# A Multi-Agent System for Reporting Suspected Adverse Drug Reactions

Yanqing Ji[1,*], Fangyang Shen[2], See-Yan Lau[3], John Tran[4]

1. Department of Electrical and Computer Engineering, Gonzaga University, Spokane, WA
2. Department of Computer Systems Technology, New York City College of Technology, Brooklyn, NY
3. Department of Pharmacy, Abertsons Inc., Spokane, WA
4. Frontier Behavioral Health, Spokane, WA, USA

## Abstract

Adverse drug reactions (ADRs) represent a serious worldwide public health problem. Current post-marketing ADR detection approaches largely rely on spontaneous reports filed by various healthcare professionals such as physicians, pharmacists et.al.. Underreporting is a serious deficiency of these methods - the actually reported adverse events represent less than 10% of all cases. Studies show that two important reasons that cause the underreporting are: 1) healthcare professionals are unaware of encountered ADRs, especially for those unusual ADRs; 2) they are too busy to voluntarily report ADRs since it takes a lot of time to fill out the reporting forms. This paper addresses these two issues by developing a multi-agent ADR reporting system. The system 1) helps healthcare professionals detect the potential causal relationship between a drug and an ADR by analyzing patients' electronic records via both case-based analysis and statistical data mining approach; 2) allows healthcare professionals to add new rules to signal potential ADRs based on their medical expertise; 3) makes the reporting much easier by automatically linking the patients' electronic data with the reporting form. A functioning prototype of the system has been developed. The proposed data analysis approaches as well as the performance of the system have been tested. The results indicate that this system has a great potential to improve the spontaneous reporting rate of suspected adverse drug reactions.

Keywords: Adverse Drug Reactions, Post-marketing Drug Surveillance, Underreporting, Intelligent Agents, Multi-Agent Systems.

Abbreviations: Food and Drug Administration (FDA), Adverse Event Reporting System (AERS), Structured Query Language (SQL), Association Rule Mining (ARM), Multi-Agent System (MAS)

Correspondence: ji@gonzaga.edu*

## Introduction

Thousands of drugs are currently available on the U.S. market. These drugs can occasionally cause adverse drug reactions (ADRs) in patients. An ADR is defined by the World Health Organization (WHO) as "an effect which is noxious and unintended, and which occurs at doses used in man for either prophylaxis, or diagnosis, or therapy." [1]. ADRs could worsen the patient's medical problems, place patients in life-threatening situations, lead to increased health

care costs, and extend patients' length of stay in hospitals. For example, a recent study of ADRs in U.S. hospitals [2] indicates that about 2 million hospitalized patients per year experience serious ADRs and approximately 100,000 die from ADRs, which rank the fifth in leading causes of death as well as being one of the more frequent causes of hospitalization. In addition, patients experiencing an ADR while hospitalized have a substantially increased risk of death (approximately 20%) and incur significantly higher costs for hospitalization and drug treatments (estimated at $516 million in total costs and $37 million in additional drug costs for Medicare patients alone [2]).

Even though premarketing clinical trials are required for all new drugs before they are approved for marketing, these trials are necessarily limited in size and duration and are not capable of detecting relatively infrequently occurring adverse reactions caused by normal use of drugs. In general, if the occurrence rate of a potential ADR is less than 0.1%, it cannot be recognized in randomized and controlled clinical trials due to limitations in trial size [3]. As an example, bromfenac (marketed as Duract) was a nonsteroidal antiinflammatory (NSAID) agent that was removed by the Food and Drug Administration (FDA) from the market in 1998, less than 1 year after it was introduced, because it caused liver injury in 1 in approximately 20,000 patients taking the drug for longer than 10 days [4]. This type of rare and delayed drug reaction is very difficult to detect during preapproval clinical trials since most new drugs are approved after an average of only 1,500 patient exposures.

Drug safety depends heavily on postmarketing surveillance - the systematic detection and evaluation of medicines once they have been marketed [3]. Current postmarketing methods largely rely on voluntary, spontaneous reports of suspected ADRs to be filed by various healthcare professionals such as physicians, pharmacists, and others. To ensure quality of patient care, many healthcare facilities (e.g., hospitals) have a drug safety office which collects spontaneous ADR reports and monitors the safety of drugs, particularly recently marketed drugs. At the national level, many countries have their own spontaneous reporting systems. For example, the United States has an ADR reporting system called MedWatch, which is managed by the FDA. Internationally, the World Health Organization (WHO) also maintains a similar reporting system that collects individual case safety reports from healthcare providers in member countries of the WHO Program.

Because ADR reports are filed at the discretion of the users in the reporting system, there is gross underreporting [5-7]. Hazell and Shakir examined 37 studies from 12 countries and found that the median underreporting rate across these studies was 94% (interquartile range 82-98%) [6]. Because of underreporting, long delays occur before a sufficient number of suspect cases (signals) are accumulated to raise concern among clinicians and regulators. Consequently, the current post-marketing surveillance approach may require years to identify and withdraw problematic drugs from the market [8,9], and result in unnecessary mortality, morbidity, and cost of healthcare.

Studies have shown that lack of time and uncertainty of reactions being caused by drugs are two major reasons that hinder reporting [10-12]. As indicated by the study of [10], more than 35% of medical practitioners have the opinion that reporting of ADRs takes too much time, and it is too bureaucratic. For instance, the FDA's reporting form has four sections: patient information, adverse event, suspect medication(s), and reporter information. For each section, a lot of specific details are requested. For example, the adverse event section requests outcomes attributed to

2

adverse event, date of event, date of report, description of event, relevant tests/laboratory data including dates, and other relevant history including preexisting medical conditions (e.g. allergies, race, pregnancy, smoking and alcohol use, hepatic/renal dysfunction, et.al.). Finding these details and putting them in a reporting form are both tedious and time-consuming.

Uncertainty as to whether the reaction is caused by a drug is another important factor for underreporting [10]. Elucidating causal relationships between a drug and an outcome of interest is often clouded by broad boundaries of uncertainty. Reasons for uncertainty include (i) multiple conditions may result in similar outcomes (e.g., a possible ADR); (ii) a drug may be essential for producing an outcome or condition of interest; however, it is not a sufficient cause in and of itself. That is, there may be other circumstances such as genetics or other underlying conditions that participate in concert with the drug to produce the adverse outcome.

Intelligent agents provide a promising approach to ameliorating the underreporting problem because they can act on behalf of human users to find and filter information, automate complex tasks, and collaborate with other software agents to assist complex decision-making. Even though intelligent agents have been used for different purposes, they usually share some common characteristics such as autonomy, collaboration, delegation, and communication. A multi-agent system may be defined as "a collection of autonomous agents that communicate between themselves to coordinate their activities in order to be able to solve collectively a problem that could not be tackled by any agent individually" [13].

The multi-agent technology provides a new paradigm for developing artificial intelligence applications in a variety of fields. In the health care domain, exemplary applications include the AADCare (agent architecture for distributed medical care) system for integrating the patient management processes [14], the Organ Transplant Management system for the management of organ and tissue transplants among different medical centers [15], the internal hospital tasks management systems for monitoring the application of medical protocols [16], the ADR-Monitor system for monitoring potential ADRs of interest using electronic patient records [17], and the multi-agent system used for finding similar patients [18]. Up to date, the number of biomedical applications is still very limited.

In this paper, we develop a functioning prototype of an agent-based spontaneous ADR reporting system. The system provides both case-based analysis and statistical data mining approach to help healthcare professionals establish the causal relationship between a drug and a potential ADR. In addition, the system has a rule-adding framework that allows healthcare professionals to add new rules to signal potential ADRs based on their medical expertise. Moreover, once a healthcare professional decides to file an ADR report, his/her assistant agent will retrieve the required data from related electronic health databases and automatically fill a significant portion of the reporting form. As a result, filing a report would become much easier and takes significantly less time than existing methods.

## Methods

### Multi-Agent System Architecture

Our agent system architecture designed for a healthcare organization is shown in Figure 1. Note that a healthcare organization can be a hospital, a medical center, a sizable clinic, or any other

health organizations in which people would like to file ADR reports. There are four types of agents: assistant agent, data mining agent, management agent, and data collection and monitoring agent. The first three types of intelligent agents are accessible by any computer connected to the Internet through web pages. All these agents can communicate with each other through agent communication languages. The databases at a healthcare organization contain detailed information of its patients. The data are important since they provide, among other information, direct or indirect evidence of the clinical state, and comorbid conditions.
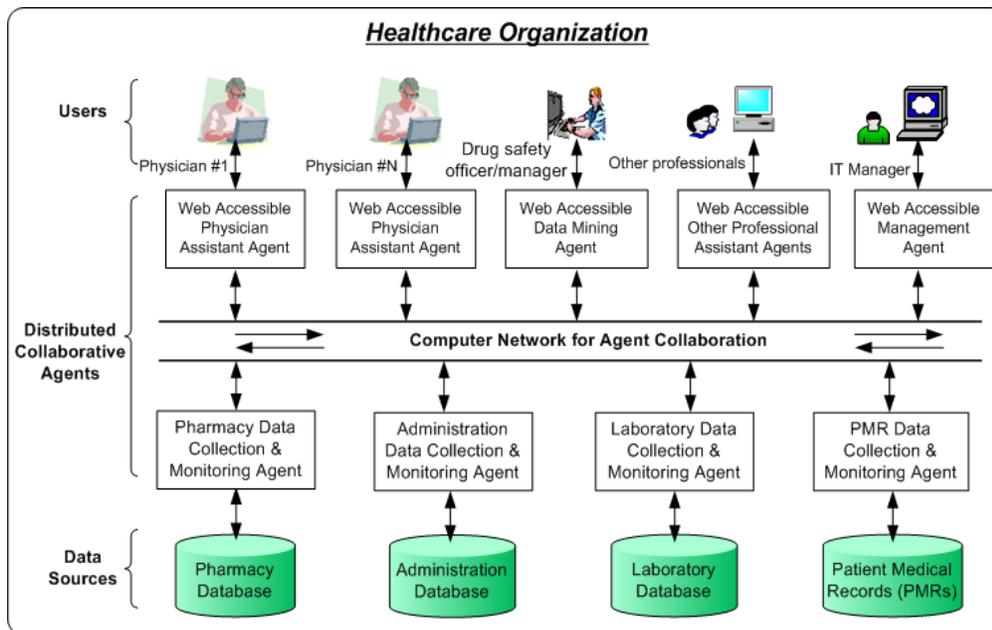


**Figure 1:** Multi-agent system architecture

Physicians or other healthcare professionals have their own intelligent assistant agents, which would help them effectively file ADR reports. First, each assistant agent can communicate with various data collection and monitoring agents and obtain the electronic data of the involved patient for a particular ADR. More importantly, the agent is powered by a rule-based algorithm which can assess the degree of causal association between a drug and a potential ADR based on the retrieved patient data related to a particular patient. Second, a healthcare professional can add new rules (through his/her assistant agent) to signal potential ADRs based on his/her expertise and experience in a particular class of drugs or ADRs. Third, an assistant agent can contact the data mining agent to get high-level statistical information regarding the strength of the association of a drug-ADR pair. Fourth, if a healthcare professional decides to file an ADR report, his/her agent would automatically feed the available data to the reporting form. More details regarding the rule-based algorithm, the data mining method, the rule-adding framework, and the auto-filling process will be described in the next section.

The data mining agent is maintained by a safety officer/manager who has background in epidemiology or biostatistics. The agent collects statistical data through data collection and monitoring agents and calculates the value of an adopted interestingness measure for suspected drug-ADR signal pairs. The value of the measure indicates the strength of the association of a drug-ADR pair from a statistical perspective. The agent keeps a list of signal pairs and

periodically updates the measure's value for all pairs. If a healthcare professional wants to look at the statistical strength of a particular signal pair, the data mining agent would search its list and send the information to his/her assistant agent.

The management agent provides agent management, a white-page service through which an agent can register and discover other agents, and a yellow-page service through which an agent can publish and search for "services." An agent is allowed to subscribe services. Whenever a service registration or modification is made, the agent will be notified.

The data collection and monitoring agents convert information requests from other agents into appropriate SQL (Structured Query Language) statements and retrieve the data from various databases. Moreover, they help other agents track further development of the suspected patient cases. For example, a physician may be interested in what happens to the patients after the suspected drug is discontinued. In general, the physicians and other healthcare professionals could more easily determine the causal relationship between a potential ADR and a drug by monitoring the related cues.

## System Design and Functionalities

In this section, we present the detailed system design and functionalities. Specifically, we discuss how the agent-based system assists healthcare professionals in submitting ADR reports, by providing both case-based and statistical decision support, by allowing healthcare professionals to add new rules for evaluating the potential causal association between a drug and an ADR, and by automatically filling out most of the reporting form using electronic health data.

### *Decision Support*

To help healthcare professionals confirm potential ADRs, our reporting system provides two types of decision support. First, each assistant agent can utilize a rule-based algorithm to assess the degree of causal association between a drug and an ADR based on the electronic data of a single patient. Second, the data mining agent can also do a numerical analysis of the correlation between a drug and ADR using a data mining method.

1) Single Patient Evaluation Using a Rule-Based Algorithm

The physician and pharmacist in the project team developed a rule-based algorithm (Table 1) in order to evaluate the potential association between a drug and an ADR on the basis of a patient's electronic data. The design of the algorithm was based on their own expertise and the literature research [19]. The algorithm poses a list of questions that indicate the correlation of a drug-ADR pair from different aspects. Each question has a set of answers, each with a score or "weight". For instance, the answer to an example question "Is the patient 65 or older?" might be "Yes [+1]", "No [-1]", or "unknown [+0]". The number in the bracket is a score assigned to the corresponding answer. One can see that each answer represents an IF-THEN rule. For instance, the answer "Yes [+1]" to the above question indicates "IF the patient is 65 or older, THEN the score of this question is +1 for the patient". The score for a patient case is calculated by summing up the scores of each question. Depending on the total number of scores accrued, one of the four labels – "definite", "probable", "possible", or "doubtful" - is assigned to the ADR.

2) Statistical Association Analysis Using a Data Mining Method

Studies have shown that data mining techniques can be very helpful in analyzing the electronic health data and discovering potential ADRs [20-24]. However, mining the association between a drug and an ADR is not easy due to the large size of the data. Since electronic patient data are generally stored in relational databases, we developed a SQL-based association mining algorithm that can deal with large amount of data. The algorithm employs risk ratio to evaluate the strength of association between each drug-ADR pair given all the electronic data in a database.

Risk ratio (RR) provides a measure of association between an exposure and the outcome of interest. In mathematical epidemiology, it is defined as a ratio of the probability of the event occurring in the exposed cohort versus a non-exposed cohort. Risk ratio is used widely in the statistical analysis of binary outcomes where the outcome of interest has relatively low probability. The magnitude of risk ratio may vary from 0 to infinity. When the risk ratio is over 1.0, the incidence of the outcome is greater in the exposed cohort than the unexposed cohort. In our study, this method is used to investigate the potential association between the drug of interest (exposure) and a reaction (outcome). Specifically, it is employed to compare the risk of developing an ADR in people receiving a drug as compared to the people who are not receiving the drug. Its calculation is based on a 2x2 contingency table as shown in Table 2. The formula is given below:

$$RR = \frac{a/(a+c)}{b/(b+d)} \qquad (1)$$

**Table 1:** Rule-based algorithm

| |
|---|
| When did the adverse drug reaction occur?<br> • Prior to Drug Administration [-1]<br> • Unknown [0]<br> • Less than two weeks later [+3]<br> • Less than one month later [+2]<br> • Less than six months later [+1]<br> • More than six months later [0] |
| Is the patient 65 or older?<br> • Yes [+1]<br> • Unknown [0]<br> • No [-1] |
| Did the reaction stop if the drug was no longer administered?<br> • Yes, within one week [+2]<br> • Yes, within one month [+1]<br> • Yes, more than one month later [0]<br> • No [0] |
| If the patient was given the drug again, did the reaction re-occur?<br> • Yes, within one month [+2]<br> • Yes, more than one month later [0]<br> • No [-3] |

6

The total score calculated from the above questions defines the category of an ADR as follows:

 <= 1 [Doubtful]
 2 – 4 [Possible]
 5 – 8 [Likely]
 9 – 10 [Definite]

However, the risk ratio value alone is insufficient to determine association since we must ask whether an estimate for the risk ratio is statistically significant. Can the estimate that is greater than 1.0 be attributed to chance, or is it sufficiently greater to require further scrutiny? For statistical inference in this study, we also calculate the Chi-Squared ($\chi^2$) value as shown below:

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \qquad (2)$$

Where *n* represents the number of cells in the table and it is equal to 4. That is, the summation applies over all four cells of the 2x2 contingency table. $O_i$ is the observed frequency and $E_i$ is the expected (theoretical) frequency asserted by the null hypothesis. For example, $O_1$ and $E_1$ for the first cell should be calculated as

$$O_1 = a \qquad E_1 = \frac{(a+b)(a+c)}{(a+b+c+d)}$$

**Table 2:** Two-by-two contingency table for reactions

|  | Patients with drug of interest | Patients with all other drugs |
|---|---|---|
| Patients with reaction of interest | *a* | *b* |
| Patients with all other reactions | *c* | *d* |

For the remaining cells it takes the value contained in the cell (i.e. b, c, d) in turn.

To confirm a drug-reaction signal pair, we also include another factor: the number of patients with the reaction of interest in the database (i.e., *a* in the Table 2). The reason is that, even a fairly small *a* can possibly produce a high risk ratio. But, in post-marketing surveillance, a signal pair generated by an interestingness measure is generally not considered as valid if only one or two patient cases contain the pair [25].

Considering the above three factors, we mark a drug-reaction pair as a signal when it satisfies: 1) $RR > 2$, 2) $\chi^2 > 4$, and 3) $a \geq 3$. Hence, the association mining task is to find all the drug-reaction pairs that satisfy the three criteria in a relational database. We use $D(d_1, d_2, \dots d_i, \dots d_m)$ to represent a list of all drugs in the pharmacy database. Similarly, $R(r_1, r_2, \dots r_j, \dots r_n)$ is utilized to represent a list of all reactions (or medical conditions) covered by the patient medical record database. Each pair $< d_i, r_j >$ forms a potential association $d_i \rightarrow r_j$ whose *RR* value needs to be

7

computed based on the contingency table shown in Table 2. We assume that all the data needed to calculate the RR value are in three tables: Demography Table containing patient's demographic information such as age, gender, weight, etc., Drug Table consisting of drug-related data such as drug name, dose, start date, etc., and Reaction Table containing reaction-related data such reaction name, start date, and so on.

Our association mining algorithm is shown in Algorithm 1 where the same variables (i.e. *a, b, c,* and *d*) in Table 2 are used in the description of the algorithm. To compute *RR* based on Eq (1), four values need to be obtained; i.e., *a, a+c, b,* and *b+d*. Six getxxx() functions are utilized to get all drugs and reactions as well as all the information needed to compute the four values. The two functions getAllDrugs() and getAllReacs() retrieve all the distinctive drugs and reactions from Drug Table and Reaction Table, respectively (line 1 and 2). The function getTotalNumOfPatients() is used to obtain the total number of patients (line 3) from Demography Table. Note that each patient is uniquely identified by a patient identification (PID) number that represents the primary link field (primary key) between database tables. In addition, one or more drugs and reactions may appear in a single patient record. For each distinctive drug $d_i$, the function getTotalNumOfPatientsOnDrug ($d_i$) is used to obtain *a+c* (from the Drug Table) which represents the total number of patients who have taken the drug (line 5). Since the total number of patients is equal to *a+c+b+d*, *b+d* can be obtained by subtracting *a* and *c* from the total number of patients (line 6). Given a pair $< d_i, r_j >$, the function getTotalNumOfPatientsOnDrug&Reac($d_i$,$r_j$) is used to get *a* (line 8) based on both Drug Table and Reaction Table. The value of *b* is obtained by first getting the total number of patients on the reaction $r_j$ (which represents *a+b*) and then subtracting *a* from this number (line 9). After all the four values for *a, a+c, b,* and *b+d* are obtained, *RR* and $\chi^2$ can be computed using Eq (1) and (2), respectively.

---

**Algorithm 1** SQL-based association mining algorithm

Input: relational databases
Output: *RR,* $\chi^2$, and *a* for each drug-ADR pair

```
1:  D(d₁, d₂, … dᵢ, … dₘ) = getAllDrugs(); // get all distinctive drugs
2:  R(r₁, r₂, … rⱼ, … rₙ) = getAllReacs(); // get all distinctive reactions
3:  totalNumOfPatients = getTotalNumOfPatients();
4:  for each drug dᵢ ∈ D, do
5:      a+c = getTotalNumOfPatientsOnDrug(dᵢ);
6:      b+d = totalNumOfPatients − (a+c);
7:      for each reaction rⱼ ∈ R, do
8:          a = getTotalNumOfPatientsOnDrug&Reac(dᵢ, rⱼ)
9:          b = getTotalNumOfPatientsOnReac(rⱼ) - a
10:         RR = (a/(a + c))/(b/(b + d));
11:         compute χ² based on Eq.(2)
12:         output RR, χ², and a for <dᵢ, rⱼ>.
13:     end for
14: end for
```

---

The data mining agent maintains a list of drug-reaction pairs. For each pair, its *RR*, $\chi^2$, and *a* are updated periodically (e.g. each week) under the help of data collection and monitoring agents. A pair is marked as a signal when it satisfies the criteria defined above. A higher risk ratio value indicates a stronger signal when the other two criteria are satisfied. If a healthcare professional wants to confirm a potential ADR before reporting it, his/her assistant agent can always contact the data mining agent and bring the above three values to him/her.

### *Rule-adding Framework*

To make our multi-agent system more flexible and useful, a rule-adding framework is designed to allow healthcare professionals to adjust existing rules or adding new rules in the rule-based algorithm based on their knowledge of a particular class of drugs or ADRs. For example, a class of drugs frequently used in a particular medical domain (e.g., Psychiatry) may cause an ADR only after the drug toxins are accrued to a certain level after patients have taken the drug(s) for a long time. In such a case, healthcare professionals in the domain may want to adjust the scores assigned to each answer for the question "When did the adverse drug reaction occur?" in Table 1. In addition, other factors (e.g., abnormal lab results) can indicate the occurrence of ADRs caused by a particular class of drugs. In these cases, new questions and rules can be added to the algorithm through the framework.

Virtually any "question" that can be expressed in English and answered using information from a hospital database can be added to the framework. Healthcare professionals have the ability to select which signals would be used to indicate a potential ADR and add their medical expertise to our framework using their assistant agents. Any rule, whether a part of the rule-based algorithm or a custom rule added by a healthcare professional, is stored in a table in a database. Each rule consists of four parts: *rule ID*, *descriptions*, *status*, and *code*. The *rule ID* represents a unique identifier of a rule. The *descriptions* part provides specific descriptions of a rule in English language. The *status* (either active or non-active) of a rule indicates whether a rule is used when calculating a patient's total score. The *code* portion is critical, as it formally describes what variable is evaluated, the upper or lower limit, and the score if the condition of a rule is satisfied.

We developed a custom language to make up the *code* section of a rule. The custom language is translated into a SQL command that is then executed. Generally, the code follows this format: First of all, a database table name and a column name within that table are listed. Then an operator and a constant is listed, followed by a score. The table/value, the operator, and the constant combine to form the IF-condition of a rule. If the condition is answered "yes", then an appropriate score is assigned to the rule. The *code* for an example question "Is the patient greater than 65 years old?" is expressed as "Patient Table|Age|>|65|1", where the vertical bars (|) represent the delimiters of different components of the *code*. The "Patient Table" refers to a table in the hospital database from which data are retrieved and evaluated. The "Age" represents a column name in the "Patient Table". The symbol ">" is an operator that represents "greater than". The two numbers "65" and "1" represent the lower limit of "Age" and the score, respectively. Table 3 is a list of operators supported by the custom language. Note that users do not need to learn the custom language and manually create the *code* for each rule. An easy-to-use web interface is provided to allow users to select the patient table, column name and operator,

9

and enter the upper and/or lower limit and the score. The system then automatically generate the *code* based on the inputs.

**Table 3:** List of operators in the custom language

| Operator | Description |
|---|---|
| < | Less than |
| <= | Less than or equal to |
| > | Greater than |
| >= | Greater than or equal to |
| = | Equal to |
| != | Not equal to |
| Between | Between two specified values |
| Diff | Difference between two values |

### *Auto-Filling*

Analysis of various ADR reporting forms indicates that the majority of information requested can be obtained from the healthcare organization's electronic health databases. This includes patient demographics and disease history, reporter identification, and general information regarding medication. By developing a system that automatically retrieves the information and fills them in, we can reduce the amount of work required by the reporters. As a result, the professionals would become more willing to file an ADR report.

We have implemented the auto-filling property using the FDA's spontaneous reporting form. The user of our prototype system is provided with a web-based interface through which he/she can login the reporting system. After that, the user is taken to a new page and linked with his/her assistance agent. For a new user of the system, he/she needs to apply for an account and then a new unique assistant agent will be created for him/her by the system. Note that an agent keeps running in the background unless it is deleted. To report an ADR, the user enters two identification numbers which identify the patient and medication suspected (i.e., Patient_ID and Medication_ID). From there the assistant agent will retrieve all the information needed to fill out the FDA form except the Date of Event and a Description of the Event. This is a significant reduction in the amount of information the reporter is required to fill out. The system will then return to the user a form with all of the information filled out so that the user can review the form before submitting it.

## System Implementation

The implementation of the ADR reporting system follows the Model-View-Controller (MVC) pattern as shown in Figure 2 [26]. The *model* consists of the intelligent agents of the system. The *view* defines the presentation layer of an application which consists of web pages in our system. The *controller* manages the flow of the system, and this work is done by servlets. A servlet is a small Java program that receives requests from Web clients, usually across the HyperText Transfer Protocol (HTTP). The requests can be forwarded to the *model*, which may trigger data-based analysis. The analysis results will be returned to the servlet and then be forwarded to the *view* for presentation. The *Data Store* in Figure 2 represents the various electronic health

10

databases in our system. Since we use a couple of tools that are implemented in Java, we build the system using the Java programming language.
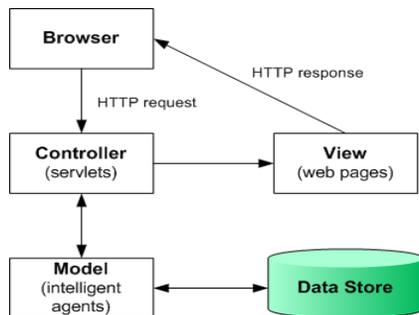


**Figure 2:** Block diagram of MVC model

### *Web-Based Interface*

We implemented our web interface using Apache Tomcat, an open source software implementation of the Java Servlet and JavaServer Pages (JSPs) technologies. Tomcat is used as both the web server and the servlet and JSP engine for our web-based interface. JSPs allow us to embed basic Java code in HTML pages, while servlets do the processing for the dynamic web pages of our system. NetBean was used as the development environment for developing the whole system including JSPs, servlets, intelligent agents, and interface to databases.

### *Intelligent Agents*

Agent-related implementation issues include agent creation, communication, interoperability, and service registry. We found that the Java agent development environment (JADE [27]), one of the agent system development software packages, could readily handle them. We utilized it to implement these aspects of the intelligent agents in our reporting system. JADE is an integrated tool suite for constructing, managing, and testing intelligent software agents. It fully complies with the Foundation for Intelligent Physical Agents (FIPA) specifications and adopts FIPA ACL as agent communication language. The FIPA standards allow agents effectively interoperate with each other. JADE provides both white-page and yellow-page services for agent management. These features free us from many of the implementation details of developing and managing agents, and allow us to focus our time and energy on developing their behaviors and interactions.

### Implementation of Rule-adding Framework

As explained earlier, each rule in the rule-based algorithm is formally represented by its *code* which will be automatically translated into a SQL statement. For example, the *code* for question "Is the patient greater than 65 years old?" is "Patient Table|Age|>|65|1". It would be translated into a SQL statement "SELECT PID, Age FROM `Demography Table` WHERE Age>65". The patient identifier (i.e., PID) is automatically added to the statement and it is known from the context. That is, when a healthcare professional intends to file an ADR report, he/she must identify the patient relevant to the report so that his/her assistant agent can analyze the patient's records using the rule-based algorithm. This SQL query would return each row in the database that contains the desired patient's id and an age value greater than 65. If a row was correctly returned, we know the patient is older than 65 years, so the appropriate score should be assigned.

We implemented web pages for the rule-adding framework through which users can edit existing rules or add new ones based on their medical expertise.

# Results

## Experiment Data

To test our ADR reporting system, we created four relational MySQL databases (i.e., pharmacy database, administration database, laboratory database, and patient medical record database), each of which contains certain types of data. The creation of the databases were based on real data from the FDA's Adverse Event Reporting System (AERS) and our past experience in working on real health datasets retrieved from the Veterans Affairs Medical Center in Detroit, Michigan. The AERS is a database that contains approximately 4.5 million ADR reports, with approximately 100,000 additional reports released each quarter. The AERS data is freely available in plaintext format on the FDA's website. When importing the AERS raw data into our databases, we made some changes. For example, we standardized ages by converting from days, weeks, months, or decades into years. We also pruned out patients missing any data regarding drug name, or therapy start/end dates. As a result, we obtained 801,623 patient cases in our databases based on the AERS data released in the last two years.

## Experiment Results

Using the above data, we tested the two decision-making strategies (i.e., case-based analysis using the rule-based algorithm and statistical analysis using *RR*) imbedded in our system, the scalability of the association mining algorithm as well as how indexes affect the performance of the algorithm. The experiments were done on a PC with a 3.4GHz Intel i7-3770 processor and 8 GB of RAM, running Windows 7.

We first examined the degrees of agreement between the causality categories (i.e., "Definite", "Likely", "Possible", and "Doubtful") assigned by the rule-based algorithm and those by an experienced physician. To do this, we randomly selected 100 patient cases, within each of which a distinctive drug-ADR pair was reported. According to the HIPAA (Health Insurance Portability and Accountability Act) privacy regulations, a healthcare professional can only access the data of his/her own patients. We assume that the selected patients are associated with a particular healthcare professional. These 100 patients were used to validate the assignment of causal link strengths by the proposed rule-based algorithm. Since our patient data were derived from FDA's spontaneous reports, the reported drug and potential ADR are known for each patient case. The algorithm was used to identify and assign likelihoods of causal associations between the drug (cause) and ADR (effect) for the 100 patients. To provide preliminary validation for the capacity of the rule-based algorithm, the physician on our project team participated in the experiment. He was provided a general overview of the objectives of the study and was asked to review each of the 100 patient cases and make a judgment regarding the likelihood of a significant causal association between each drug-ADR pair. To make the following agreement assessment easier, both the rule-based algorithm and the physician assigned a numerical number between 1 and 4 based on the strength of the perceived causality, where 4 means "definite causal relationship," 3 is "likely casual relationship," 2 indicates "possible causal relationship," and 1 stands for "doubtful causal relationship.

Because scoring of the causal association was based on ordinal data, we utilized the weighted Kappa statistic [28,29] to estimate the levels of agreement between the rule-based algorithm and the physician. The Kappa coefficient is an estimate of the agreement between two raters after chance agreement is controlled. Kappa scores range between 1 (complete agreement) and 0. Because there are no "absolute" interpretations of the Kappa coefficients, experts have offered opinions regarding interpretations of agreement. Landis and Koch [30] suggest that for values of Kappa greater than 0.75, there is excellent agreement. For values less than 0.4, there is poor agreement, and for values between 0.40 and 0.75, there is fair to good agreement. The calculated Kappa between the algorithm and the physician is 0.894. This coefficient suggest excellent agreement between the proposed algorithm and the physician. We computed the 95% confidence intervals using both the Asyptotic formula and the Jackknife method. The results of these two methods agreed to the second decimal as shown in Table 4.

**Table 4:** 95% confidence intervals for weighted Kappa coefficients from Asymptotic formula and Jackknife method

|  | Kappa Coefficient | Asymptotic 95% confidence interval | Jackknife 95% confidence interval |
|---|---|---|---|
| Physician vs. rule-based algorithm | 0.894 | (0.837, 0.952) | (0.838, 0.954) |

We also examined the consistence between the case-based analysis using the rule-based algorithm and the statistical analysis using $RR$. For each drug-ADR pair in the 100 patients, its $RR$, $\chi^2$ and $a$ were calculated based on all the 801,623 patient cases. Whether a pair represents an ADR signal is determined by the criteria ( $RR > 2$ ; $\chi^2 > 4$ ; $a \geq 3$ ) established earlier. Table 5 gives the number of drug-ADR pairs in each causality category based on the rule-based algorithm as well as how many pairs in each category represent ADR signals according to the statistical analysis criteria. One can see that, among the 100 drug-ADR pairs in the selected patients, 68 of them represent ADR signals that are confirmed by the statistical approach. In addition, if a pair exhibits a stronger causality according to the case-based analysis, it more likely represents a signal established by the statistical analysis.

**Table 5:** The number of drug-ADR pairs in each causality category based on the rule-based algorithm as well as the number pairs (in each category) that represent ADR signals according to the statistical analysis criteria

|  | "Definite" | "Likely" | "Possible" | "Doubtful" | Total |
|---|---|---|---|---|---|
| Number of pairs in each category based on the rule-based algorithm | 18 | 43 | 27 | 12 | 100 |
| Number of pairs that represent ADR signals according to the statistical analysis. | 17 | 34 | 15 | 2 | 68 |
| Percentage of pairs that represent ADR signals | 94.4% | 79.1% | 55.6% | 16.7% | N/A |

Since our association mining algorithm is SQL-based its performance can be greatly affected by database indexes. In addition, the number of database connections can also affect the performance of a SQL-based algorithm. To query data, one approach is to create a connection to the related database for each SQL statement and then close the connection after the query. Another approach is to create a single connection for each database that will be used by all the SQL statements in the algorithm. We tested how different combinations of these two factors (i.e., index and number of connections) affect the performance of our algorithm. The experiment data contains 5,281 drugs and 7,943 reactions. Since the execution time of the program without indexes is very long we randomly chose 100 drugs and 100 reactions for this experiment. That is, 10,000 (100x100) association rules were evaluated by calculating their $RR$, $\chi^2$ and $a$ values. The experiment results are provided in Table 6. As expected, the results show that indexes dramatically affect the overall performance of our SQL-based algorithm. The number of connections has less effect on the performance. However, if indexes are set properly, the time consumed by making a lot of connections are still relatively high (see the last two columns of Table 6). Note that a user of the system normally checks one drug-reaction pair each time, which takes a very small amount of time (comparable to 1 millisecond with indexes and one connection for each database). In addition, the data mining agent can be configured to automatically run the algorithm after every certain period of time (e.g., one month) after new data are entered into the database. Then, a query of a pair would be simply to check a table that contains the $RR$, $\chi^2$ and $a$ values for each pair.

**Table 6:** Execution time of the association mining algorithm under different experiment conditions

| | Experiment Conditions | | | |
|---|---|---|---|---|
| | No Indexes, One Connection for Each Statement | No Indexes, One Connection for Each Database | With Indexes, One Connection for Each Statement | With Indexes, One Connection for Each Database |
| Execution Time (seconds) | 7935 | 7793 | 45 | 7.2 |

As indexes are so important for reducing the execution time of the algorithm, we decided to test how different index storage types affect the performance using the same dataset. The MySQL utilized in this study supports two index types: HASH and BTREE. Note that when a new index type is applied to the database, the first run of the program would take much longer time than the following runs since the computer needs time to build the indexes and store them into a new data structure. We examined the execution time for the first run of the program and the average time for the next 10 runs with each index type. The results are presented in Table 7. The table indicates no significant difference of the execution time when using different index types.

**Table 7:** Execution time of the association mining algorithm under different index storage types

| | Index Storage Type | |
|---|---|---|
| | HASH | BTREE |
| Execution Time for the First Run | 34.1 | 34.8 |

| Average Execution Time for the Next 10 Runs | 7.3 | 7.2 |
|---|---|---|

As each health organization may have a lot of patient data, the scalability of our SQL-based association mining algorithm is very important. We tested the scalability of Algorithm 1 in terms of the size of the data. We started with 100,000 patient cases, and then added additional 100,000 cases each time the program was executed. The execution time against the number of patient cases was plotted as shown in Figure 3. One can see that our algorithm is scalable in terms of the size of the patient data.
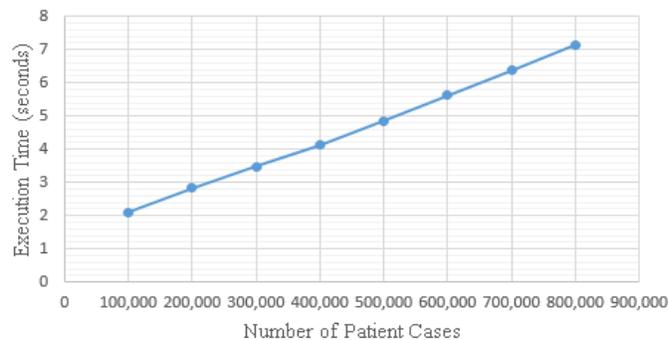


**Figure 3:** Scalability of the association mining algorithm in terms of the number of patient cases

## Discussion

The proposed rule-based algorithm can be used to assess the degree of association between a drug and a potential ADR within a single patient case. However, even if the causality category assigned by the algorithm is "likely" or "definite", that may not represent real causal association due to the complexity of the ADR problem. For example, a syndrome or an adverse effect may be caused by various other conditions (e.g., underlying diseases), even though the syndrome and the associated drug satisfy one or more rules in the algorithm given the patient data. In addition, a reaction of a drug may indirectly cause an adverse effect. In this sense, the rule-based algorithm should not be used for the purpose of accurately establishing the causal relationship between a drug and an ADR. Instead, it is an assistant tool that would encourage healthcare professionals to file spontaneous reports of suspected ADRs.

The statistical analysis results given by the proposed data mining method provide another perspective for healthcare professionals to examine the drug-ADR pair of interest. Please note that some false ADR signals may be generated due to the nature of the ADR problem and the complication of medical data. First, ADRs normally occur very infrequently. Thus, the adverse effects caused by ADRs can be hidden by other high frequency symptoms. Second, frequently prescribed drugs are often associated with common symptoms by coincidence. These data noises could reduce the effectiveness of any association mining methods. Third, the electronic data may be incomplete and has potential bias and even errors [31]. For example, some potential adverse effects may not be recorded in a health database if the results caused by drug reactions are not serious. Please note that, in post-marketing surveillance, data mining is primarily used to generate ADR signals that will be subject to further analysis. To establish true causal associations, other labor-intensive approaches such as epidemiology study, case review, and

15

interpretation by drug safety professionals experienced in the nuances of pharmacoepidemiology and clinical medicine are often used [32].

The rule-based algorithm and the data mining method are complementary to each other. While the former examines the causal association between a drug and an ADR based on the data within a single patient case, the latter assesses the association based on all the data from a statistical perspective. They can be used together to assist healthcare professionals to decide whether a drug-ADR pair has causal association. Note that even if a signal generated by the data mining method represents true causal association, the relevant ADR is not necessary to occur on a patient who has taken the drug of interest due to the variation of human bodies. Hence, the rule-based algorithm has to be checked when a healthcare professional decides whether a spontaneous report should be filed for a particular patient case. In addition, when there exist few patient cases in which the drug and ADR coexist, the data mining method become less reliable.

The auto-filling feature of the proposed multi-agent system can improve the quality and consistency of the submitted ADR reports. By looking at FDA's spontaneous reports, we found that there existed a lot of typo-errors and some information such as age, gender, drug dose, and so forth are often missing. Since our system can automatically fill out the major portion of the reporting form, these issues can be effectively avoided.

Our system can also be easily adapted to improve other drug safety-related practices in a health organization. For example, it could help physicians make more appropriate choices in drug prescription and avoid preventable ADRs. It could also help make formulary decisions. To ensure quality and cost effectiveness of patient care, it has been a common practice in the past decades that healthcare organizations (e.g., hospitals) establish their formularies [33]. The basic premise is that within a given drug class, there is significant overlap and duplication relative to efficacy and safety. A single drug that is representative of a therapeutic class can meet the needs of the vast majority of patients. Selecting only one drug streamlines inventory, purchasing and maintenance of databases, diminishes the likelihood of medication errors, and creates a competitive environment for pricing. A major barrier to effective formulary decision making is the lack of sufficient published data on the safety of drugs, particularly recently marketed drugs. Our system could provide continuous, active surveillance and timely identification of potential safety issues following the introduction of a new drug to a formulary. Such a system could lead to safer drug use policy, more cost-effective formulary decisions, better healthcare, and earlier detection of ADRs.

## Conclusion

We have developed an agent-based ADR reporting system that can potentially improve the spontaneous reporting rate by providing decision-making assistance and automatically feeding most data into the reporting form. This system will be practically tested in the Frontier Behavioral Health in Spokane, Washington.

## References

1. Anonymous. 1969. International drug monitoring. The role of the hospital. *World Health Organ Tech Rep Ser*. 425, 5-24. PubMed

2.  Bond CA, Raehl CL. 2006. Adverse drug reactions in United States hospitals. *Pharmacotherapy*. 26, 601-08. PubMed http://dx.doi.org/10.1592/phco.26.5.601

3.  Talbot J, Waller P. Stephens' Detection of New Adverse Drug Reactions. 5 ed: John Wiley & Sons; 2004.

4.  Friedman MA, Woodcock J, Lumpkin MM, Shuren JE, Hass AE, et al. 1999. The safety of newly approved medicines: do recent market removals mean there is a problem? *JAMA*. 281, 1728-34. PubMed http://dx.doi.org/10.1001/jama.281.18.1728

5.  Hartmann K, Doser AK, Kuhn M. 1999. Postmarketing safety information: how useful are spontaneous reports? *Pharmacoepidemiol Drug Saf*. 8, S65-71. PubMed http://dx.doi.org/10.1002/(SICI)1099-1557(199904)8:1+<S65::AID-PDS403>3.3.CO;2-V

6.  Hazell L, Shakir SAW. 2006. Under-Reporting of Adverse Drug Reactions - A Systematic Review. *Drug Saf*. 29, 385-96. PubMed http://dx.doi.org/10.2165/00002018-200629050-00003

7. Palleria C, Leporini C, Chimirri S, Marrazzo G, Sacchetta S, Bruno L, et al. Limitations and obstacles of the spontaneous adverse drugs reactions reporting: Two "challenging" case reports. Journal of pharmacology & pharmacotherapeutics. 2013;4:S66-72.

8.  Biriell C, Edwards R. Reasons for Reporting Adverse Drug Reactions - Some Thoughts Based on an International Review. Pharmacology and Drug Safety 1997;21-6.

9.  Lasser KE, Allen PD, Woolhandler SJ, Himmelstein DU, Wolfe SM, et al. 2002. Timing of new black box warnings and withdrawals for prescription medications. *JAMA*. 287, 2215-20. PubMed http://dx.doi.org/10.1001/jama.287.17.2215

10. Eland I, Belton K. 1999. van GA, Meiners A, Rawlins M, Stricker B. Attitudinal survey of voluntary reporting of adverse drug reactions. *Br J Clin Pharmacol*. 48, 623-27. PubMed http://dx.doi.org/10.1046/j.1365-2125.1999.00060.x

11. Molokhia M, Tanna S, Bell D. 2009. Improving reporting of adverse drug reactions: Systematic review. *Clinical Epidemiology*. 1, 75-92. PubMed http://dx.doi.org/10.2147/CLEP.S4775

12. Lopez-Gonzalez E, Herdeiro MT, Figueiras A. Determinants of under-reporting of adverse drug reactions: a systematic review. Drug safety: an international journal of medical toxicology and drug experience. 2009;32:19-31.

13. Jennings NR. 2000. On agent-based software engineering. *Artif Intell*. 117, 277-96. http://dx.doi.org/10.1016/S0004-3702(99)00107-1

14. Huang J, Jennings NR, Fox J. 1995. An agent-based approach to health care management. *Int Journal of Applied Artificial Intelligence*. 9, 401-20. http://dx.doi.org/10.1080/08839519508945482

15. Vázquez-Salceda J, Padget JA, Cortes U, Lopez-Navidad A, Caballero F. 2003. Formalizing an electronic institution for the distribution of human tissues. *Artif Intell Med*. 27, 233-58. PubMed http://dx.doi.org/10.1016/S0933-3657(03)00005-8

17

16. Alsinet T, Ansotegui C, Bejar R, Fernandez C, Manya F. 2003. Automated monitoring of medical protocols: a secure and distributed architecture. *Artif Intell Med*. 27, 367-92. PubMed http://dx.doi.org/10.1016/S0933-3657(03)00010-1

17. Ji Y, Ying H, Farber MS, Yen J, Dews P, et al. 2010. A Distributed, Collaborative Intelligent Agent System Approach for Proactive Postmarketing Drug Safety Surveillance. *IEEE Trans Inf Technol Biomed*. 14, 826-37. PubMed http://dx.doi.org/10.1109/TITB.2009.2037007

18. Mansour A, Ying H, Dews P, Ji Y, Yen J, et al. Finding similar patients in a multi-agent environment. Fuzzy Information Processing Society (NAFIPS), 2011 Annual Meeting of the North American2011. p. 1-6.

19. Naranjo CA, Busto U, Sellers EM, Sandor P, Ruiz I, et al. 1981. A method for estimating the probability of adverse drug reactions. *Clin Pharmacol Ther*. 30, 239-45. PubMed http://dx.doi.org/10.1038/clpt.1981.154

20. Zorych I, Madigan D, Ryan P, Bate A. 2013. Disproportionality methods for pharmacovigilance in longitudinal observational databases. *Stat Methods Med Res*. 22, 39-56. PubMed http://dx.doi.org/10.1177/0962280211403602

21. Norén GN, Hopstadius J, Bate A, Star K, Edwards IR. 2010. Temporal pattern discovery in longitudinal electronic patient records. *Data Min Knowl Discov*. 20, 361-87. http://dx.doi.org/10.1007/s10618-009-0152-3

22. Huidong Jin, Jie Chen, Hongxing He, Kelman C, McAullay D, et al. 2010. Signaling Potential Adverse Drug Reactions from Administrative Health Databases. *IEEE Transactions on Knowledge and Data Engineering*. 22, 839-53. http://dx.doi.org/10.1109/TKDE.2009.212

23. Harpaz R, DuMouchel W, Shah NH, Madigan D, Ryan P, et al. 2012. Novel data-mining methodologies for adverse drug event discovery and analysis. *Clin Pharmacol Ther*. 91, 1010-21. PubMed http://dx.doi.org/10.1038/clpt.2012.50

<jrn>24. Ji Y, Ying H, John T, Dews P, Mansour A, et al. 2013. A Method for Mining Infrequent Causal Associations and Its Application in Finding Adverse Drug Reaction Signal Pairs. *IEEE Transactions on Knowledge and Data Engineering*. 25, 721-33. http://dx.doi.org/10.1109/TKDE.2012.28</jrn>

25. Evans SJ, Waller PC, Davis S. 2001. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. *Pharmacoepidemiol Drug Saf*. 10, 483-86. PubMed http://dx.doi.org/10.1002/pds.677

26. Murach J, Steelman A. murach's Java Servlets and JSP. 2nd ed: Mike Murach & Associates; 2008.

27. TILAB. Java Agent DEvelopment Framework (JADE). 2011.

28. Fleiss JL. Statistical methods for rates and proportions New York: John Wiley & Sons; 1981.

29. Siegel S, Jr. NJC. Nonparametric Statistics for The Behavioral Sciences 2ed: McGraw-Hill Humanities/Social Sciences/Languages; 1988.

30. Landis JR, Koch GG. 1977. The measurement of observer agreement for categorical data. *Biometrics*. 33, 159-74. PubMed http://dx.doi.org/10.2307/2529310

31. Ling Y, An Y, Liu M, Hu X. An error detecting and tagging framework for reducing data entry errors in electronic medical records (EMR) system. Bioinformatics and Biomedicine (BIBM), 2013 IEEE International Conference on2013. p. 249-54.

32. Hauben M, Zhou X. 2003. Quantitative methods in pharmacovigilance: focus on signal detection. *Drug Saf*. 26, 159-86. PubMed http://dx.doi.org/10.2165/00002018-200326030-00003

33. 1986. ASHP statement on the formulary system. *Am J Hosp Pharm*. 43, 2839-41. PubMed