a lower value. To overcome these issues, [2] proposes two new interestingness criteria (local support and exclusiveness) that are suitable for finding interesting rules in datasets with highly unbalanced classes.

As an extension to the work done in [2], the same authors in [3] used local support and risk ratio as criteria to find interestingness rules. The main contribution in [3] is representing rules by two kinds of probability trees that aim to guide the drug prescribers to the risk of specific adverse drug reactions for some categories of patients. Each rule is represented as a probability tree, where at each node; the population support and the risk ratio of an individual component are shown. The root node represent the entire population and the right node represents the risk ratio for the complementary population (i.e., the population where the condition of the left node does not apply). Thus, by further breaking down the risk caused by individual component factors, the medical practitioners can have sufficient knowledge about the side effects of some drugs on specific categories of patients.

Spontaneous Reporting System (SRS) databases contain a collection of reports provided voluntary from health-care professionals and patients or mandatory from drug manufacturers. SRSs are usually used by pharmacovigilance to detect possible adverse drug reactions. Most of the previous research in the area of pharmacovigilance focused on a single drug-adverse association analysis. In [4], [5], and [6] multi-item drug-adverse association analysis are considered. All of them tailored the Apriori association rule mining algorithm for extracting multi-variate association rules between drugs and adverse events from the SRSs. Instead of using the confidence measure as in the classical Apriori algorithm, [4] used the Relative Reporting Ratio (RR) measure for the judgement of a rule strength and interestingness. In contrary, [5] and [6] used the Proportional Reporting Ratio (PRR), which is a disproportionality measure.

[7] Deployed a hybrid scheme to detect potential ADRs from SRSs. The proposed scheme is composed of three stages. In the first stage, each report that has more than one drug or more than one adverse event in it is broken into small tuples, such that each new tuple has only an association between one drug and one adverse event. For example, a report that has mentions of three drugs and four adverse events is broken into twelve tuples. In the second stage, the ARs are mined using the classical Apriori algorithm from the resulted tuples of the first stage. In the last stage, the mostly known word embedding vector model, Word2Vec, is applied on the mined ARs to detect the potential ADRs. Word2Vec is mostly used in Natural Language Processing (NLP) and text mining to find the most similar words. It can learn the semantically related words when applied on a large corpus of text. [7] Used this word-embedding model to improve the performance of detecting the ARs signals by finding the most related drugs and adverse events. The proposed scheme is evaluated by conducting experiments on two kinds of drugs and then comparing the detected ARs signals with the existing pharmacovigilance literature.

Traditional association rule mining algorithms may not be efficient when applied on some transactional databases, for example, publication-like ones [8]. In publication-like databases, each transaction consists of a set of items, where each item has a specific exhibition period. It is not fair in this kind of databases to consider only the occurrence of an item with respect to the size of database when computing the support and confidence metrics. Instead, [8] introduces temporal association rule mining algorithm that incorporate the exhibition periods of items when computing the support and confidence metrics. [9] Utilizes the algorithm proposed in [8] to mine the Unexpected Temporal Association Rules (UTARs) to detect the unexpected drug-drug adverse effects.

[10] Conducted a comparative study to evaluate the performance of several algorithms when used for detecting the signals of adverse drug events in SRSs. The main goal of the study is to see whether the ARM algorithm is superior to those other algorithms. The study compared the performance of the following algorithms: the Proportional Reporting Ratio (PRR), the adjusted PRR (adjPRR), the Reporting Odds Ratio (ROR), the Bayesian Confidence Propagation Neural Network (BCPNN), and the ARM algorithm. The performance study is conducted on both simulated and real datasets. It showed that the ARM algorithm is superior to the other algorithms in detecting the signals of ADRs with similar performance results.

[11] Proposed a method to refine ADR signals that may occur due to confounding. In other words, the proposed method removes out all signals that are not true adverse reactions and that can be explained due to other previous health issues. It works only when an electronic health records (EHRs) databases are available for patients with a few years of detailed medical records. In this study, ARMA is applied on the EHRs database to mine for common drug-adverse patterns, and the extracted rules are used to refine side effect signals. Thus, by refining some of the ADR signals, we reduce the time needed to evaluate the suspected signals to see if they are really related to true ADRs.

Similarly, to detect only true casual associations of drug-drug interactions from SRSs, [12] proposed an algorithm based on Casual Bayesian Network (CBN) called Causal Association Rule Discovery (CARD). The idea behind the work done in [12] is based on the well-known phrase “correlation does not imply causation”. In other words, if two drugs are associated together with a specific adverse event, this does not mean that the correlation between the two drugs is the cause of the adverse event, because the event might be due to another unknown or confounding variable, like a previous disease, for example. Thus, the proposed CARD algorithm in [12] augments association rule mining with some properties of CBN to guide the mining for causal association rules instead of regular association rules. It then compared the results of the proposed method with the baseline ARM algorithm. The results showed that the proposed method is able to find more expected drug-drug interactions but less unexpected drug-drug interactions.

Health social forums, like Daily Strength, contains valuable information about the experience of patients with some drugs. Through their comments in such forums, patients usually mention the effects they experience while taking specific drugs. In [13], the authors propose a new text mining method to extract ADRs automatically from health social networks using association rule mining (ARM). As most of the comments in such forums are informal texts, the focus of [13] is to discover ADRs mentions from colloquial text. Thus, the proposed method uses both Natural Language Processing techniques and ARM algorithm to extract the potential colloquial expression patterns about the ADRs. The discovered patterns can be applied to new comments to discover mentions of ADRs. However, as a pre-condition, experts annotate the user comments by specify the mentions of ADRs for each comment.

Similar to [13], ADRs are extracted from user comments in Spanish health-related forums in [14]. In addition, [14] also extracts the drug indications in addition to the ADRs from comments. It also proposes an automatic construction of the first Spanish database (SpanishDrugEffectBD) for drug indications and reactions.

[15] applied ARM node in SAS Enterprise Miner on the Pediatric Health Information System (PHIS) database to analyze usually combined drugs (i.e., polypharmacy) and then visualizes then these drug-combinations using Link Analysis node of SAS Enterprise Miner. The discoveries out of this analysis can be used for many purposes, for

examples, to detect the characteristics of in and out pediatric patients who have specific polypharmacy patterns.

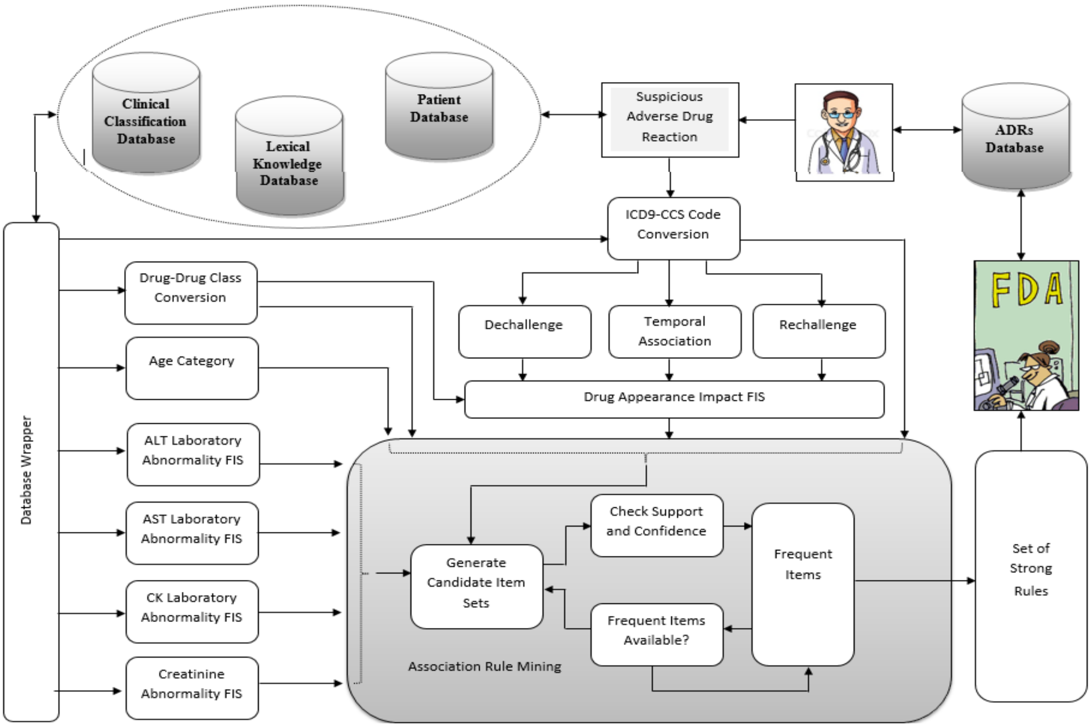

Clinical Classification Database
Lexical Knowledge Database
Patient Database
Suspicious Adverse Drug Reaction
ADRs Database
FDA
ICD9-CCS Code Conversion
Database Wrapper
Drug-Drug Class Conversion
Age Category
Dechallenge
Temporal Association
Rechallenge
Drug Appearance Impact FIS
ALT Laboratory Abnormality FIS
AST Laboratory Abnormality FIS
CK Laboratory Abnormality FIS
Creatinine Abnormality FIS
Check Support and Confidence
Generate Candidate Item Sets
Frequent Items
Frequent Items Available?
Association Rule Mining
Set of Strong Rules

Fig. 1. The develop detection system

The Developed System

The developed ADR signal pairs detection methodology is based on many cues: Drug Appearance Impact, Abnormality in AST, ALT, CK and Creatinine laboratory tests, Age Category and Drug classes. The cues represent the higher-level information that are obtained from the patients' elementary and diagnostic data. The developed system is shown in Figure 1.

The developed system relies on association rule mining that finds the relationships between a set of inputs. The set of inputs is used to generate candidate items sets. The generated item sets are analyzed and checked for high support and high confidence to achieve a set of strong rules. The rules show the correlation of a suspected adverse drug and a taken medication. This with decision-making. The initial diagnostic factors are extracted from different databases. Then these factors are processed using fuzzy set theory to generate secondary factors that are more accurate than the initial factors in order to obtain more effective decision-making rules.

The Adverse drug reaction appears normally when a patient takes a medication to treat a specific disease. For example, if the Panadol is taken as an analgesic to treat headaches, then it turns out after a period of time that the patient suffers from a stomach ulcer. It is highly important to know whether the suspected relationship between the Panadol and stomach ulcer is true or not. The developed system detects such relationships between an adverse effect and a drug under different factors that the patient was exposed to during a treatment period.

Databases contains a Patient Database, Lexical Knowledge Database, a Clinical Classification Database and ADRs Database. This Lexical knowledge base consists of linguistic knowledge, such as synonyms of medical words, grammatical patterns in which they can appear, possible abbreviation of the medical words and complex medical terminology. The Clinical Classification Database offers a wrapper that maps the human description of a case to different medical codes (ICD-9, Clinical Classification, etc.) that are commonly used in U.S. Such codes describe the medical conditions of patients and symptoms phenomenon.

The patient database contains the information related to the patients such as the taken medications, gender, age, laboratory tests results, symptoms and the procedures

Followed by physician in order to diagnose a case. The ADRs database contains the adverse drug reactions corresponds to a drug. The system are supported with ability of mapping free text terms to unique concepts. It uses the Lexical Variants Generator (LVG) program provided by the National Library of Medicine. LVG is the most powerful solution for lexical variations at the individual word level. This allows the system to deal with inflectional variants, spelling variants, acronyms and abbreviations, expansions, derivational variants, synonyms as well as combinations of these. For example the system deals with the following spelling variant words: lab, laboratory, labs, and laboratory (with spilling mistake) as laboratory term. Table 1 shows the used ADR parameters. The Database Wrapper provides a JDBC (Java Database Connectivity) wrapper library, a simple abstraction layer that encapsulates standard database language (e.g., SQL) and provides frequently-used methods for database connectivity, working with database schema information and database data.

ICD-9 to CCS code Conversion

In order to represent the morbidities, the *International Classification of Diseases, 9th Revision, and Clinical Modification* (ICD-9-CM) code is used. The ICD-9 code provides codes to classify diseases and a wide variety of signs, symptoms, abnormal findings, complaints, social circumstances, and external causes of injury or disease. Every health condition is assigned a unique category and given a unique code. For example, if a patient is diagnosed with Hepatitis C, he/she will be given the ICD-9 code “070.51“. If the diagnosis is for something acute, something that goes away with treatment like a rash or the flu, then the ICD-9 code will be less important because the illness or condition will go away. However, if the patient is diagnosed with a chronic or lifelong problem, like heart disease or diabetes, the ICD-9 code will be more important and will affect his future medical care.

Table 1. The ADR parameters

<table>
<tr><td colspan="6">Llaboratory Tests Abnormality Levels</td><td>Symptoms and Morbidities</td></tr>
<tr><td>AST</td><td>ALT</td><td>CK</td><td colspan="2">K</td><td>CR</td><td rowspan="6">CCS Codes</td></tr>
<tr><td>VeryLow</td><td>VeryLow</td><td>VeryLow</td><td colspan="2">Low</td><td>Low</td></tr>
<tr><td>Low</td><td>Low</td><td>Low</td><td colspan="2">Medium</td><td>High</td></tr>
<tr><td>Medium</td><td>Medium</td><td>High</td><td colspan="2">High</td><td></td></tr>
<tr><td>High</td><td>High</td><td>Very High</td><td colspan="2"></td><td></td></tr>
<tr><td>Very High</td><td>Very High</td><td></td><td colspan="2"></td><td></td></tr>
<tr><td colspan="2">Patient Info</td><td colspan="5">Drug Info</td></tr>
<tr><td>Age</td><td>Sex</td><td colspan="2">Drug Association Impact</td><td colspan="2">Drug Quantity</td><td rowspan="5">Drug Categories</td></tr>
<tr><td>young</td><td>Male</td><td colspan="2">Unlikely</td><td colspan="2">Low</td></tr>
<tr><td>middle age</td><td>Female</td><td colspan="2">Possible</td><td colspan="2">Medium</td></tr>
<tr><td>old age</td><td></td><td colspan="2">Likely</td><td colspan="2">High</td></tr>
<tr><td>elderly</td><td></td><td colspan="2"></td><td colspan="2"></td></tr>
</table>

Since different ICD-9 codes may represent the same (or similar) diagnoses, ICD-9 codes are clustered into a manageable number of categories based on the clinical classifications system (CCS) for the ICD-9-CM fact sheet Developed at the Agency for Healthcare Research and Quality, the CCS groups over 13,600 ICD-9 codes into 285 mutually exclusive and clinically meaningful categories. The clinical classifications system makes it easy for physicians and hence the developed detection model to understand patient cases and analyze them for similarity task.

Drug Association Impact

Drug Association Impact is based on three important cues, which are Temporal Association, Dechallenge and Rechallenge.

Temporal association gives the relationship between the time of taking a drug and the time of symptom occurrence. This time duration is called Symptom Appearance Duration. It should be noted that in the case of a potential ADR, exposure to a drug should always precede the effect (symptom). This distinction is important because the effect might result from entirely different causes (e.g., underlying diseases or reception of another medication).

Figure 2 shows an Example of possible ADR event. In this case, since suspected ADR occurs after the start date of medication, it can be considered as an adverse event caused by that drug Based on the experience of the physicians on the team, we define the following fuzzy rules to link cause (drug) to effect (ADR):

If Symptom Appearance Duration is Short Then Temporal Association is Likely.

If Symptom Appearance Duration is Medium Then Temporal Association is Possible.

If Symptom Appearance Duration is Long Then Temporal Association is Unlikely.

Both Symptom Appearance Duration and Temporal Association are variables characterized by triangular fuzzy sets. The fuzzy parameters and fuzzy sets were optimized using genetic algorithm as in [16], [17].

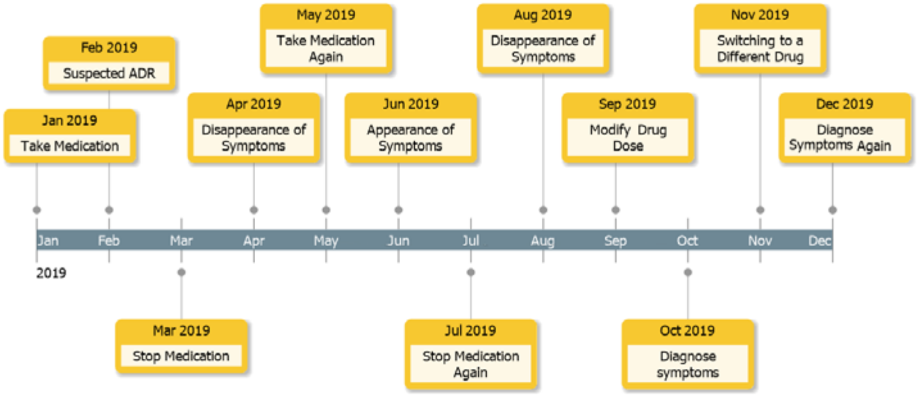


Fig. 2. Example of possible ADR event

Medication Dechallenge refers to the relationship between discontinuity of the drug and abatement of the apparent ADR. Dechallenge is a fuzzy variable characterized by triangular fuzzy sets labeled as *“Unlikely*,” ”*Possible*,” and *“Likely*”. It is difficult to directly evaluate Dechallenge of a pair since the drug stop date is usually unavailable in electronic health databases. However, it can be indirectly assess the existence of Dechallenge of a pair if a symptom occurs after the drug start date and another drug in the same class was prescribed after the appearance of the symptom. This is because the physicians often stop a drug and prescribe another drug in the same class to avoid apparent adverse effect found on a patient.

Also, if the temporal association is *Unlikely*, then Dechallenge is *Unlikely*. In some cases the patient stops taking the drug for a period greater than 120 days then the stop date can be considered as the previous start date plus the number of days the patient took that medication. In such cases five fuzzy rules will be applied to get the strength of Dechallenge. Here are the rules:

- If Time Duration between stopping the drug and the abatement of the apparent symptoms is *small,* then Dechallenge is *Likely*.
- If the Time Duration between stopping the drug and the abatement of the symptoms is Medium then Dechallenge is Possible.
- If the Time Duration between stopping the drug and the abatement of the symptoms is Large then Dechallenge is Unlikely.
- If the reaction does not abate after withdrawal of drug then Dechallenge is Unlikely.
- If the reactions occurred again after the drug was discontinued then Dechallenge is Unlikely.

Time Duration between stopping the drug and the abatement of the symptoms is a fuzzy variable represented by triangular membership functions.

Medication Rechallenge depicts the relationship between re-introduction of the drug discontinued before and recurrence of an ADR. Rechallenge is determined by the temporal associations of the two consecutive occurrences of the same pair one after taking the medication and the other one after the reintroduction of the medication. Let Temporal Association of time t_1 and Temporal Association of time t_2 represent the two temporal associations, respectively. Then the following fuzzy rules are used to assess the value of the Rechallenge of a pair.

- If Temporal Association of time t_1 is *Likely* and Temporal Association of time t_2 is *Likely* Then Rechallenge is *Likely*.
- If Temporal Association of time t_1 1 is *Likely* and Temporal Association of time t_2 is *Possible* Then Rechallenge is *Likely*.
- If Temporal Association of time t_1 is *Likely* and Temporal Association of time t_2 is *Unlikely* Then Rechallenge is *Possible*.
- If Temporal Association of time t_1 is *Possible* and Temporal Association of time t_2 is *Likely* Then Rechallenge is *Likely*.
- If Temporal Association of time t_1 is *Possible* and Temporal Association of time t_2 is *Possible* Then Rechallenge is *Possible*.
- If Temporal Association of time t_1 is *Possible* and Temporal Association of time t_2 is *Unlikely* Then Rechallenge is *Possible*.
- If Temporal Association of time t_1 is *Unlikely* and Temporal Association of time t_2 is *Likely* Then Rechallenge is *possible*.
- If Temporal Association of time t_1 is *Unlikely* and Temporal Association of time t_2 is *Possible* Then Rechallenge is *Possible*.
- If Temporal Association of time t_1 is *Unlikely* and Temporal Association of time t_2 is *Unlikely* Then Rechallenge is *Unlikely*.

Both Temporal Association and Rechallenge are fuzzy variables. Rechallenge is fuzzified by three fuzzy sets Likely, Possible and Unlikely. Table 2 shows the variables of Drug Appearance Impact and their fuzzy set parameters. All are triangular membership functions.

The Drug Appearance Impact is calculated as a linear combination of Dechallenge, Temporal Association and Rechallenge. The aggregated Drug Appearance Impact value is computed as the following:

Drug Appearance Impact value = Dechallenge x w_1 + Temporal Association x w_2 + Rechallenge x w_3

where $w_1 + w_2 + w_3 = 1$ and the weights control the importance of the sub values.

Table 2. Variables of Drug Appearance Impact and their fuzzy set parameters

Variable	**Fuzzy Set Name**	**Triangular Fuzzy Set Parameters**		
		a	b	c
Symptom Appearance	Short		0	20
	Medium	0	20	40
	High	20	40	
Temporal Association	Unlikely		0	0.65
	Possible	0.25	0.65	1
	Likely	0.65	1	
Abatement Duration	Small		0	10
	Medium	0	10	25
	High	10	25	
Dechallenge	Unlikely		0	0.65
	Possible	0.25	0.65	1
	Likely	0.65	1	
Rechallenge	Unlikely		0	0.65
	Possible	0	0.65	1
	Likely	0.65	1	

Laboratory Tests Abnormalities

This section shows how to determine the effects of Laboratory Tests and its abnormalities, which include Creatine phosphokinase (CPK) Laboratory Test (also known as Creatine Kinase (CK)), Transaminases Laboratory Test (either ALT or AST), Creatinine Laboratory Test and Potassium Laboratory Test. These laboratory tests are mainly used in ADRs detection studies [16], [17]. Most people with adverse event in the early stages feel well and have no clear symptoms that would lead a health care provider. This place a large emphasis on laboratory tests to diagnose, predict or evaluate a medical problem since they are indicative of extensive problems in the patient. The liver for example is one of the organs that can be affected from medications. The liver has several functions and it is usually called the body's manufacturing and filtering unit. The Alanine Aminotransferase (ALT) and Aspartate Aminotransferase (AST) laboratory test are typically used to evaluate liver functions or liver injury. Elevation of these tests is reflects a damage to the liver cell. Another example is the CK test. The elevation of CK laboratory result rise when muscle or heart cells are injured. Abnormality of a laboratory test shows the degree of elevation of a laboratory test result.

For each laboratory test, the laboratory result will be converted to its abnormality value. The interpretation of abnormality value of laboratory test results is very important to understand the situation of the patients. The abnormality value will be zero for the laboratory results in normal ranges. For other ranges, The Abnormality value will be calculated using fuzzy Inference system. The laboratory results will be the input to the system and the Abnormality value will be the output. Both the input and the output are fuzzy variables. The variables used to determine Abnormality in Laboratory Tests are defined in Table 3.

Age Category

Age is an important factor in determining suspected ADR. For example, acute diarrhea in an adult is not that danger as in an elderly patient which could produce dehydration more quickly; Based on the literature and the physicians on our team, the age of a patient will be classified to one out of four groups. These groups are:

- Group 1: Age is below 35 years.
- Group 2: Age is between 35 and 69 years.
- Group 3: Age is Between 70 and 90 years.
- Group 4: Age is greater than 90 years.

Table 3. Variables of Laboratory Tests Abnormalities and their fuzzy set names, types and parameters

Variable/Fuzzy Set Type	Fuzzy Set Name	Fuzzy Set Parameters		
		a	b	c
AST Test/ Triangular	Very Low		60	100
	Low	60	100	130
	Medium	100	130	160
	High	130	160	200
	Very High	160	200	
AST Abnormality/ Triangular	Very Low		0	0.25
	Low	0	0.25	0.5
	Medium	0.25	0.50	0.75
	High	0.50	0.75	1
	Very High	0.75	1	
ALT Test/ Bell	Very Low	70	16	2
	Low	102.5	16	2
	Medium	135	16	2
	High	167.5	16	2
	Very High	200	16	2
ALT Abnormality/ Bell	Very Low	0	0.15	2
	Low	0.25	0.15	2
	Medium	0.5	0.15	2
	High	0.75	0.15	2
	Very High	1	0.15	2
CK Test/ Bell	Very Low	200	75	2
	Low	350	75	2
	High	500	75	2
	Very High	750	75	2
CK Abnormality/ Bell	Very Low	0	0.2	2
	Low	0.4	0.2	2
	High	0.8	0.2	2
	Very High	1	0.2	2
Potassium Test/ Gaussian	Low	5	0.5	
	Medium	6	0.5	
	High	7	0.5	
Potassium Abnormality/ Gaussian	Low	0	0.3	
	Medium	0.5	0.3	
	High	1	0.3	
Creatinine Test/ Bell	Low	1.5	0.75	3
	High	3	0.75	3
Creatinine Abnormality/ Bell	Low	0	0.5	3
	High	1	0.5	3

Drug to Drug Class Conversion

A drug may be classified by the chemical type of the active ingredient or by the way it is used to treat a particular condition. To achieve this, a strategy of classifying drugs is needed in order to be used in the developed framework. The medications are catalogued according to the Anatomical Therapeutic Chemical classification. This system is recommended by the WHO for drug utilization studies. In the Anatomical Therapeutic Chemical classification system, the active substances of a medication are divided into different groups according to the functional system they have effects on besides the chemical properties of the medication. For example, in the Anatomical Therapeutic Chemical system captopril and enalapril which are inhibitors medications are given the code C09AA.

The Developed Association Rule Detection Algorithm

A strong association rule is the one with a confidence value that satisfies min_conf. The problem of mining association rules, from a database D, is a process that composes two steps [18]:

1. Generate frequent or large sets.
2. Generate strong association rules from frequent sets.

A frequent set is the one that satisfies min_sup. To find all frequent sets in D, Apriroi algorithm scans D many times. In the first scan, candidate 1-sets, denoted as C1, are generated and their support count are computed. Next, the frequent 1-sets, denoted as L1, are generated from C1 after removing the C1's sets with support values less than min_sup. Afterwards, L1 is used to generate candidate 2-sets, denoted as C2. This process continues until no more candidate or frequent sets are left. Therefore, Apriori algorithm requires a full scan on D to find each of the Lk sets. The last frequent k-sets denoted as, Lk, is then used to generate the strong association rules, where k represents the number of full database scans the algorithm needs to extract this final Lk.

To avoid generating and computing supports for too many candidate sets, Apriori algorithm generates candidate sets only from the frequent sets found in the previous pass on D. In addition, it prunes any candidate sets with a subset that is not frequent. This is called the support-based pruning Apriori property, a one that makes it a so popular algorithm.

The second step of the process of mining association rules is to generate the strong association rules out of the frequent or large sets we have got from the first step. As we stated before, an association rule is considered a strong rule if it satisfies both min_sup and min_conf. As we have created the association rule from the frequent sets, then it will automatically satisfy min_sup. To verify whether an association rule satisfies min_conf or not, we find the confidence of the association rule using Equation 1. We generate the strong association rules from the frequent sets L as follows [Han and Kamber, 2006]:

- First, we find all non-empty subsets of l, for each l ϵ L.
- Next, we generate a candidate association rule for each non-empty subset s of l in the format (s ⇒ (l - s)) and compute its confidence according to Equation 1.
- Finally, we only output the association rules, which satisfy min_conf, and these are the strong association rules that we generated from the frequent sets.

We illustrate the process of mining association rules by the following two examples. The first example illustrate how we generate frequent sets from a given database D, and the second example illustrate how to generate the strong association rules from the frequent sets, L, found in the first example.

Example 1: Generate frequent sets in Apriori algorithm

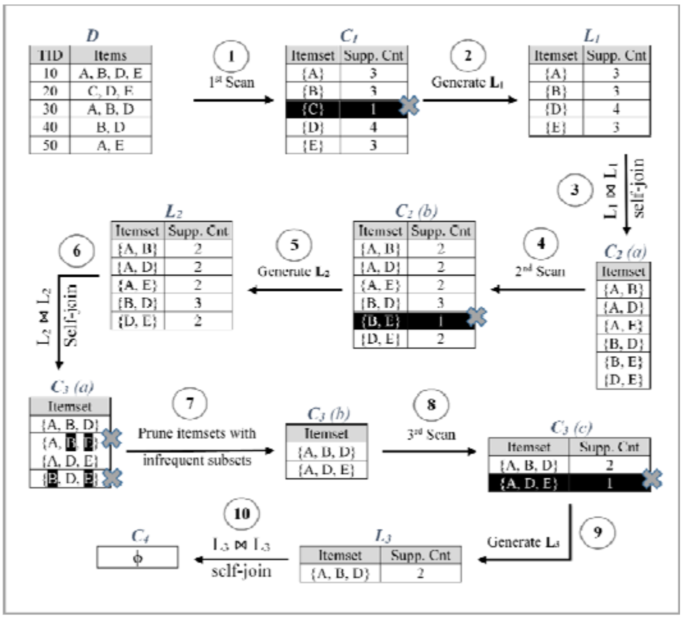
Consider the small database, D, shown in Figure 3 (D). It has five records with the TIDs (10, 20, 30, 40, and 50) and five unique cues (A, B, C, D, and E). A, B, C, D, and E are Drug Appearance Impact, Abnormality in laboratory tests, Age Category, Drug classes and a suspected ADR (symptom). Let us apply the Apriori algorithm shown in Figure 3 on D, as shown in Figure 3, to mine frequent sets and then generate association rules out of them in the second example. Assume that min_sup count is two, which is 50% of the number of total records, and min_conf is 70%. Here, we use the support count instead of percentage as a threshold for sake of simplicity.

Next, we explain the process of generating frequent sets from D.

1. In the first scan on D, we find the candidate 1-sets along with their support count and store them in C1, as shown in Figure 3 (C1).
2. For each 1-set in C1, if the set's support count satisfies min_sup, add it to the frequent set L1. Here, we remove the set {C} as its support count is one, which is less than min_sup. The result is shown in Figure 3 (L1).
3. To find the set of frequent 2-sets, L2, we do a self-join

on L1 (L1 ⋈ L1) to produce the set of candidate 2-sets, C2. The result is shown in Figure 3 (C2, a).

4. Next, we scan D for the second time to compute the support count for each set in C2. The result is shown in Figure 3 (C2, b).
5. L2 is then generated by adding to it only C2's sets that satisfy the minimum support, as shown in Figure 18 (L2). So here, we do not add the 2-set {B, E} to L2, because its support is less than two.
6. Then, we find C3 by doing a self-join on L2, to get the result as shown in Figure 3 (C3, a).
7. Using the Apriori property, we prune any candidate 3-set with a subset of 2-set that is infrequent. Therefore, we remove {A, B, E} and {B, D, E} because they both have the infrequent 2-set subset {B, E}. So the remaining of C3 is shown in (C3, b).
8. Next, we scan D for the third time to compute the support counts for C3's sets, as shown in Figure 3 (C3, c).
9. We then generate L3 by adding to it only C3's sets that satisfy the minimum support, as shown in Figure 3 (L3). So here, we do not add the 3-set {A, D, E} to L3, because its support is less than two.
10. C4 is then generated by doing a self-join in L3. However, because L3 has only one set, C4 becomes empty (ϕ), and the algorithm terminates after finding all frequent sets as shown in the last frequent sets, L3.



D

TID	Items
10	A, B, D, E
20	C, D, E
30	A, B, D
40	B, D
50	A, E

C1

Itemset	Supp. Cnt
{A}	3
{B}	3
{C}	1
{D}	4
{E}	3

L1

Itemset	Supp. Cnt
{A}	3
{B}	3
{D}	4
{E}	3

C2 (a)

Itemset
{A, B}
{A, D}
{A, E}
{B, D}
{B, E}
{D, E}

C2 (b)

Itemset	Supp. Cnt
{A, B}	2
{A, D}	2
{A, E}	2
{B, D}	3
{B, E}	1
{D, E}	2

L2

Itemset	Supp. Cnt
{A, B}	2
{A, D}	2
{A, E}	2
{B, D}	3
{D, E}	2

C3 (a)

Itemset
{A, B, D}
{A, B, E}
{A, D, E}
{B, D, E}

C3 (b)

Itemset
{A, B, D}
{A, D, E}

C3 (c)

Itemset	Supp. Cnt
{A, B, D}	2
{A, D, E}	1

L3

Itemset	Supp. Cnt
{A, B, D}	2

Fig. 3. Generating Frequent Sets in Apriori Algorithm

Example 2: Generate Strong Association Rules from Frequent Sets

We refer to Example 1 and Figure 3, where we have generated the frequent sets, L, which only has one frequent set, *l*, where *l* = {A, B, D}.

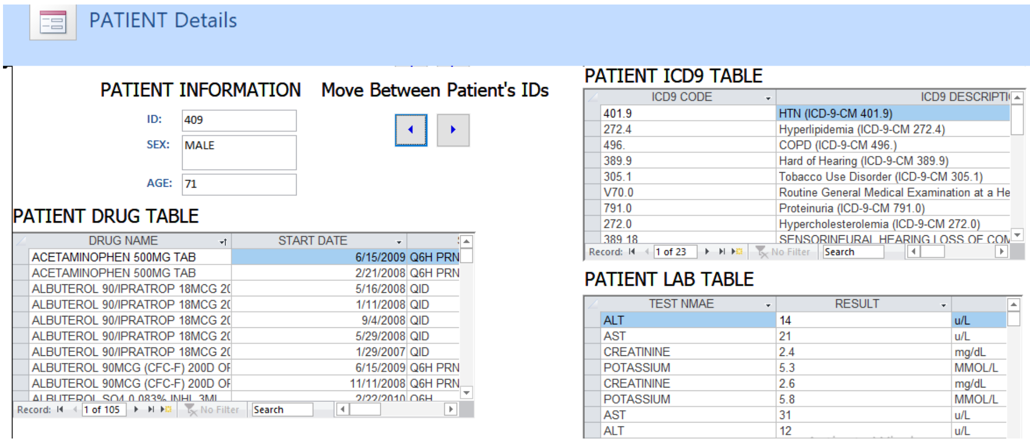
To generate the strong association rules from *l*, first, we find all non-empty subsets of *l*. Next, we find all candidate association rules of *l* and compute their confidences. The ones having confidence ratios equal or above *min_conf* are considered the strong association rules. Thus, as shown in Table 4, the only strong association rules are ({A, B} ⇒ {D}) and ({A, D} ⇒ {B}).

Table 4. Generating strong association rules from frequent sets for example 1

Non-empty subsets of *l* = *{A, B, D}*	Candidate Association Rules	Rule Con-fidence	Rule Con-fidence %	Strong ?
{A}	{A} ⇒ {B, D}	2/3	67%	No
{B}	{B} ⇒ {A, D}	2/3	67%	No
{D}	{D} ⇒ {A, B}	2/4	50%	No
{A, B}	**{A, B} ⇒ {D}**	**2/2**	**100%**	**Yes**
{A, D}	**{A, D} ⇒ {B}**	**2/2**	**100%**	**Yes**
{B, D}	{B, D} ⇒ {A}	2/3	67%	No

Experiments and Results

The purpose of the experiment is to examine the developed algorithm. The used database (Figure 4) is a real database from a Veterans Affairs Medical Center in Detroit during the time period from January 1, 2005 to December 31, 2008. The retrieved patient data includes dispensing of drug, office visits, symptoms experienced, and laboratory testing. For each event certain details were obtained. The data for dispensing of drug includes name of the drug, quantity of the drug dispensed, dose of the drug, drug start date, and the number of refills. The office visits data includes treatment regimens, treatment start dates and stop dates. The symptoms experienced data includes the symptoms appearance date, the symptoms ICD-9 codes and the ICD-9 code description. The system is implemented using Waikato environment for knowledge analysis (Weka) [19] and Fuzzy Jess software packages [20]. Access database 2019 was adopted for the development of the database. The total number of retrieved patients was 20,000 (19,102 males and 898 females). All the data was stored in a Microsoft Access database. Figure



PATIENT INFORMATION

ID: 409
SEX: MALE
AGE: 71

PATIENT DRUG TABLE

DRUG NAME	START DATE	
ACETAMINOPHEN 500MG TAB	6/15/2009	Q6H PRN
ACETAMINOPHEN 500MG TAB	2/21/2008	Q6H PRN
ALBUTEROL 90/IPRATROP 18MCG 2(	5/16/2008	QID
ALBUTEROL 90/IPRATROP 18MCG 2(	1/11/2008	QID
ALBUTEROL 90/IPRATROP 18MCG 2(	9/4/2008	QID
ALBUTEROL 90/IPRATROP 18MCG 2(	5/29/2008	QID
ALBUTEROL 90/IPRATROP 18MCG 2(	1/29/2007	QID
ALBUTEROL 90MCG (CFC-F) 200D OF	6/15/2009	Q6H PRN
ALBUTEROL 90MCG (CFC-F) 200D OF	11/11/2008	Q6H PRN
ALBUTEROL SO4 0.083% INHL 3ML	2/22/2010	Q6H

Record: 1 of 105 No Filter Search

PATIENT ICD9 TABLE

ICD9 CODE	ICD9 DESCRIPTI
401.9	HTN (ICD-9-CM 401.9)
272.4	Hyperlipidemia (ICD-9-CM 272.4)
496.	COPD (ICD-9-CM 496.)
389.9	Hard of Hearing (ICD-9-CM 389.9)
305.1	Tobacco Use Disorder (ICD-9-CM 305.1)
V70.0	Routine General Medical Examination at a He
791.0	Proteinuria (ICD-9-CM 791.0)
272.0	Hypercholesterolemia (ICD-9-CM 272.0)
389.18	SENSORINEURAL HEARING LOSS OF COM

Record: 1 of 23 No Filter Search

PATIENT LAB TABLE

TEST NMAE	RESULT	
ALT	14	u/L
AST	21	u/L
CREATININE	2.4	mg/dL
POTASSIUM	5.3	MMOL/L
CREATININE	2.6	mg/dL
POTASSIUM	5.8	MMOL/L
AST	31	u/L
ALT	12	u/L

Fig. 4. Graphical user Interface of the Database

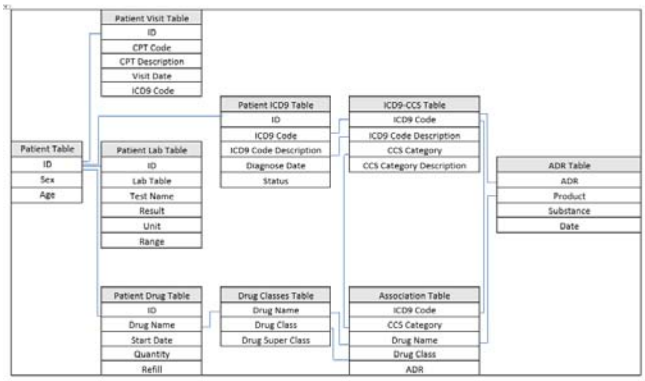


Fig. 5. System Database diagram

An example of the used ADRs database and ICD9-CCS database are shown in Table 5 and Table 6 respectively.

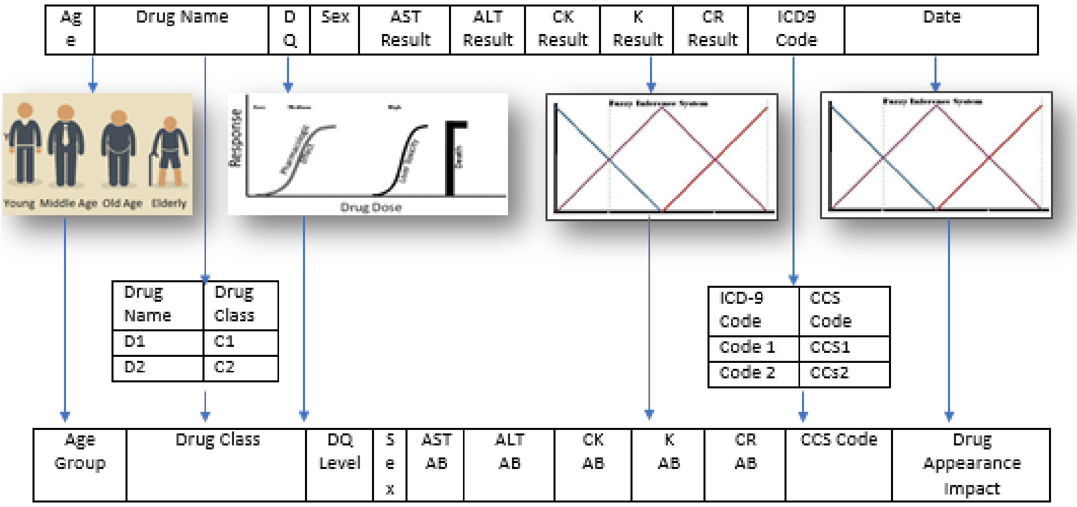
The preprocessed stage before using the association rule is shown in Figure 6 which was described in details in previous sections. Table 7 shows the Association rule input that leads to Suspected ADRs. The resulted ADRs will be added to ADR database after confirming them from expert physicians

Table 5. The ADRs Database

SUBSTANCE	DATE OF MOST RECENT SPC	ADR AS IT APPEARS IN THE SPC
HUMAN NORMAL IMMUNOGLOBULIN	10/27/2016	INCREASED BLOOD CREATININE
DARUNAVIR	1/4/2017	(DRUG) HYPERSENSITIVITY
DARUNAVIR	2/26/2016	(DRUG) HYPERSENSITIVITY
DARUNAVIR, COBICISTAT	1/26/2017	(DRUG) HYPERSENSITIVITY
ETHINYLESTRADIOL, NORELGESTROMIN	1/16/2014	(VULVO)VAGINAL FUNGAL INFECTION
VALSARTAN,	5/5/2015	ABASIA
VALSARTAN,	5/5/2015	ABASIA
VALSARTAN	5/5/2015	ABASIA
PENTOSAN POLYSULFATE	6/2/2017	ABDOMEN ENLARGED
HUMAN FIBRINOGEN, HUMAN THROMBIN	3/29/2017	ABDOMINAL ABSCESS
HUMAN THROMBIN	10/24/2016	ABDOMINAL ABSCESS

Table 6 The ICD9-CCS Database

'ICD-9-CM CODE'	'CCS CATEGORY'	'CCS CATEGORY DESCRIPTION'	'ICD-9-CM CODE DESCRIPTION'
'01000'	'1 '	'Tuberculosis'	PRIM TB COMPLEX-UNSPEC
'01001'	'1 '	'Tuberculosis'	PRIM TB COMPLEX-NO EXAM
'01104'	'1 '	'Tuberculosis'	TB LUNG INFILTR-CULT DX
'01105'	'1 '	'Tuberculosis'	TB LUNG INFILTR-HISTO DX
'01106'	'1 '	'Tuberculosis'	TB LUNG INFILTR-OTH TEST
'1122 '	'4 '	'Mycoses'	CANDIDIAS UROGENITAL NEC
'1123 '	'4 '	'Mycoses'	CUTANEOUS CANDIDIASIS
'1125 '	'4 '	'Mycoses'	DISSEMINATED CANDIDIASIS
'11282'	'4 '	'Mycoses'	CANDIDAL OTITIS EXTERNA
'11284'	'4 '	'Mycoses'	CANDIDAL ESOPHAGITIS (Begin 1992)
'0786 '	'7 '	'Viral infect'	HEM NEPHROSONEPHRITIS
'0787 '	'7 '	'Viral infect'	ARENAVIRAL HEM FEVER
'07881'	'7 '	'Viral infect'	EPIDEMIC VERTIGO
'1322 '	'8 '	'Oth infectns'	PHTHIRUS PUBIS
'1323 '	'8 '	'Oth infectns'	MIXED PEDICUL & PHTHIRUS
'0999 '	'9 '	'Sexual Infxs'	VENEREAL DISEASE NOS
'79505'	'9 '	'Sexual Infxs'	CERVICAL (HPV) DNA POS (Begin 2004)
'79515'	'9 '	'Sexual Infxs'	VAG HI RISK HPV-DNA POS (Begin 2008)
'79519'	'9 '	'Sexual Infxs'	OTH ABN PAP SMR VAG/HPV (Begin 2008)

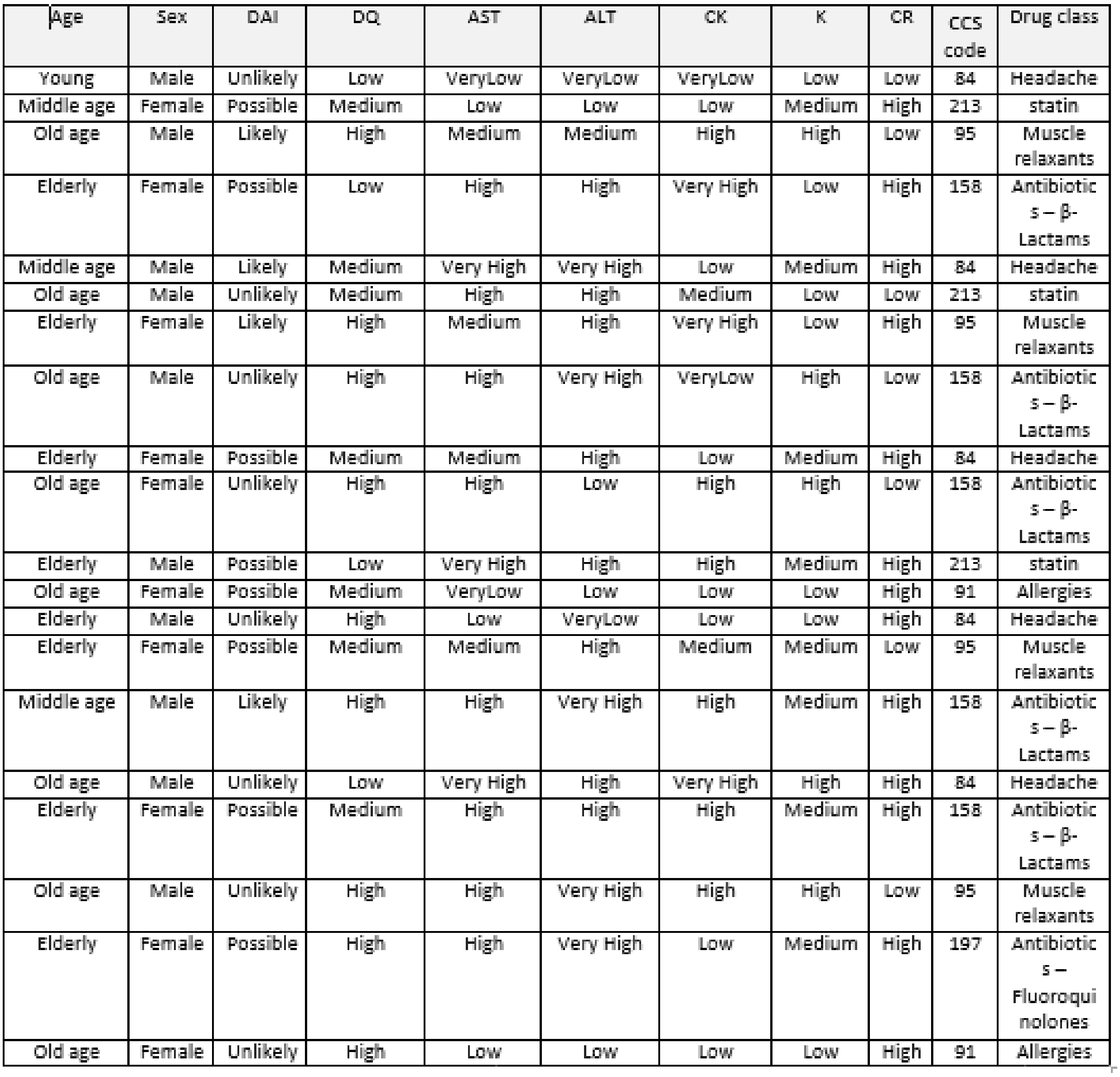


Drug Name	Drug Class
D1	C1
D2	C2

ICD-9 Code	CCS Code
Code 1	CCS1
Code 2	CCs2

Fig. 6. Database reprocessing stage

Table 7. The Association rule inputs



Age	Sex	DAI	DQ	AST	ALT	CK	K	CR	CCS code	Drug class
Young	Male	Unlikely	Low	VeryLow	VeryLow	VeryLow	Low	Low	84	Headache
Middle age	Female	Possible	Medium	Low	Low	Low	Medium	High	213	statin
Old age	Male	Likely	High	Medium	Medium	High	High	Low	95	Muscle relaxants
Elderly	Female	Possible	Low	High	High	Very High	Low	High	158	Antibiotics – β-Lactams
Middle age	Male	Likely	Medium	Very High	Very High	Low	Medium	High	84	Headache
Old age	Male	Unlikely	Medium	High	High	Medium	Low	Low	213	statin
Elderly	Female	Likely	High	Medium	High	Very High	Low	High	95	Muscle relaxants
Old age	Male	Unlikely	High	High	Very High	VeryLow	High	Low	158	Antibiotics – β-Lactams
Elderly	Female	Possible	Medium	Medium	High	Low	Medium	High	84	Headache
Old age	Female	Unlikely	High	High	Low	High	High	Low	158	Antibiotics – β-Lactams
Elderly	Male	Possible	Low	Very High	High	High	Medium	High	213	statin
Old age	Female	Possible	Medium	VeryLow	Low	Low	Low	High	91	Allergies
Elderly	Male	Unlikely	High	Low	VeryLow	Low	Low	High	84	Headache
Elderly	Female	Possible	Medium	Medium	High	Medium	Medium	Low	95	Muscle relaxants
Middle age	Male	Likely	High	High	Very High	High	Medium	High	158	Antibiotics – β-Lactams
Old age	Male	Unlikely	Low	Very High	High	Very High	High	High	84	Headache
Elderly	Female	Possible	Medium	High	High	High	Medium	High	158	Antibiotics – β-Lactams
Old age	Male	Unlikely	High	High	Very High	High	High	Low	95	Muscle relaxants
Elderly	Female	Possible	High	High	Very High	Low	Medium	High	197	Antibiotics – Fluoroquinolones
Old age	Female	Unlikely	High	Low	Low	Low	Low	High	91	Allergies

Using 400 cases not previously used in the training to check the performance of the system. The physicians examines 400. We examined agreement between the results generated by the developed system and the one by the physicians. We constructed the confusion matrix for each class (Non-ADR or Possible ADR). The confusion matrix is shown in Table 8.

Table 8. Confusion matrix.

	Actual (physicians)		
		Non -ADR	Possible ADR
System Detection	Non-ADR	193	11
	Possible ADR	7	189

The performance measurements used for this paper were recall, precision, classifier F- measure and accuracy. These measurements are defined in [16]. The performance measurements result is shown in Table 9. The develop detection system shows high accuracy on (up to 95.5%) and high agreement with physicians (0.91) on the test data.

Table 9. Performance Measurements.

Precision	**Recall**	**F- Measure**	**Class**
94.608%	96.5%	95.5%	Non- ADR
96.429%	94.5%	95.4%	Possible ADR

These measurements suggest excellent agreement between the proposed model and the physicians. According to this experiment, the system showed a superior performance and it was able to solve the problem efficiently.

Conclusion

A new methodology is developed to detect adverse drug reactions. The proposed method is unique in the sense it implements a model that uses a set of relations generated from real diagnostic data to detect suspected ADRs. In this paper, a set of rules for different types of medications are extracted to improve the Adverse drug reaction detection methodology. Association rule mining combined with fuzzy set theory is used to build the developed methodology. The developed system shows great results, so that the ADR detection proess is enhabced by the generated ADR signal pairs rules. The model shows high accuracy (up to 95.5%) in the test data. The resulted rule were validated by two expert physicians to confirm the generated ADRs. The result shows totally matching with ADRs defined by drug companies, medical assocations and FDA.

Acknowledgment

The authors wish to express their thanks for the support of Tafila Technical University and Al al-bayt university.

***Authors**: Dr Ayman M Mansour, Department of Communication, Electronics and Computer Engineering, Faculty of Engineering, Tafila Technical University, Tafila 66110, Jordan ; Dr. Fayez Khazalah, Department of Information Systems, Information Technology College, Al al-Bayt University, Mafraq 25113, Jordan; Dr. Mohammad A Obeidat, Department of Electrical Power and Mechatronics Engineering, Faculty of Engineering, Tafila Technical University, Tafila 66110, Jordan, Email: mansour@ttu.edu.jo*

REFERENCES

[1] I. R. Edwards and J. K. Aronson, "Adverse drug reactions: definitions, diagnosis, and management," Lancet, vol. 356, pp. 1255-9, Oct 7 2000.

[2] Gu, Lifang and Li, Jiuyong and He, Hongxing and Williams, Graham and Hawkins, Simon and Kelman, Chris, *Association rule discovery with unbalanced class distributions,* Australasian Joint Conference on Artificial Intelligence, Springer, (2003),221-232.

[3] Chen, Jie and He, Hongxing and Li, Jiuyong and Jin, Huidong and McAullay, Damien and Williams, Graham and Sparks, Ross and Kelman, Chris, *Representing association classification rules mined from health data,* International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, Springer, (2005), 1225-1231.

[4] Harpaz, Rave and Chase, Herbert S and Friedman, Carol, *Mining multi-item drug adverse effect associations in spontaneous reporting systems,* BMC bioinformatics, Springer 11(2010), No.9,1-8

[5] Sindhu, MS and Kannan, B, *Detecting signals of drug-drug interactions using association rule mining methodology,* Int J Comput Sci Inf Technol, Citeseer, 4 (2013), No.4 , 590-594.

[6] Ibrahim, Heba and Saad, Amr and Abdo, Amany and Eldin, A Sharaf, *Mining association patterns of drug-interactions using post marketing FDA's spontaneous reporting data*, Journal of biomedical informatics, Elsevier, 60(2016),294-308

[7] Guo, Kai and Lin, Hongfei and Xu, Bo and Yang, Zhihao and Wang, Jian and Sun, Yuanyuan and Xu, Kan, *Detecting potential adverse drug reactions using association rules and embedding models,* International Symposium on Bioinformatics Research and Applications, Springer, (2017), 373-378.

[8] Lee, Chang-Hung and Chen, Ming-Syan and Lin, Cheng-Ru, *Progressive partition miner: an efficient algorithm for mining general temporal association rules ,* IEEE Transactions on Knowledge and Data Engineering, 15 (2003), No. 4, 1004-1017.

[9] Shanmugapriya, K and Shanmugapriya, D and Parveen, H Summia and Niranjani, V, *N-Unexpected temporal association rule for diagnosing adverse drug reaction from health database,* International Proceedings of Computer Science and Information Technology (IPCSIT),18 (2011).

[10] Wang, Chao and Guo and et al, *Exploration of the association rules mining technique for the signal detection of adverse drug events in spontaneous reporting systems,* PloS one, Public Library of Science San Francisco, 7 (2021), No.7, e40561

[11] Reps, Jenna M and Aickelin, Uwe and Ma, Jiangang and Zhang, Yanchun , *Refining adverse drug reactions using association rule mining for electronic healthcare data,* IEEE International Conference on Data Mining Workshop, (2014), 763-770.

[12] Cai, Ruichu and Liu, Mei and Hu, Yong and Melton, Brittany L and Matheny, Michael E and Xu, Hua and Duan, Lian and Waitman, Lemuel R , *Identification of adverse drug-drug interactions through causal association rule discovery from spontaneous adverse event reports,* Artificial intelligence in medicine, Elsevier, 76 (2017), 7-15.

[13] Nikfarjam, Azadeh and Gonzalez, Graciela H, *Pattern mining for extraction of mentions of adverse drug reactions from user comments,* AMIA annual symposium proceedings, American Medical Informatics Association, (2011), 1019-1026.

[14] Segura-Bedmar, Isabel and de la Pena Gonzalez, Santiago and Martinez, Paloma , *Extracting drug indications and adverse drug reactions from Spanish health social media*, Proceedings of BioNLP, (2014), 98-106.

[15] Dingwei Dai and Chris Feudtner , *Association Rule Mining of Polypharmacy Drug Utilization Patterns in Health Care Administrative Data Using SAS Enterprise Miner*, sas-global-forum-proceedings, (2018),1-17.

[16] Mansour, Ayman M, *Decision tree-based expert system for adverse drug reaction detection using fuzzy logic and genetic algorithm*, International Journal of Advanced Computer Research, 8(2018), No. 36,110-128.

[17] Mansour, Ayman and Ying, Hao and Dews, Peter and Ji, Yanqing and Massanari, R Michael , *Fuzzy Rule-Based Approach for Detecting Adverse Drug Reaction Signal Pairs,* 8th Conference of the European Society for Fuzzy Logic and Technology, (2013), 384-391.

[18] Agrawal, Rakesh and Srikant, Ramakrishnan and others , Fast algorithms for mining association rules, Proc. 20th int. conf. very large data bases, Citeseer, 1215 (1994), 487-499.

[19] University of Waikato. Weka Software. https://www.cs.waikato.ac.nz/ml/weka/. Accessed 27 September 2021.

[20] R. Orchard, "Fuzzy reasoning in Jess: the Fuzzy J toolkit and Fuzzy Jess," 2001.