Published in final edited form as:

IEEE Trans Inf Technol Biomed. 2010 May ; 14(3): 826–837. doi:10.1109/TITB.2009.2037007.

A Distributed, Collaborative Intelligent Agent System Approach for Proactive Postmarketing Drug Safety Surveillance

Yanqing Ji,

Department of Electrical and Computer Engineering Gonzaga University, Spokane, WA 99258
USA (ji@gonzaga.edu)

Hao Ying [Senior Member, IEEE],

Department of Electrical and Computer Engineering, Wayne State University, Detroit, MI 48202
USA (hao.ying@wayne.edu)

Margo S. Farber,

Drug Information/Use Policy, Detroit Medical Center, Detroit, MI 48201 USA (mfarber@dmc.org)

John Yen [Fellow, IEEE],

College of Information Sciences and Technology, The Pennsylvania State University, University
Park, PA 16802 USA (jyen@ist.psu.edu)

Peter Dews,

Department of Medicine, Wayne State University, Detroit, MI 48201 USA; Department of
Medicine, St John Health Providence Hospital, Southfield, MI 48075 USA
(peter.dews@stjohn.org)

Richard E. Miller, and

John D. Dingell Veterans Affairs Medical Center, Detroit, MI 48201 USA
(remiller@med.wayne.edu)

R. Michael Massanari

Research for the Critical Junctures Institute, Bellingham, WA 98225 USA
(michael.massanari@wwu.edu)

Abstract

Discovering unknown adverse drug reactions (ADRs) in postmarketing surveillance as early as possible is of great importance. The current approach to postmarketing surveillance primarily relies on spontaneous reporting. It is a passive surveillance system and limited by gross underreporting (<10% reporting rate), latency, and inconsistent reporting. We propose a novel team-based intelligent agent software system approach for proactively monitoring and detecting potential ADRs of interest using electronic patient records. We designed such a system and named it ADRMonitor. The intelligent agents, operating on computers located in different places, are capable of continuously and autonomously collaborating with each other and assisting the human users (e.g., the food and drug administration (FDA), drug safety professionals, and physicians). The agents should enhance current systems and accelerate early ADR identification. To evaluate the performance of the ADRMonitor with respect to the current spontaneous reporting approach, we conducted simulation experiments on identification of ADR signal pairs (i.e., potential links between drugs and apparent adverse reactions) under various conditions. The experiments involved over 275 000 simulated patients created on the basis of more than 1000 real patients treated by the drug cisapride that was on the market for seven years until its withdrawal by the

FDA in 2000 due to serious ADRs. Healthcare professionals utilizing the spontaneous reporting approach and the ADRMonitor were separately simulated by decision-making models derived from a general cognitive decision model called fuzzy recognition-primed decision (RPD) model that we recently developed. The quantitative simulation results show that 1) the number of true ADR signal pairs detected by the ADRMonitor is 6.6 times higher than that by the spontaneous reporting strategy; 2) the ADR detection rate of the ADRMonitor agents with even moderate decision-making skills is five times higher than that of spontaneous reporting; and 3) as the number of patient cases increases, ADRs could be detected significantly earlier by the ADRMonitor.

Keywords

Adverse drug reactions (ADR); fuzzy logic; intelligent agents; postmarketing surveillance; recognition-primed decision (RPD) model

I. Introduction

ADVERSE drug reactions (ADRs) refer to the drug-associated adverse incidents in which drugs are used at an appropriate dose and indication. They can complicate a patient's medical condition or contribute to increased morbidity, even death. Drug-induced morbidity and mortality occurs on a daily basis. The latest data that we can find in the literature show that in year 2000, there were about 100 000 deaths in the U.S. due to medical errors, of which about 7000 were attributed to drug reactions [1]. According to Laser *et al.* [2], between 1975 and 1999, 548 new drugs were approved by the food and drug administration (FDA), 16 (i.e., 2.9%) of which were subsequently withdrawn from the market because of ADRs. Forty-five (8.2%) of the 548 drugs acquired at least one black box warning for an ADR that was not known when the drug was approved by the FDA for marketing (a black box warning is required by the FDA to appear in the drug package insert as well as in the *physician's desk reference* if substantial risk to the patient may occur or if additional information or monitoring of drug use might prevent an adverse event). Laser *et al.* [2] also pointed out that "Many serious ADRs are discovered only after a drug has been on the market for years. Only half of newly discovered serious ADRs are detected and documented in the *Physician's Desk Reference* within 7 years after drug approval."

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 FDA's spontaneous reporting system MedWatch. The limitations of this system are well described [4]. MedWatch is a *passive* system in that it depends on voluntary, spontaneous reports of suspected ADRs to be filed by healthcare professionals, drug manufactures, and/or consumers using the system's online forms. Detection of an ADR generally relies on FDA's retrospective or concurrent review of patient cases. Because ADR reports are filed at the discretion of the users of the system, there is gross underreporting [5], [6]. It was estimated that less than 10% of all ADR cases were reported to MedWatch [2]. Moreover, it depends on human recognition of a potential link between a drug and an apparent adverse reaction (called signal pair), and on the time and will to report the observation [7]. In addition, the rate at which cases are reported is dependent on many factors, including the time period since the drug was released into the market place, pharmacovigilance-related regulatory activity, the indications for use of the drug (which impacts prescribing frequency), and media attention [8]. Finally, the passive surveillance system is limited by latency and inconsistent [9]. Consequently, the current approach may require years to identify and withdraw problematic drugs from the market, and result in unnecessary mortality, morbidity, and cost of healthcare.

Systematic methods for the detection of suspected safety problems from spontaneous reports have been studied and practically implemented [10]. For example, the FDA currently adopts a data mining algorithm called multiitem gamma Poisson shrinker [11] for detecting potential signals from its spontaneous reports. Another important signal detection strategy is known as the Bayesian confidence propagation neural network that has been used by the Uppsala Monitoring Center in routine pharmacovigilance with its World Health Organization database [12]. However, a large number of cases are needed in order for the neural network approach to work. Various other methods, such as proportional reporting ratios [13], empirical Bayes screening [14], reporting odds ratios [15], and incidence rate ratios [16] have been used in the spontaneous reporting centers of other nations (e.g., England and Australian). By utilizing data mining or Bayesian techniques, these methods have shown better performance than traditional methods even though the ability of detecting a signal varies among them [17]. However, the performance of these techniques could be highly situation dependent due to the weaknesses and potential biases inherent in spontaneous reporting [18]. When confounding exists, the signals generated by a data mining algorithm may not represent a true causal-effect relationship between ADRs and the drugs.

In this paper, we present the design of an innovative team-based intelligent agent software system, ADRMonitor, at conceptual level for actively monitoring and detecting signal pairs of ADRs. While current data mining methods for ADR detection rely on the spontaneous reports, our approach utilizes a different source of information—the electronic patient records, representing a new direction of ADR signal generation. As a result, it is possible for the proposed system to detect ADR signal pairs earlier when only a small number of cases exist that can be too few evidence for the data mining methods even to be applicable. Like the other methods, the signal pairs found by our approach will be subject to further analysis (e.g. epidemiology study) and case review and interpretation by expert drug safety professionals experienced in the nuances of pharmacoepidemiology and clinical medicine [19]. To the best of our knowledge, an agent approach to drug safety has never been suggested in the literature.

Intelligent agents are “smart” software programs that act on behalf of human users to find and filter information, negotiate for services, automate complex tasks, and collaborate with other software agents to solve complex problems [20]. Intelligent agents share some common characteristics, including autonomy, collaboration, delegation, and communication skills. A team-based agent system (i.e., multiagent 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” [21]. In this paper, we use the terms “multiagent system” and “team-based agent system” interchangeably. The multiagent technology provides a new paradigm for developing software applications in a variety of fields, such as industry, commerce, entertainment, and military [22], [23]. In the healthcare domain, exemplary applications include the agent architecture for distributed medical care (AADCare) system for integrating the patient management processes [24], the Organ Transplant Management system for the management of organ and tissue transplants among different medical centers [25], and the internal hospital tasks management systems for monitoring the application of medical protocols [26]. To date, the number of biomedical applications is still very limited.

We will present a multiagent system design of the ADRMonitor that covers both the national level (i.e., the FDA) and the healthcare system level. At the current stage, we focus on three important system-level questions concerning the proposed system and will provide quantitative answers via simulation study in this paper. (Now that the simulation outcomes have pointed to the value of ADRMonitor, as a logical step, we are implementing it in a

clinical setting. We will report the evaluation results in the future). The questions are as follows.

1. Is the detection performance of the ADRMonitor better than that of the spontaneous reporting approach? What are quantifiable differences?
2. What should the levels of the agents' decision-making skills be in order for the ADRMonitor to outperform the spontaneous reporting approach at the best reporting rate in practice (i.e., 10%)?
3. How does the number of patient cases affect the detection performance of the ADRMonitor?

The remainder of the paper is organized as follows. Section II introduces our multiagent system design. Section III describes in detail the experimental design and simulation to address these three issues. In Section IV, the simulation results are reported. We wrap up with conclusions in Section V.

II. Design of the Collaborative Team-Based System Adrmonitor

In this section, we introduce the overall architecture and functionality of the ADRMonitor, its agent organization and management, as well as the cognitive decision model used by an agent.

A. ADRMonitor Architecture and Functionalities

1) Overall Architecture of ADRMonitor—The overall structure of the ADRMonitor is shown in Fig. 1. The FDA is at the top of this structure, since it is the drug administration authority. Many healthcare systems are connected with the FDA through the Internet. Note that a healthcare system can be a hospital, a medical center, a sizable clinic, the Medicare program, a health insurance company, or any other organizations that can provide health-related electronic data for identifying potential drug safety signals. Experts in the FDA and healthcare professionals in local healthcare systems have their own intelligent agents, which would collaborate with each other in order to effectively identify unknown ADRs. These intelligent agents are accessible by any computer connected to the Internet. Both the FDA and the healthcare systems have their own databases. The database in the FDA stores the spontaneous reports filed by healthcare professionals, consumers, etc. These reports are valuable for experts at the FDA to conduct population level analyses (e.g., data mining and epidemiological study) as well as case analysis. The databases at each local healthcare system contain detailed information (e.g., pharmacy data and clinical laboratory data) of its patients. The data are important since they provide, among other information, direct or indirect evidence of the clinical state, comorbid conditions, and the progress of the patients [27].

Safety evaluators at the FDA regularly assess the spontaneous reports. Each of them would have an intelligent agent to help him/her perform his/her duty (e.g., requesting details of a particular suspected patient case and/or similar patient cases from local healthcare systems via their intelligent agents). Detecting unknown ADRs is a complex process, often requiring expertise from a variety of specialists (e.g., physicians, pharmacists, chemists, and epidemiologists). At the FDA, each of these experts would have his/her own intelligent agent to assist in performing various analyses (e.g., epidemiologic studies) and gathering necessary information. There can be multiple safety evaluator agents, pharmacist agents, and epidemiologist agents, although only one of each is illustrated in Fig. 1.

2) ADRMonitor Architecture at the Local Level—The detailed agent architecture in a healthcare system is presented in Fig. 2. Note that we use a medical center as an example to

illustrate the agent system architecture. For other types of healthcare systems (e.g., health insurance company), the agent architecture would be similar while the local users and their intelligent agents could be different. In a medical center, there can be multiple safety officer agents and more than one other professional agent, although only one of each is illustrated in Fig. 2. The agent users in such a healthcare system include physicians, drug safety officers/managers, and other healthcare professionals (e.g., pharmacists and nurses). These agents would collect useful information and help the local users identify signal pairs of potential ADRs. There are four different data collection and monitoring agents, each of which corresponds to a specific type of data source for ADR detection (i.e., administration database, pharmacy database, laboratory database, and electronic patient medical records). These agents will receive data collection tasks from other agents, and then, query the corresponding databases for desired data. Furthermore, they continuously monitor their respective databases and proactively provide data to agents who need them. Administrative coding of diagnoses and procedures is a routine part of billing procedures for medical care and is a universally available electronic data source that can be utilized in ADR surveillance. These data are coded in the form of *International Classification of Diseases*, Ninth Edition, Clinical Modification (ICD-9-CM) and physicians' Current Procedural Terminology (CPT), two widely used coding standards. Pharmacy data, clinical laboratory data, and patient medical records are the three other sources of information.

All the professional agents (e.g., the physician agents) use the recognition-primed decision (RPD) model while the other agents, such as the management agents, and data collection and monitoring agents, do not. Since the human professionals are responsible for submitting spontaneous reports, the RPD model would help them assess the strength of causality of signal pairs (i.e., ADRs). Different types of professional agents (e.g., physician agents versus pharmacist agents) have different domain knowledge bases to map the fact that different types of professionals have different expertise and would evaluate signal pairs from different perspectives. (Experts' initial knowledge about a new drug can be obtained by analyzing the results of the premarketing clinical trials and the medical literature, as well as by using their prior ADR experience in the other drugs that are in the same class as the new drug.) For the same type of agents, we assume that different agents use the same cues to assess signal pairs, but they may view the cues differently in order to reflect different levels of skills of different professionals. For example, different physician agents may have different weights on the cues or different definitions of the membership functions for the same fuzzy cue. Note that there may exist inconsistencies among the agents in that different agents have different expertise and different levels of skills. One assumption of this study is that medical expertise and skills are not uniform across agents. This is true for the ADR problem, for example, all the spontaneous reports submitted to the FDA are based on personal recognition and assessment of perceived adverse effects. However, we do believe that collaborations among intelligent agents with complementary expertise would help to make more accurate evaluation of signal pairs. We will explore this in our future work.

We suppose that no conflicts exist in the agent community over resource sharing because the system is not very time critical and a slight ADR detection delay (e.g., in minutes) is acceptable. We believe that all the competing requests from different agents to access the resources can be adequately managed if the system is properly designed.

3) Functionalities of ADRMonitor—With this system design, the safety evaluators at the FDA have at least three ways to discover unknown ADRs, which are complementary to one another. First, they could apply a current data mining algorithm to collected spontaneous ADR reports. Second, they could perform patient-specific ADR reviews by analyzing the details (e.g., temporal relationship between taking a drug and occurrence of an apparent adverse event, concurrent medications, and important laboratory data) of the relevant

individual patient cases collected by their agents. Third, their agents could help them perform epidemiologic analyses at the population level by collecting similar cases distributed across all the healthcare systems, and in association with epidemiologists, assist them to design epidemiologic studies (e.g., case-control or cohort studies). The safety evaluators' intelligent agents could also track further development of the suspected patient cases under the help of the agents in the local healthcare systems. For example, a safety evaluator may be interested in what happens to the patient after the suspected drug is discontinued. In general, the safety evaluators could more easily determine the causal relationship between a potential ADR and a drug.

In the proposed agent architecture, all the intelligent agents at the FDA and different healthcare systems would collaborate with each other and assist the FDA's safety evaluators to more effectively make decisions. For example, after analyzing a spontaneous report concerning an unknown potential ADR, the safety evaluator may need similar cases in order to further analyze whether a true causal relationship exists in this particular drug-ADR pair. Therefore, the safety evaluator could send the information request via his/her intelligent agent to the management agents at different local healthcare systems, which will then forward the request to the corresponding data collection and monitoring agents. These agents would search the local databases and provide similar cases to the evaluator. Furthermore, the data collection and monitoring agents could *continuously* monitor their databases in order to find future similar cases, and proactively forward them to the relevant safety evaluator agents. After receiving the collected information, the safety evaluator will make a further analysis. He/she may need the inputs of other experts at the FDA (e.g., pharmacists and epidemiologists) before deciding what to do with this signal pair. At the local healthcare system level, different intelligent agents can also help each other (e.g., a physician agent may need inputs from a safety officer agent before it files a report).

There are several ways to trigger the collaborative surveillance process of the ADRMonitor at the local healthcare system level or at the national level. First, the system can always monitor some potential adverse effects like death and hospitalization for any new drugs of interest. These adverse effects can be predetermined and entered into the system by a safety evaluator at the FDA or a safety officer in a local healthcare system, which trigger the surveillance process. Second, a safety evaluator or a drug safety officer/manager can anticipate and enter (through his/her software agent) a list of suspected drug-ADR signal pairs based on premarketing clinical trials, and the drug's pharmacodynamic and pharmacokinetic characteristics, which are available at the FDA. Third, a safety evaluator or a safety officer/manager can actively trigger this system at any time when he/she receives reports or information from other sources (e.g., the safety alerts issued by the FDA and new case reports in the literature) that raises concerns about possible ADRs. Fourth, once a physician finds potential adverse effects on his/her patient, he/she can, with the help of his/her agent, search more information or simply file a report to the safety officer or to the FDA, which will trigger the surveillance process for that drug.

In general, intelligent agents at different levels could collaborate with each other in order to monitor different data resources, collect information, make decisions, and share decision-making results. Communications among agents and the networking scalability are two major advantages of the ADRMonitor system. When more healthcare systems are connected to the ADRMonitor, the ability to detect rare ADR signal pairs would be greatly enhanced.

B. Agent Organization and Management

The organization and management of intelligent agents is a very important issue in a large distributed multiagent system. To address this issue, we propose a hybrid agent organization strategy based on our analysis of the system's requirements and constraints (e.g. privacy,

scalability, security, and generality). Within each healthcare system, the unique management agent and many healthcare professional agents (each assisting a unique healthcare professional) take the classic client/server model, which is widely used for distributed applications. In this model, while the server node provides most of the services or capabilities of the system, the client nodes just access and use them. In our case, the management agent works as a server that provides services for all the healthcare professional agents in the same healthcare system. The healthcare professional agents work as clients who can initiate requests. In the FDA, the National Regulatory Authority Agent and all the safety expert agents also use the client/server model.

At the healthcare system level, a client/server model would not be desirable because a healthcare system is expected to be an autonomous peer that can initiate communication and provide capabilities for other healthcare systems. One possible solution is to use the peer-to-peer (P2P) model, where all the peers are treated equally and the service is distributed among all the peers of the network. However, for the pure P2P model, it is difficult to maintain the coherence of the network and discover new peers. Also security is an important issue for the pure P2P model, since each node is allowed to join the network without any control mechanism. We thus take the hybrid model, a model between the client/server model and the pure P2P model. That is, a high-level node (i.e., the National Regulatory Authority Agent) is added to the top of all the healthcare systems and at the same time different healthcare systems can still communicate with each other through their management agents. The National Regulatory Authority Agent has the highest authority from the perspective of agent management in the whole agent framework. For example, all the intelligent agents at the FDA and all the management agents in healthcare systems (see Fig. 2) must be registered and authenticated through this Agent before they can be recognized by other agents in the framework. To avoid being a possible communication bottleneck, the National Regulatory Authority Agent only provides necessary services (e.g., providing name-physical address associations for new comers). Moreover, the caching mechanism can be implemented for each healthcare system so that the name-physical address associations are stored locally and the National Regulatory Authority Agent is contacted only when the address information is missing.

Scalability is another major issue in a large distributed system. To reduce communication overhead, it is reasonable to divide people as well as their supporting agents into different groups or teams. For instance, a group of experts and their agents focusing on a particular ADR can form a (virtual) team. Of course, a human expert and his supporting agent could be a member in multiple teams. The formation of a virtual team would be transparent to the end users. When a safety expert makes a query or a healthcare professional provides supporting information, he/she can be automatically added to a team according to the suspicious drug (or its class) and the suspected adverse outcome. To achieve this, we need to find a strategy of classifying drugs on the market and embed the classification information in our proposed framework. The National Drug Code (NDC) system is chosen as a therapeutic or pharmacological classification scheme in which each NDC, a unique ten-digit number, serves as a universal product identifier for drugs reported to the FDA.

A more concrete and technical description of the agents, their functions and roles, as well as hierarchy formalization in the context of ADR detection can be found in our previous reports [28], [29].

The remaining issues include agent communication, interoperability, and service registry. We have found that the Java agent development environment (JADE [30]), one of the agent system development software packages, can readily handle them. We are using it to implement these aspects of ADRMonitor. 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 also 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.

C. Fuzzy RPD Model

Each intelligent agent of the healthcare professionals and the safety experts is equipped with a cognitive decision-making model to characterize the agent’s decision-making behavior. There exist two different categories of cognitive decision-making models. One is the rational decision-making approach [31], and the other is naturalistic decision making. The RPD model [32] is a popular model of the latter category with a solid theoretical foundation. It is particularly useful when used to model how human experts make decisions based on their prior experiences. The RPD model focuses on situation assessment due to the fact that experts spend most of their time and energy on understanding the decision situation, and that once the situation is recognized, decision making is almost automatic. One study showed that 50% to 80% of all decisions were made in this way [33], while another investigation indicated that 95% of all the decisions made by the naval officers on a cruiser followed the RPD model [34]. We chose to use the RPD model to represent the decision-making behavior of the ADRMonitor’s agents. The original RPD model is descriptive and is not directly implementable on a computer. Hence, we developed a fuzzy RPD model [35], which is not only computational, but also capable of handling vague and subjective information using fuzzy logic [36], [37].

RPD is an experience-based model that employs “situation-experience matching” decision rules to determine, which prior experience can be utilized to solve the current problem. The experiences acquired by solicitation with the physicians on our project team were stored in an experience knowledge base. Fig. 3 is a sample experience illustrated in natural language for easier understanding.

As shown, an experience consists of four components—cues, goals, actions, and expectancies. Cues represent the higher-level information (synthesized from elementary or environmental data) that a decision maker must pay attention to. Expectancies describe what will happen next as the current situation continues to evolve in a changing context. Goals represent an end state that the decision maker is trying to achieve. Actions represent a set of potential decisions that the decision maker can take in the current situation. Cues are used to match the current situation with prior experiences. This sample experience has four cues: temporal association, other explanation, dechallenge, and rechallenge. Temporal association refers to the temporal relationship between taking the drug and occurrence of the adverse event. Other explanations denote alternative explanations by concurrent disease or other drugs. Dechallenge is defined as the relationship between withdrawal of the drug and abatement of the adverse effect. Rechallenge describes the relationship between reintroduction of the drug and recurrence of the adverse event. The type of a cue could be quantitative, nominal, or fuzzy in the proposed computational fuzzy RPD model. For instance, the cue temporal association may have fuzzy values (e.g. unlikely, possible, and likely). The weights for these cues are design parameters and are assigned by domain experts. Table I shows that how the four cues are related to degree of causality of a signal pair and gives four examples.

The experience knowledge base is an integral part of the fuzzy RPD model [38]. It is for the task of assessing suspected ADR signal pairs, and is built through the joint efforts of our

engineering and medical team members after careful analysis of the relevant literature. According to the classification scheme in [39], a particular pattern of cue values characterizes a specific degree of causality, which may require certain courses of action to handle the ADR. Therefore, we can define various experiences, each of which is associated with a degree of causality (e.g., very likely, probable, and possible). These experiences form the experience knowledge base. The number of experiences in the base is related to the complexity of the ADR detection objectives and the base is scalable in that more (or less) experiences can be added (or deleted) when needed. The addition and deletion can be manual, or conceptually automatic.

After representing the experiences, the next step is to extract the cue values from elementary data for the current situation. Fuzzy rules and fuzzy reasoning are used to achieve this task. For example, to obtain the fuzzy value of temporal association, one of the fuzzy rule “if the time duration between taking the drug and the occurrence of the adverse event is short, then temporal association is likely” is used. An embedded fuzzy inference engine is employed to perform fuzzy reasoning. The inference engine is what drives the RPD model, updating cues once new information is detected, and monitoring expectancies and goals. After the cue values of the current situation are known, similarity measures are needed to measure the degree of matching between the current situation and prior experiences. We developed our own similarity measures to identify the experience that was applicable to the current situation in a specific decision-making context. Once the most applicable experience is identified, the action(s) in that experience is used to solve the current problem.

Fuzzy rules and fuzzy reasoning were implemented using the freeware fuzzyjess [40], a Java-based fuzzy inference engine. It allows the user to use Java language to define membership functions, set antecedent, and consequent of a fuzzy rule, and make a fuzzy inference. This made it easier to integrate the reasoning process with the other parts of the RPD model, which were also implemented with Java.

From the perspective of reusing prior experiences/cases to solve current problems, the RPD model is similar to case-based reasoning. However, the RPD is different from case-based reasoning in several aspects: 1) RPD originates from cognitive studies about how human experts make decisions. In some sense, it is a more complete and systematic model developed to specifically describe the cognitive decision-making process; 2) the decision making in RPD is more viewed as a gradual, iterative procedure instead of one critical momentary decision—this feature makes it more appropriately to capture the decision-making process in domains like medicine; and 3) RPD stresses the real-world decision-making settings with time constraints, voluminous information, ill-structured problems, etc. In addition, our integration of RPD and intelligent agents enables constant tracking and monitoring of the decision process. For instance, once new information violates the “expectancies” of the model, prior decisions will be reevaluated. For more details of the fuzzy RPD model as well as concrete examples, the reader is referred to our previous study [35].

The fuzzy RPD model was preliminarily validated by first optimizing it, and then, using the resulting model to calculate the extent of causality between cisapride and some of its adverse effects for 100 simulated patients created based on the profiles of the same 1015 Veterans Affairs (VA) Medical Center patients mentioned earlier. There were 18 model parameters, which included four weights for different cues and 14 parameters to characterize the fuzzy sets used in the fuzzy RPD model. They were optimized using a genetic algorithm so that the model could best mimic the decisions of the two physicians at the same time (i.e., mimic their underlying consensus). We tried different settings of the genetic algorithm so as to obtain an optimal result within a modest computing time. The genetic algorithm’s (GA)

elitism rate and cullage rate were set at 0.05 and 0.4, respectively. A two-point crossover and a mutation with mutation rate of 0.01 were applied to generate new members.

The model's validity was then established by comparing the decisions made by the optimized model and those by two independent experienced internists for the 100 simulated patients. The levels of agreements were measured by the weighted Kappa statistic, which is an estimate of agreement between two raters after chance agreement is controlled. Kappa statistic ranges from 1 (complete agreement) to 0 (totally disagreement). As suggested by Landis and Koch [41], agreement is considered to be excellent when the weighted Kappa statistic is over 0.75, fair to good when it is between 0.4 and 0.75, and poor when it is less than 0.4. In our case, the Kappa statistics for physicians 1 and 2 versus the genetic-algorithm-optimized model were 0.939 and 0.7, respectively, while the statistic for physician 1 versus physician 2 was 0.657. The results suggested good to excellent agreements [35].

III. Simulation Study

Reporting rate of the ADRMonitor is effectively 100% because it continuously monitors the data sources and the agents collaborate consistently to detect potential ADRs. Conceptually, it should have the potential to outperform the current spontaneous reporting approach given that the intelligent agents could actively track patient cases, alert physicians regarding potential new ADRs, and provide more detailed information to safety officers. We will quantify the performance through computer simulation. Another important issue is as follows. From the perspective of artificial intelligence (AI), it is very difficult to equip each of the intelligent agents with the same high level of decision-making skills that human experts possess. The individual agents' decision-making skills will likely be inferior to experienced healthcare professionals. Logically, in order for the ADRMonitor to be superior to the current system, it is essential that the agents collectively have better signal pair detection performance than experienced human experts who fail to report over 90% potential ADR cases (i.e., less than 10% reporting rate). We now explore the answer to this question using simulation. Last but not least, how does the increase of patient cases affect the ADR detection performance of the ADRMonitor? Conceptually speaking, ADRs are difficult to detect when the number of patient cases is small. If more physician groups, and in turn more patient cases, are included in the multiagent network, the rate of detection will be greatly accelerated. To explore this issue quantitatively, we will simulate and compare ADRMonitor' detection process when different numbers of physician groups are connected.

In all of the simulations concerning these three issues, we assume that agents collaborate efficiently and that they can access the desired information instantaneously. A similar assumption is applied to the spontaneous reporting approach—the healthcare professionals have all the necessary information handy for filing a report. Note that, in this paper, we focus on evaluating the performance of a group of *physician agents*. Other professional agents are not evaluated because at this stage our primary goal is to design a multiagent framework (i.e., ADR-Monitor) and establish its superiority in detecting drug-ADR pairs relative to the spontaneous reporting approach using simulation. Physician agents are chosen because physicians hold the key to ADR detection and are reporting many more (possible) ADRs than other healthcare professionals. For example, after studying all those spontaneous reports submitted to the FDA by different healthcare professionals in 2008, we found that physicians reported 57.5% of the total cases, while pharmacists and other professionals reported 11.4% and 31.1% of the total cases, respectively.

A. Real Patient Data as a Basis for Simulation Investigation

The drug cisapride was introduced into the marketplace in 1993 upon approval by the FDA for the treatment of gastroesophageal reflux. Seven years later the drug was removed from

the marketplace because of idiosyncratic, high-risk adverse reactions. With approvals from the Human Investigation Committees of Wayne State University (WSU) and the VA Medical Center in Detroit, we retrieved all the patients treated with cisapride at the VA Medical Center between 1993 and 2000. The sources for the data included the standard hospital discharge abstract database and the pharmacy database. A statistically deidentified dataset was produced for analysis. A total of 1015 patients were found that had received cisapride on one or more encounters in the institution. The total number of prescriptions was 2803. We found that the number of apparent signal pairs of potential ADRs (i.e., cisapride and Torsades de Pointe, and/or syncope) in the patient data is very limited. Therefore, we used the real patients to create simulated patient cases in order to conduct the simulation.

B. Creation of Simulated Patient Cases

To create simulated patient cases, we utilized the real patient cases to identify the important parameters that would provide useful information for weighing potential causal linkages among signal pairs. The determination of the value range for each parameter was based on the analysis of real patient data by the experienced physicians on the project team. We generated simulated patient cases by assigning the parameters with random values within the ranges. The case generation was accomplished in a way that simulated the observed distribution of cisapride prescriptions over time, but at a rate that was increased 100-fold (i.e., 280 300). The much higher number of cases was needed to represent the number of prescriptions that would be ordered by physicians in 100 medical centers individually comparable to our VA Medical Center. In our simulation study, we assumed that these medical centers were connected by the ADRMonitor through computer networks. This assumption was appropriate and reasonable because in reality it is difficult to detect ADR signal pairs when the number of cases is small as the occurrence rate of most ADRs is low. As a result, we obtained a total of 275 400 distinctive simulated patient cases. Based on our literature research and the opinions of our expert physicians and biostatisticians, the categories of signal pair strength for these simulated cases and their distribution were estimated as: very likely (0.01% of the total cases), probable (0.02%), possible (0.02%), unlikely (0.02%), and no pair (99.93%). For the purpose of our simulation study, a signal pair that was labeled or evaluated as “very likely” was treated as an ADR. Thus, the ADR occurrence rate was 0.01% (i.e., total 27 cases) among the simulated cases. All the simulated patients were stored in a microsoft access database.

C. Simulated Physicians and Simulated ADRMonitor Agents

1) Simulated Physicians—The optimized fuzzy RPD model introduced in Section II-C was treated as a gold standard in this paper. We used it to create 200 different simulated physicians in the following manner (we employed 200 because the number of internal medicine and family medicine physicians in the WSU outpatient physician group was about 200, which was considered to be representative for major healthcare centers in the nation). We randomly generated 50 000 sets of the parameters of the fuzzy RPD model, all within reasonable ranges. The model with a specific set of parameters represented a particular physician. Different parameter sets represent different physicians with various levels of experience and decision-making skills. In order to select 200 physicians, whose decision-making skills have a reasonable distribution, each simulated physician evaluated the 100 simulated patient cases, and its agreement with the optimized model, as measured by weighted Kappa statistic, was calculated. The skills of the 200 simulated physicians were selected based on their Kappa statistics in a way that the number of physicians in each Kappa category would roughly follow a sort of normal distribution: 15% (0–0.4), 70% (0.4–0.75), and 15% (0.75–1.0).

To include physicians' potential errors occurring in the process of evaluating patient cases, we assumed that patient cases with no signal pairs could be mistakenly classified as ADR signal pairs in a random fashion at a rate of 0.001%.

To address the three questions raised in the Section I, one way would be to use 100 physician groups of the same size (i.e., $200 \times 100 = 20\,000$ physicians), since the number of simulated patient cases was comparable to that treated by that many physicians in 100 medical centers. However, we found the computing time required to evaluate all the 275 400 simulated patients prohibitively long on our 2.0 GHz PC. Alternatively, we randomly assigned the patients in an indirect way. We selected only 2754 from the 275 400 patient cases to form a database, and only these 2754 cases were used in the simulation. In selection, we kept all the cases in the categories of "very likely," "probable," "possible," and "unlikely" in the 275 400 cases and randomly picked up some cases in the category of "no pair." Hence, the distribution of the new patient pool was: very likely (1%), probable (2%), possible (2%), unlikely (2%), and no pair (93%). The total number of "no pair" patients excluded was 272 646 (i.e., $2\,757\,400 \times 99.93\% - 2754 \times 93\%$). Nevertheless, the effect that these excluded "no pair" patients would exert on the signal pair detection was still reflected in the simulation study. It was represented via the 0.001% physician random error rate. The number of excluded "no pair" patients that could be mistakenly evaluated as ADR signal pairs by the simulated physicians is 2.7 cases (i.e., $272\,646 \times 0.001\%$), which were included in the simulation. By such an arrangement, simulations involving the 2754 simulated patients evaluated by the 200 simulated physicians would yield results comparable to those achieved by directly evaluating the 275 400 simulated patients by 20 000 physicians.

The same setting was applied to the ADRMonitor agents.

D. Simulated ADRMonitor Agents

When creating simulated ADRMonitor agents, we supposed that 1) the dedicated ADRMonitor agent of a real physician possessed a comparable decision-making skill as the physician did; and 2) the agent could be set up to substitute the physician for making decisions. These assumptions are based on the expectation that an agent can gradually and continuously learn a physician's decision-making skills by adjusting the parameters that model its decision behaviors. That is, every time the physician makes a decision, the agent would compare its own decision with the physician's decision and accordingly adjust its parameters using a learning algorithm. When the agent's decision-making skills are close enough to the physician's decision-making skills, the physician would be confident to let the agent substitute him/her for making decisions. These assumptions led us to use the same 200 RPD models representing 200 physicians as 200 ADRMonitor agents (the random human error component was removed as the agent was a computer program that would not make such a mistake). These 200 agents were utilized when studying the first and third questions mentioned earlier. For the second question, different agents were necessary because we needed different groups of agents, each of which contained 200 agents with the same level of decision-making skills. We realized this goal by choosing those RPD models whose Kappa statistics were the same with respect to the optimized RPD model.

E. Simulation Settings

We compared the performance of the 200 simulated physicians with that of the 200 simulated ADRMonitor agents (there were multiple groups of them) in assessing the strength of the signal pairs between cisapride and Torsades de Pointe, and/or syncope. The decisions made by the simulated physicians and the simulated ADRMonitor agents were compared with those of the optimized RPD model, which was the gold standard (recall that

this optimized model extracted and represented the consensus of two experienced physicians).

During the simulation, the simulated patient cases were randomly retrieved one by one from the database. The simulated physicians as well as the simulated ADRMonitor agents were randomly assigned with the cases. A simulated physician or agent could be selected by chance for more than one time for different patient cases. This implies that some simulated physicians or agents would never be picked up due to the randomness. The simulated physician first checked whether a drug–ADR pair of interest existed in the patient case. If it exist, the physician then assessed the strength of association of that pair. A randomly selected ADRMonitor agent then performed the same task on the case. Regardless which of the three issues raised in Section I was studied, all the signal strengths determined by the agents would always be used (i.e., 100% reporting rate). On the other hand, whether the assessment generated by a physician would be included depended on reporting rate that we set.

To study the first issue, we set the simulated physicians' reporting rate at various levels between 1% and 10% in order to investigate how the reporting rates affect the ability of detecting ADR signal pairs (keep in mind that 10% is at the high end of spontaneous reporting rates estimated in the literature). When the rate was set at, say 10%, it means that the signal strengths of 90% patient cases were randomly ignored by the simulated physicians for reporting. To explore the second issue, we created seven groups of 200 ADRMonitor agents, using the RPD models with seven Kappa statistics levels from 0.4 to 1 with increment of 0.1. As stated in aforementioned Section III-C, while the agents with the Kappa statistic value of 1 have excellent decision-making skills, and the agents with the value of 0.4 have poor decision-making skills compared with the gold standard (the optimized RPD model). To address the last question, the same 200 simulated ADRMonitor agents with different decision-making skills created previously were used. Four additional patient pools in different sizes were created. The sizes were proportional to the number of physician groups that were assumed to work together. For example, one patient pool was for the situation when ten physician groups was supposed to work together, which is 10% of the 100 physician groups assumed in the study of the first two issues aforementioned. Then only 10% of the patients was selected to the pool.

The ADRMonitor system and the spontaneous reporting approach were compared in terms of ADR detection rate, false positive rate, and false negative rate. Because of the random nature of selecting a simulated physician or an ADRMonitor agent, we ran each simulation experiment ten times and used the averages as results. The simulations were implemented using the Java programming language. In addition, several existing Java packages (e.g., fuzzyjess [40]) were utilized in order to handle fuzzy logic. The implementation of the genetic algorithm was based on a Java open-source package [42]. We expanded the package to meet some specific requirements of our simulations.

IV. Simulation Results

Fig. 4 shows the comparison of the total numbers of accumulated potential ADR signal pairs, including false positive ADR signal pairs, as detected by the simulated ADRMonitor system and the spontaneous reporting approach with 10% reporting rate between 1993 and 2000. It is for the study of the first question mentioned in Section I. According to the figure, no potential ADR signal pairs were detected during the first 26 months after the drug became available because only a small number of patients had taken the drug. Owing to the 100% reporting rate, the ADRMonitor was able to detect many more potential ADR signal pairs than the current spontaneous reporting strategy could do best (i.e., at the highest

reporting rate in practice). Fig. 5 gives the comparison of the accumulated *true* ADR signal pairs (by our definition given earlier) detected by the optimized RPD model (gold standard), the ADRMonitor, and the spontaneous reporting. It reveals that 20.6 out of 27 ADR signal pairs were detected by the ADRMonitor, leading to a detection rate of 76.3%. In contrast, only 2.9 out of the 27 pairs were detected by the spontaneous reporting approach, resulting in a detection rate of merely 10.7%. That is, the number of true ADR signal pairs detected by the ADRMonitor was 6.6 times higher than that by the spontaneous reporting strategy. Obviously, the ADR signal pairs could be detected earlier using the multiagent system approach.

Table II presents the detailed detection performance of the 200 simulated physicians under different reporting rates from 1% to 10%. It shows how the increase of reporting rate affects the detection performance. For example, if the reporting rate increases from 3% to 7%, the detection rate will increase from 1.5% to 4.8%. This suggests that ADR signal pairs could be detected earlier by human experts with the help of the ADRMonitor system if the experts would take advantages of the ADRMonitor (e.g., intelligence, proactivity, and real-time surveillance) and file more spontaneous reports.

Now, let us examine the simulation results concerning the second issue. Table III lists the detailed detection performance when 200 simulated ADRMonitor agents with same Kappa statistics were used to evaluate the simulated patient cases. The experiment was repeated under the seven different levels of Kappa statistics. Note that the agents with lower Kappa statistics have worse decision-making skills as compared with the optimized RPD model (the gold standard). The table shows that the detection rate rises, and at the same time the false positive rate and false negative rate decrease when the Kappa statistics level increases. Comparing Table II with Table III, one sees that even if agents' Kappa statistic is only 0.4, its detection rate (64.4%) is still much higher than the 200 physicians with 10% reporting rate can achieve (10.7%). That is, the ADR detection rate of the ADRMonitor agents with even moderate decision-making skills was five times higher than that of spontaneous reporting. Understandably, the false positive rate (61.3%) and false negative rate (21.3%) of the ADRMonitor are somewhat worse than the respective 46.3% and 16.7% of the simulated physicians. A logical conclusion is that even when the ADRMonitor's individual agents' decision-making skill is moderate only, their collective ability in identifying potential ADR signal pairs can still be better than the spontaneous reporting strategy. The underlying reason for the advantage is the equivalent 100% reporting rate of the ADRMonitor agents.

The simulation results covering the third issue are depicted in Fig. 6, where it shows the accumulated average true ADR signal pairs detected when different numbers of physician groups are supposedly linked via the computer networks by the ADRMonitor system. As expected, the more the physician groups are connected, the earlier the ADR signal pairs are detected. The value of Fig. 6 is that it quantifies this expectation. For instance, suppose that to make a decision on a drug (e.g., withdrawal or assignment of a black box warning), at least five true ADRs would need to be detected. Then, not enough evidence could be collected for making the decision in the time period of interest if only ten physician groups are connected. However, if 30 groups would work together, sufficient information for the decision would be collected 65 months after the drug is used. If 70 groups collaborate instead, a decision would be reached 15 months earlier.

V. Conclusion

We have developed a collaborative, team-agent framework that aims to help safety experts in the FDA and healthcare professionals in healthcare systems to achieve better postmarketing surveillance and earlier detection of potential ADRs. In this framework, a

group of collaborative agents would effectively search and track relevant patient information, provide alerts on significant or unexpected adverse reactions, and interact with drug safety experts (e.g., drug safety evaluators). Using the fuzzy logic-based computational RPD model, emulating different decision-making skills of physicians and ADRMonitor agents, our simulation study explored the system's detection performances. Our quantitative simulation results show that 1) the agent framework has the potential to identify possible ADR signal pairs more effectively than the current spontaneous reporting approach; 2) even intelligent agents with moderate decision-making skills can collectively detect more ADR signal pairs and accelerate the reporting process because their reporting rate approaches always 100%; and 3) the more physicians that are linked with and use the ADRMonitor, the earlier the pertinent ADR signal pairs can be detected. Although a particular drug (i.e., cisapride) is used, our approach is generally applicable to detection of the ADR signal pairs caused by any other drugs. This is because the cues (e.g., temporal association, rechallenge, and dechallenge) are universal in detecting ADR signals regardless of drugs. Having established the advantages of ADRMonitor over the spontaneous reporting approach quantitatively, we are now in the early stage of implementing ADRMonitor in a clinical setting. We will continue to use the knowledge-based approach (e.g., the fuzzy RPD model and the fuzzy expert system technique) to detect ADR signal pairs. The advantages of our approach over other AI approaches (e.g., neural networks) include: 1) explicit and effective utilization of expert knowledge in the form of if-then rules; and 2) the reasoning chain between the cause and the effect is always available for human to examine and understand. Issues that cannot be properly investigated in simulation are being addressed. They include design of adequate graphical user interface, improvement of ADR detection rate, and minimization of the false positive detection rate. The overall goal is to enhance the system's clinical utility to ensure end user acceptance. We will report the results when they become available in our future publications.

Acknowledgments

The authors would like to thank Dr. J. Ager for his help in statistics and E. Floyd, J. Small, C. Grace, and R. Johnson for their help in retrieving the patient information.

This work was supported in part by the National Institutes of Health under Grant 1 R21 GM082821-01A1.

Biographies

Yanqing Ji received the B.Eng. degree in industrial automation from Qingdao University, Qingdao, China, in 1997, the M.Sc. degree in physical electronics from the University of Science and Technology of China, Hefei, China, in 2001, and the Ph.D. degree in computer engineering from Wayne State University, Detroit, MI, in 2007.

He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, Gonzaga University, Spokane, WA. His research interests include multiagent systems, parallel and distributed systems, data mining, and their biomedical applications.

Dr. Ji is a member of the Program Committees of various international conferences and workshops including the 2009 International Workshop on Graphics Processing Unit Technologies and Applications, and reviewer for various international journals.

Hao Ying (S'88–M'90–SM'97) received the B.S. degree in electrical engineering, the M.S. degree in computer engineering from Donghua University, Shanghai, China, in 1982 and 1984, respectively, and the Ph.D. degree in biomedical engineering from The University of Alabama, Birmingham, in 1990.

He is currently a Professor with the Department of Electrical and Computer Engineering, Wayne State University, Detroit, MI, where he is a Full Member with the Barbara Ann Karmanos Cancer Institute. During 1992–2000, he was the Faculty of The University of Texas Medical Branch, Galveston. During 1998–2000, he was an Adjunct Associate Professor of the Biomedical Engineering Program, The University of Texas at Austin, Austin. He has authored or coauthored one research monograph/advanced textbook entitled *Fuzzy Control and Modeling: Analytical Foundations and Applications* (IEEE Press, 2000), 92 peer-reviewed journal papers, and more than 120 conference papers. He was invited to serve as a Reviewer for more than 60 international journals.

Prof. Ying is an Associate Editor or a member of Editorial Board for nine international journals. He is a member of the Fuzzy Systems Technical Committee of the IEEE Computational Intelligence Society and chairs its Task Force on Competitions. He is an Elected Board Member of the North American Fuzzy Information Processing Society (NAFIPS). He was a Program Chair for The 2005 NAFIPS Conference as well as for The International Joint Conference of NAFIPS Conference, the Industrial Fuzzy Control and Intelligent System Conference, and the NASA Joint Technology Workshop on Neural Networks and Fuzzy Logic, in 1994. He was a Publication Chair for the 2000 IEEE International Conference on Fuzzy Systems and was a member of Program Committee for more than 35 international conferences.

Margo S. Farber, photograph and biography not available at the time of publication.

John Yen, photograph and biography not available at the time of publication.

Peter Dews, photograph and biography not available at the time of publication.

Richard E. Miller, photograph and biography not available at the time of publication.

R. Michael Massanari, photograph and biography not available at the time of publication.

References

- [1]. Kohn, LT.; Corrigan, JM.; Donaldson, MS. *To Error is Human: Building a Safer Health System*. National Academy Press; Washington, DC: 1999.

- [2]. Lasser KE, Allen PD, Woolhandler SJ, Himmelstein DU, Wolfe SM, Bor DH. Timing of new black box warnings and withdrawals for prescription medications. *J. Amer. Med. Assoc.* 2002; 287:2215–2220.
- [3]. Talbot, J.; Waller, P. Stephens' Detection of New Adverse Drug Reactions. 5th ed. Wiley; New York: 2004.
- [4]. Klein DF. The flawed basis for FDA post-marketing safety decisions: The example of anti-depressants and children. *Neuropsychopharmacology*. Apr.2006 31:689–99. [PubMed: 16395296]
- [5]. Hartmann K, Doser AK, Kuhn M. Postmarketing safety information: How useful are spontaneous reports? *Pharmacoepidemiol. Drug Safety*. Apr.1999 8:65–71.
- [6]. Hazell L, Shakir SAW. Under-reporting of adverse drug reactions—A systematic review. *Drug Safety*. 2006; 29:385–396. [PubMed: 16689555]
- [7]. Granas AG, Buajordet M, Stenberg-Nilsen H, Harg P, Horn AM. Pharmacists' attitudes towards the reporting of suspected adverse drug reactions in Norway. *Pharmacoepidemiol. Drug Safety*. Apr.2007 16:429, 34.
- [8]. Goldman SA. Limitations and strengths of spontaneous reports data. *Clin. Ther.* 1998; 20:C40–C44. [PubMed: 9915089]
- [9]. Biriell C, Edwards R. Reasons for reporting adverse drug reactions—Some thoughts based on an international review. *Pharmacol. Drug Safety*. 1997; 6:21–26.
- [10]. Bennett CL, Nebeker JR, Yarnold PR, Tigue CC, Dorr DA, McKoy JM, Edwards BJ, Hurdle JF, West DP, Lau DT, Angelotta C, Weitzman SA, Belknap SM, Djulbegovic B, Tallman MS, Kuzel TM, Benson AB, Evens A, Trifilio SM, Courtney DM, Raisch DW. Evaluation of serious adverse drug reactions: A proactive pharmacovigilance program (RADAR) vs safety activities conducted by the food and drug administration and pharmaceutical manufacturers. *Arch. Intern. Med.* May 28.2007 167:1041–1049. [PubMed: 17533207]
- [11]. Szarfman A, Tonning JM, Doraiswamy PM. Pharmacovigilance in the 21st century: New systematic tools for an old problem. *Pharmacotherapy*. Sep.2004 24:1099–104. [PubMed: 15460169]
- [12]. Lindquist M, Edwards IR, Bate A, Fucik H, Nunes AM, Stahl M. From association to alert—A revised approach to international signal analysis. *Pharmacoepidemiol. Drug Safety*. 1999; 1:15–25.
- [13]. Evans SJ, Waller PC, Davis S. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. *Pharmacoepidemiol. Drug Safety*. Oct-Nov; 2001 10:483–486.
- [14]. DuMouchel W. Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. *Amer. Stat.* 1999; 53:177–190.
- [15]. Purcell P, Barty S. Statistical techniques for signal generation: The Australian experience. *Drug Safety*. 2002; 25:415–421. [PubMed: 12071778]
- [16]. Heeley E, Wilton LV, Shakir SA. Automated signal generation in prescription-event monitoring. *Drug Safety*. 2002; 25:423–432. [PubMed: 12071779]
- [17]. Kubota K, Koide D, Hirai T. Comparison of data mining methodologies using Japanese spontaneous reports. *Pharmacoepidemiol. Drug Safety*. Jun.2004 13:387–394.
- [18]. Hauben M. Early postmarketing drug safety surveillance: Data mining points to consider. *Ann. Pharmacother.* Oct.2004 38:1625–1630. [PubMed: 15304626]
- [19]. Hauben M, Zhou X. Quantitative methods in pharmacovigilance: Focus on signal detection. *Drug Safety*. 2003; 26:159–186. [PubMed: 12580646]
- [20]. Ferber, J. Multi-Agent Systems: An Introduction to Distributed Artificial Intelligence. Addison-Wesley; New York: 1999.
- [21]. Jennings NR. On agent-based software engineering. *Artif. Intell.* 2000; 117:277–296.
- [22]. Parunak, HVD. Applications of distributed artificial intelligence in industry. In: O'Hare, G.; Jennings, N., editors. *Foundations of Distributed Artificial Intelligence*. Wiley; New York: 1996. p. 139-164.
- [23]. Yen J, Fan X, Sun S, Hanratty T, Dumer J. Agents with shared mental models for enhancing team decision makings. *J. Decis. Support Syst.* 2006; 41:634–653.

- [24]. Huang J, Jennings NR, Fox J. *Int. J. Appl. Artif. Intell.* 1995; 9:401–420.
- [25]. Vazquez-Salceda J, Padget JA, Cortes U, Lopez-Navidad A, Caballero F. Formalizing an electronic institution for the distribution of human tissues. *Artif. Intell. Med. Mar.2003* 27:233–58. [PubMed: 12667738]
- [26]. Alsinet T, Ansotegui C, Bejar R, Fernandez C, Manyà F. Automated monitoring of medical protocols: A secure and distributed architecture. *Artif. Intell. Med. Mar.2003* 27:367–92. [PubMed: 12667743]
- [27]. Bates DW, Evans RS, Murff H, Stetson PD, Pizziferri L, Hripcsak G. Detecting adverse events using information technology. *J. Amer. Med. Inf. Assoc. Mar-Apr;2003* 10:115–28.
- [28]. Ji, Y.; Ying, H.; Yen, J.; Fan, X.; Massanari, RM.; Barth-Jones, DC. Team-based multi-agent system for early detection of adverse drug reactions in postmarketing surveillance; *Proc. 24th North Amer. Fuzzy Inf. Process. Soc.; Ann Arbor, MI. 2005*; p. 644–649.
- [29]. Ji Y, Ying H, Yen J, Zhu S, Barth-Jones DC, Miller RE, Massanari RM. A distributed adverse drug reaction detection system using intelligent agents with fuzzy recognition-primed decision model. *Int. J. Intell. Syst.* 2007; 22:827–845.
- [30]. TILAB (Telecom Italia Lab). Java Agent DEvelopment Framework (JADE). TILAB; Torino, Italy: Jul.. 2004 Available: <http://jade.tilab.com/>
- [31]. Janis, IL.; Mann, L. *Decision Making: A Psychological Analysis of Conflict, Choice, and Commitment.* Free Press; New York: 1977.
- [32]. Klein, GA. *Decision Making In Action: Models and Methods.* Ablex Publishing; Norwood, NJ: 1993. A recognition-primed decision making model of rapid decision making; p. 138–147.
- [33]. Klein GA. Strategies of decision making. *Military Rev.* 1989; 69:56–64.
- [34]. Kaempf GL, Klein G, Thorsden ML, Wolf S. Decision making in complex naval command-and-control environments. *Human Factors.* 1996; 38:220–231.
- [35]. Ji Y, Massanari RM, Ager J, Yen J, Miller RE, Ying H. A fuzzy logic-based computational recognition-primed decision model. *Inf. Sci.* 2007; 177:4338–4353.
- [36]. Zadeh LA. Fuzzy sets. *Inf. Control.* 1965; 8:338–353.
- [37]. Ying, H. *Fuzzy Control and Modeling: Analytical Foundations and Applications.* IEEE Press; Piscataway, NJ: 2000.
- [38]. Ji Y, Ying H, Yen J, Massanari RM. A fuzzy logic-based computational recognition-primed decision model. *Inf. Sci.* 2007; 177:4338–4353.
- [39]. Edwards IR, Aronson JK. Adverse drug reactions: Definitions, diagnosis, and management. *Lancet.* Oct.2000 356:1255–1259. [PubMed: 11072960]
- [40]. Orchard, R. Fuzzy reasoning in jess: The FuzzyJ toolkit and fuzzyjess; *Proc. 3rd Int. Conf. Enterprise Inf. Syst.; Setubal, Portugal. 2001*; p. 533–542.
- [41]. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* Mar.1977 33:159–174. [PubMed: 843571]
- [42]. Faupel, M. GAJIT—Genetic Algorithm Java Implementation Toolkit. 2003. Available: <http://www.micropraxis.com/gajit/index.html>

Fig. 1.
Overall structure of the ADRMonitor.

Fig. 2.
ADRMonitor architecture at the healthcare system level.

Experience_N: The causality between the drug and a particular ADR is *very likely*

- Cues:** 1) The occurrence of ADR has a *likely* temporal association with the drug;
 2) There is *no* other explanation;
 3) The ADR has *likely* dechallenge relationship with the drug;
 4) The ADR has *likely* rechallenge relationship with the drug.

Goals: Find the strength of the causality between a drug and a particular ADR;

- Actions:** 1) Suggest for further analytical studies;
 2) Suggest filing an ADR report online.

- Expectancies:** 1) More similar cases are available;
 2) Relevant information for filing an ADR report is available;

Fig. 3.
 Sample experience.

Fig. 4.

Accumulated average potential ADR signal pairs detected by the ADRMonitor and by the spontaneous reporting approach (month 1 and month 85 correspond to Jul. 1993 and Jul. 2000, respectively—the time period that the drug cisapride was on the market).

Fig. 5. Accumulated average “true” cisapride ADR signal pairs detected by the optimized RPD model (gold standard), the ADRMonitor, and the spontaneous reporting approach (the x -axis is the same as that in Fig. 4).

Fig. 6.

Accumulated average true cisapride ADR signal pairs detected by the ADRMonitor under different numbers of connected physician groups (the x -axis is the same as that in Fig. 4).

TABLE I

Relating Cues to Degree of Causality of a Signal Pair—Four Examples

Cues	Cue value set 1	Cue value set 2	Cue value set 3	Cue value set 4
Temporal association	Likely	Likely	Possible	Unlikely
Other explanations	No	No	No	Yes
Dechallenge	Likely	Likely	Possible	Unlikely
Rechallenge	Likely	Possible	Unlikely	Unlikely
Degree of causality	Very likely	Probable	Possible	Unlikely

TABLE II

Averaged Results of ADR Signal Pairs Detected by the 200 Simulated Physicians Under Different Reporting Rates (The Total Number of True ADR Signal Pairs is 27)

Reporting rate	Average Detected Potential ADR Signal Pairs	Average Detected True ADR Signal Pairs	Average Detection Rate	Average False Positive (Rate)	Average False Negative (Rate)
1%	0.3	0.1	0.4%	0.2 (66.7%)	0.1 (33.3%)
3%	0.9	0.4	1.5%	0.5 (55.6%)	0.3 (33.3%)
5%	1.9	0.8	3.0%	1.1 (57.9%)	0.4 (21.1%)
7%	2.8	1.3	4.8%	1.5 (53.6%)	0.6 (21.4%)
10%	5.4	2.9	10.7%	2.5 (46.3%)	0.9 (16.7%)

TABLE III

Averaged Results of ADR Signal Pairs Detected by Different Groups of 200 Simulated Agents With the Same Kappa Statistics (The Total Number of True ADR Signal Pairs is 27)

Kappa statistic	Average Detected Potential ADR Signal Pairs	Average Detected True ADR Signal Pairs	Average Detection Rate	Average False Positive (Rate)	Average False Negative (Rate)
0.4	45	17.4	64.4%	27.6 (61.3%)	9.6 (21.3%)
0.5	42	18.3	67.8%	23.7 (56.4%)	8.7 (20.7%)
0.6	37.9	19.9	73.7%	18 (47.5%)	7.1 (18.7%)
0.7	37.6	21.8	80.7%	15.8 (42.0%)	5.2 (13.8%)
0.8	35.8	22.6	83.7%	13.2 (36.9%)	4.4 (12.3%)
0.9	32.8	24	88.9%	8.8 (26.8%)	3.0 (9.1%)
1.0	28.2	26.6	98.7%	1.6 (5.7%)	0.4 (1.4%)