Scientific Research Publishing

Journal of Software Engineering and Applications, 2023, 16, 605-621
https://www.scirp.org/journal/jsea
ISSN Online: 1945-3124
ISSN Print: 1945-3116

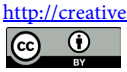
A Personalized Adverse Drug Reaction Early Warning Method Based on Contextual Ontology and Rules Learning

Haixia Zheng[1*], Wei Wei[2]

[1]School of Chinese-German Institute for Applied Engineering, Zhejiang University of Science and Technology, Hangzhou, China
[2]School of Statistics and Mathematics, Zhongnan University of Economics and Law, Wuhan, China
Email: *77618365@qq.com, 503175355@qq.com

How to cite this paper: Zheng, H.X. and Wei, W. (2023) A Personalized Adverse Drug Reaction Early Warning Method Based on Contextual Ontology and Rules Learning. *Journal of Software Engineering and Applications*, **16**, 605-621.
https://doi.org/10.4236/jsea.2023.1611031

Received: September 21, 2023
Accepted: November 25, 2023
Published: November 28, 2023



Abstract

Background: The fatality of adverse drug reactions (ADR) has become one of the major causes of the non-natural disease deaths globally, with the issue of drug safety emerging as a common topic of concern. **Objective:** The personalized ADR early warning method, based on contextual ontology and rule learning, proposed in this study aims to provide a reference method for personalized health and medical information services. **Methods:** First, the patient data is formalized, and the user contextual ontology is constructed, reflecting the characteristics of the patient population. The concept of ontology rule learning is then proposed, which is to mine the rules contained in the data set through machine learning to improve the efficiency and scientificity of ontology rule generation. Based on the contextual ontology of ADR, the high-level context information is identified and predicted by means of reasoning, so the occurrence of the specific adverse reaction in patients from different populations is extracted. **Results:** Finally, using diabetes drugs as an example, contextual information is identified and predicted through reasoning, to mine the occurrence of specific adverse reactions in different patient populations, and realize personalized medication decision-making and early warning of ADR.

Keywords

Health Information Services, Personalized, Contextual Ontology, Drug Adverse Reaction, Early Warning, Reasoning

1. Introduction

The increase of adverse drug reactions (ADR) occurrences globally and their

consequent fatality has led to raising the issue of drug safety as a common concern topic. According to Goetz and Schork [1], many population differences in drug treatment and effects exist. Biomedical research has shown that the genetic factor of individuals is the main reason for ADR. Differences in individuals such as gender, age, race, and different diseases have an effect on the dynamic treatment process of medicines, and produce different ADRs. Moreover, genetic mutation is the source of individual differences in drug reactions. Pharmacogenomics focuses only on the different medication reactions to different diseases caused by genetic mutations [2]. Therefore, in the process of promoting precise medical treatment, full consideration of patient genetic factors, gender, age, weight, physiological and pathological features, as well as the other medications should be taken. Based on this, a safe, rational, effective and economical drug treatment plan is formulated, and personalized medication is administered to prevent or reduce ADR.

Many pharmaceutical companies and medical institutions in western countries have taken the lead in personalized drug development, personalized diagnosis and treatment services, while China's personalized health care services are gradually emerging. Relevant organizations and functional departments are actively promoting the development of personalized medication in China. At present, few studies focus on the monitoring of ADR in individualized pharmaceuticals. To advocate smart medicine, information technology can support the prediction, early warning, and prevention mechanisms of ADR.

In this research, a contextual ontology was developed that reflected the characteristics of patient population differences and used the ontology rule learning to find the patterns of ADR in patients with different population characteristics, and transform those patterns to Semantic Web Rule Language (SWRL) rules that can achieve early warning of personalized ADR.

The remainder of this paper is organized as follows: previous research on context-based health information service, context modeling and semantic reasoning is reviewed in Section 2. The research framework for this study is presented in Section 3. The experiments for early warning of personalized ADR, including contextual ontology construction, ontology rule learning and inference are described in Section 4. Finally, Section 5 concludes the paper and presents future work.

2. Related Work

2.1. Adverse Drug Reaction Mining in Health Social Media

The usual pharmacy alert system is a report from doctors and pharmacists, not a direct report from the patient. Due to the filtering of doctors and pharmacists, reports sent to the pharmacovigilance system are not the first source of data. With the advent of Web 2.0, social media has accumulated a large amount of user information resources. Users are willing to share and discuss their experiences with the use of drugs on social media about health topics. The emergence of

these social platforms provides a new way to monitor ADR. Social media commonly used in ADR research include Twitter and Facebook, and online health communities such as DailyStrength and Medhelp. These data sources are the first-hand data from a wide range of drug users. Compared with traditional spontaneous event reporting systems and electronic health records, social media data has the characteristics of extensive resources, rich information, and real-time generation [3]. However, at the same time, the unstructured social media data increases the difficulty of mining. Nevertheless, the rapid development of natural language processing technology and machine learning methods have made the use of social media data to mine ADR possible. Freifeld CC *et al.* compared the use of social media data with spontaneous reporting system data for adverse reaction mining, extracted potential ADR from millions of Twitter texts, and compared the results with those in FAERS [4]. The article concluded that the distribution of adverse reactions in the two data sources was basically consistent. Pharmaceutical companies are also very interested in these immediate and direct reports of ADR from patients, because they can enable them to identify problems early and take measures to avoid being subjected to serious legal proceedings and loss of profits [5].

Healthcare research should make full use of these rich information resources. The rapid emergence of user-generated content (UGC) has become an important asset for continuous monitoring of public health resources and adverse disease events [6]. Analyzing the narrative content of patients on social media is very important for assessing their perceived risk of ADR and exploring the relationship between drugs and adverse reactions [7]. In summary, ADR reported by patients can contribute greatly to reliable pharmacovigilance [8].

2.2. Context-Based Health Information Service

Context refers to information that can be used to describe entities. An entity can be regarded as a conceptual or physical object, including people, places and calculation objects [9]. Context can be defined as the relatively static user attributes, the environment they belong to, and the current state. Context awareness can predict user behavior and need through the acquired context information, thus providing precision services. Context acquisition mainly collects user and external environment context information via diverse types of sensors.

In the field of health information services, with the development of e-health, health information services based on context technology can provide personalized and adaptive health information services to patients, transforming healthcare into user-centered self-care, home care, and mobile care. According to Liang and Li, research in this field mainly focuses on five aspects, namely electronic medical records, reminder and early warning, disease diagnosis, information recommendation and epidemiological prediction [10].

2.3. Context Modeling and Semantic Reasoning

The context based method involves two technical issues: context modeling and

context reasoning. Context modeling is the conceptual abstraction of context information, while context reasoning is the identification or prediction of high-level context through user low-level context information. These two issues are the key to system implementation based on context awareness.

Context modeling has realized the mapping of context information from the physical to virtual space. Many methods for context modeling exist. Owing to the advantages of concept description, the ability of expression and reasoning support, the ontology-based context modeling method has become the mainstream modeling method. In addition to providing modeling of contextual knowledge, ontology can also support more reasoning by adding rules description. Using ontology to model health context information can express the connection between concepts, and promote knowledge sharing and reusing [11].

Semantic reasoning supports the broader reasoning by adding rule descriptions based on semantic information organizing [12]. The ontology contains rich semantic information and can be regarded as the form of information organizing (*i.e.*, context modeling method). Through reasoning, the specific form of knowledge set in the ontology can be obtained and used to solve practical problems [13]. Commonly used inference mechanisms can be divided into ontology-based reasoning, rule-based reasoning and machine learning-based reasoning. These inference mechanisms have their corresponding advantages and disadvantages. The purpose of ontology-based reasoning is to mine the implicit knowledge in display definitions and declarations through a processing mechanism. Generally, the rules of ontology reasoning are user-defined, and are based on description logic, while machine learning and data mining algorithms can mine the rules contained in the data set to improve the efficiency of rule generations. Combining the advantages of different inference algorithms to improve the efficiency and accuracy of reasoning is a significant challenge of semantic reasoning.

Health information services should utilize semantic web technology to enable high-level mapping and reasoning of low-level context to support intelligent decision-making through ontology modeling of user context and semantic annotation of resources, thus providing personalized health care services.

3. Methods

A personalized ADR early warning framework that integrates into the characteristics of patient population is presented in this paper. First, the semantic enhancement method was used to formalize patient data, and user contextual ontology that reflected the characteristics of the patient population was constructed. Then based on the ADR contextual ontology, the contextual information was identified and predicted through reasoning, to mine the occurrence of different adverse reactions in different patient groups. This implementation of personalized ADR early warning explored the application of "smart medicine", by using machine learning methods to mine data from social media.

In the stage of rule generation of reasoning, the concept of rule learning proposed is the following: mining healthy social media through machine learning methods to generate custom rules for ontology reasoning. Machine learning methods were integrated into the generation of reasoning rules, to improve the efficiency and intelligence of reasoning. Figure 1 presents the early warning framework for personalized ADR, which contains four parts, namely data preprocessing, user contextual ontology modeling, ontology rule learning and personalized medication early warning.

This paper combines statistic-based association rule mining with ontology-based semantic reasoning to establish the early warning model. The empirical research was carried out through the semantic reasoning of ADR in diabetes.

3.1. Data Preprocessing

Personalized ADR early warning data can be divided into structured data for constructing the contextual ontology and unstructured data for the rule learning stage. Context ontology modeling is defined by domain experts and knowledge engineers, who cooperatively define the scope of ontology, and collect required drugs and their adverse reaction data from authoritative databases such as MedDRA, UMLS and SIDER.

The data preprocessing in the rule learning stage cleans and regulates the required data according to the requirements of the machine learning task. After preprocessing operations, such as word segmentation and removal of non-English characters for the data mined from the online community, the data is converted into the format of the transaction database. At the same time, incomplete or meaningless data is eliminated, while the original continuous data can be replaced by discrete attributes.

3.2. Contextual Ontology Modeling

Contextual ontology can be divided into user ontology, environment ontology and state ontology [14]. User contextual ontology was constructed in this study

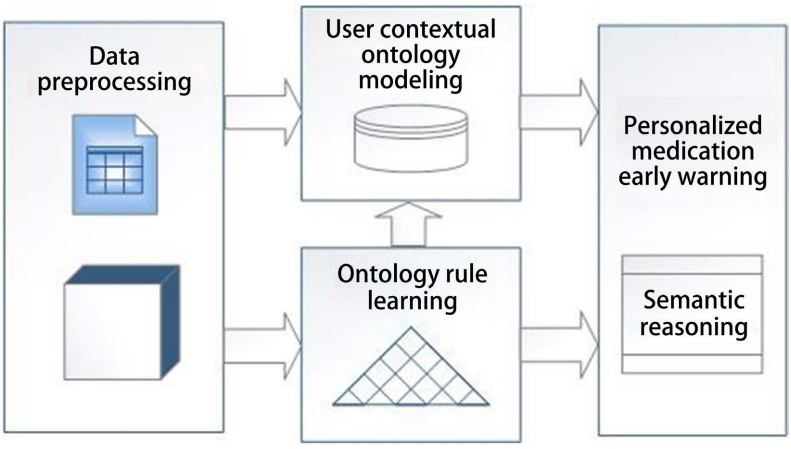


Figure 1. The early warning framework for personalized ADR.

according to the static attribute characteristics of users. "User" is the core concept in the context of context-aware computing. The "user" in this study refers to different medication groups, that is, patients.

Let the contextual ontology be $O = <O_s, O_d, R>$, where, O_s is the model layer, O_d is the data layer, and R is the correspondence between the two [15].

1) Mode layer $O_s = <C_s, P_s, E_s>$, where, C_s represents the concept set of related entities in the field such as gender, age, weight, physiological and pathological characteristics and other drugs being taken. P_s represents the attribute set of the relationship, including the parent relationship, equality relationship, causality relationship, and user-defined related properties used to describe the concept, such as the domain attribute of the drug concept. E_s represents the set of relations between concept nodes, $E_S \in C_s \times P_s \times C_s$.

2) The data layer $O_d = <C_d, P_d, E_d>$ is the instantiation of the mode layer, where, C_d represents a collection of instances (C_i) or a collection of literal values. P_d represents the attribute collection, including object attribute collection and numerical attribute collection; E_d represents the relationship collection between the data layer nodes.

3) The relationship between O_s and O_d is $R = \{$(instance, rdf: type, class$|$ instance $\in C_i$, class $\in C_s\}$.

To increase the accuracy of personalized drug information recommendation for different patient populations and reduce the incidence of ADR, inference is necessary based on user basic attribute characteristics and physiological parameters to determine the probability of adverse reactions caused by taking certain drugs.

This study used the basic attributes and physical parameters of the patient to describe the user context. The contextual factors considered included gender, age, weight, physiological and pathological characteristics, and other drugs being taken. In addition, genetic differences are the root cause of adverse reactions. Selecting therapeutic drugs based on genes can improve their effectiveness and avoid adverse reactions. Therefore, genetic testing is an important part of user's context.

When designing the model layer, the six upper-level core concepts of gender, age, weight, physiological and pathological characteristics, other drugs being taken, and pharmacogenomic testing were selected to define the contextual ontology.

Gender: Women are often excluded from clinical research during drug development, and compared with men, hormone levels in women, drug use propensity, gender differences in pharmacological effects, and personality differences make women more prone to ADR. The different mechanisms of absorption, distribution, metabolism, and excretion of drugs in men and women lead to gender differences in pharmacodynamic responses [16].

Age: Age is also an important factor affecting the efficacy of drugs. The effects of certain drugs in the older population, adolescents, and children are different

from those of adults. For example, many anesthetics can achieve the same effect by injecting fewer doses into the older people [17].

Pathophysiological features: The pathological state of different diseases will affect the function of each body, which in turn will influence the effect of drugs. For example, when liver and kidney function decline, the effects of certain drugs may be enhanced and even lead to poisoning. Physiological changes in the body may also lead to adverse reactions. For example, there are circadian rhythm fluctuations in the secretion of gastric acid, cholesterol synthesis, and asthma attacks in physiological changes. According to the pathological characteristics of different diseases and the physiological laws of the human body, choosing the best time to take medicines to synchronize with the biological rhythm of the human body, can achieve the best effect and reduce the occurrence of adverse reactions [18].

Other medicines being taken: The treatment of many diseases in the clinic is combined with drugs, and the interaction between drugs brings benefits or harm to the body. Drug interaction refers to the change in the action time or the strength of the drug, after taking other drugs sequentially or at the same time. The consequences may be expected, insignificant, or even harmful. The interaction between different drugs and drugs and food may also induce changes in pharmacokinetics or pharmacodynamics, causing adverse reactions [19].

Gene test: Pharmacogenomics reveals the differences in drug response to different diseases caused by genetic mutations, and clarifies the relationship between genetic differences and drug efficacy and safety. A DNA sample is procured from the patient by scraping the exfoliated cells of the oral mucosa, to obtain their genetic sequence. By comparing with known disease gene samples, it is possible to identify the susceptibility genes of the disease in the patient, and infer the risk of the patient suffering from a certain disease. Thus, the drug treatment plan can be selected according to the genetic characteristics and the harm from adverse reactions also can be avoided [20].

Other characteristics: Many drugs are contraindicated, if the patient is in pregnancy, lactation, menstrual period, etc. Patients such as these should be cautious in taking drugs to avoid drug damage. The difference in the nutritional status of patients can also lead to different reactions after taking certain drugs, such as patients with muscular dystrophy taking verapamil will suffer from respiratory arrest. In addition, the characteristics of the patient population such as blood type and weight still need to be researched.

The contextual ontology of this study modeled and instantiated the key information of the above user attributes, and determined the above ontology class to describe and explain the characteristics of the patient population.

3.3. Ontology Rule Learning

Although both resource description frameworks (RDFs) and Web ontology language (OWL) in the ontology have well-defined formal semantics and support

reasoning, they only describe the data used for reasoning. Due to the lack of rules for reasoning, it is impossible to determine the authenticity of a proposition only through data. The Semantic Web needs to add rule descriptions on the basis of data to support a wider range of reasoning. The rules for reasoning in ontology reasoning usually summarize the rules of the field by analyzing the relationships between entities, concepts and concepts, concepts and attributes, and instances in the real world. These rules are usually custom description logic based on experience, facts, etc.

ADR involve multiple dimensions including drugs, adverse reactions and drug users. The differences in drug use population, the complexity of adverse reactions, and the discontinuity of attribute values make conventional early warning methods unable to fully satisfy the demand of personalized health information services.

This study proposes the idea of rule learning, that is, mining social media data using machine learning methods to generate rules for ontology reasoning, thus breaking the limitations of previous reasoning rules based on custom logic descriptions, such as experience and facts. In this study, the association rule mining method was used as an example to analyze the association between the patient and the adverse drug information, and to find the association relationship between the adverse drug reaction and the gender, age, physiological and pathological characteristics and other multi-dimensional contextual attributes of the patient. In this way, rules for ontology reasoning of ADR, which can provide a reference method for personalized medication early warning, are generated.

When the Apriori algorithm searches for frequent sets layer by layer, the entire database needs to be scanned once for each frequent k-item set. In order to improve the efficiency of association rule mining, the data cube-based multi-dimensional association rule mining algorithm was used to organize the data in the data warehouse into the form of data cubes, and perform association rule mining on the data cube to reduce the item set as much as possible. In the Apriori-cube algorithm [21], the search object is a frequent predicate set. For example, the multi-dimensional association rule: $Age(x, ">40") \wedge hasDisease(x, "hypertension") \Rightarrow susceptibletoDisease(x, "diabetes")$, *the predicate set* {*Age*, *hasDisease*, *susceptibletoDisease*} *is a 3-predicate set*.

According to the cube-based multidimensional association rule mining algorithm, the implementation steps of the ADR ontology rule learning based on user context include: 1) The generation of ADR data cubes; 2) Mining frequent predicate sets that meet the minimum support on the generated data cube; 3) Generate ADR early warning rules for different drug populations in the frequent predicate sets.

The importance degree of the early warning rule is determined by the degree of support, and the accuracy of the early warning rule is determined by the degree of confidence. After further analysis and rule description, the early warning rules for ADR ontology were extracted, adding to the adverse reaction ontology

rule base for reasoning and medication warning. Simultaneously, through the dynamic mining of continuously updated training data sets, the update of the early warning rule base for ADR was achieved. It is clear though that other machine learning methods can also be used to extract more potential rules applicable to the requirements of ontology reasoning.

3.4. Personalized Medication Early Warning

The ontology reasoning model was used in this study to create the early warning mechanism. The powerful semantic expression and reasoning ability of the ontology was used to provide personalized information services for the early warning of ADR. **Figure 2** presents the ontology reasoning model, which contains three parts: the input of medical information, ontology reasoning model detection, and the output of early warning information.

The early warning of personalized drug adverse reactions is divided into two stages: the training stage, that is, the ontology rule learning, and the early warning stage, that is, the ontological reasoning and personalized medication early warning. In the training phase, the personal data of the drug users and the historical data of adverse reactions were trained, and the ADR rules for different drug population characteristics were mined. Then the rules were described and written into the ontology for ontology reasoning. In the early warning stage, patient characteristic data and drug information were input into the ontology as instances, combined with the knowledge and facts in the ADR ontology. The ontology reasoning engine searched for matching rules in the rule library for reasoning, predicted the probability of adverse reactions in the user, and provided early warning according to the set threshold.

The early warning platform can provide early warning according to the different characteristics of the patients who use the drug and possible adverse reactions. Input any dimension in the three-dimensional combination, such as patient age, if the ontology data contains the drug and adverse reaction rules for that age group, and the pre-term, post-item, rule support and warning information will be displayed in the output results. Similarly, when inputting the drug information, if the ontology contains different population characteristics of the drug and possible adverse reaction rules, the warning information will also be likewise displayed.

4. Experiment Analysis

4.1. ADR Contextual Ontology Construction

Diabetes drugs were used in the experiment as the example. According to UMLS, SIDER, ADReCS (ADR classification system), it is possible to design the diabetes ADR ontology, with the help from experts in the fields of medicine and pharmacy. **Table 1** shows important concepts and terms involved in the diabetes ADR contextual ontology. The hierarchical structure of the contextual ontology contains four major classes, namely *Drugs used in diabetes* (diabetes drugs classes),

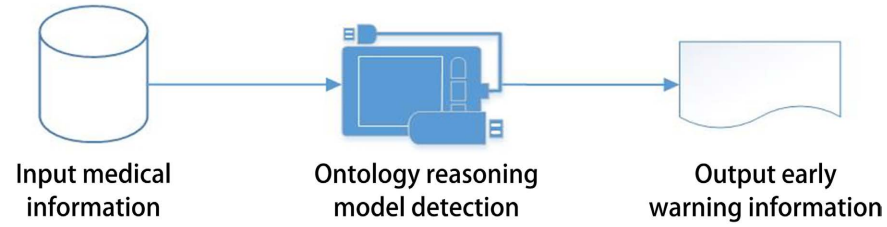


Figure 2. The ontology reasoning model.

Table 1. Important terms in the diabetes ADR ontology.

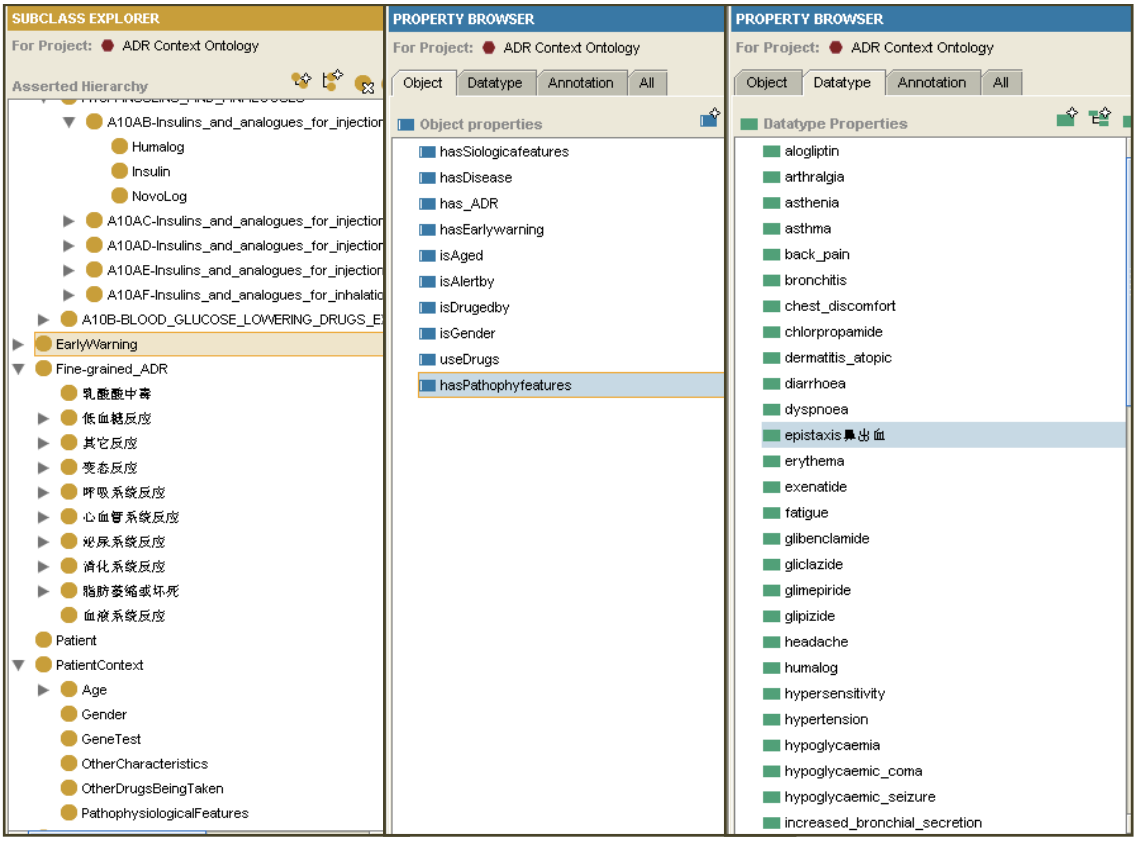
Drugs used in diabetes	Insulin, Metformin, Chlorpropamide, Glibenclamide, Gliclazide, Glimepiride, Glipizide, Tolazamide, Tolbutamide, Acarbose, Miglitol, Pioglitazone, Rosiglitazone, Troglitazone, Alogliptin, Saxagliptin, Sitagliptin, Vildagliptin, Exenatide, Liraglutide, Nateglinide, Repaglinide, Symlin, Humalog, NovoLog…
Fine-grained adverse drug reactions	Hypoglycemia reaction, allergic reaction, lactic acidosis, cardiovascular system, hematological system, alimentary system, urinary system, adipose atrophy or ecrosis, Arthralgia, Asthenia, Back pain, Chest discomfort, Fatigue, Headache, Injection site reaction, Musculoskeletal discomfort, Oedema peripheral, Swelling, Dermatitis atopic, Nausea, Diarrhoea…

Fine-grained ADR (fine-grained ADR classes), *UsertContext* (patient context class), and *Patient* (user classes). Each major class contains corresponding subclasses.

The contextual ontology refines the key information of user attributes and determines the contextual class to describe and explain the characteristics of different groups of patients. Based on the ontology of ADR that have been built in the pre-study work, user contextual characteristics including *Gender*, *Age*, *GeneTest*, *Pathophysiological*, and *other drugs being taken* were used to achieve ontology modeling. The important attributes involved in the diabetes ADR ontology include object attributes, such as *hasADR* (with adverse reactions), *useDrugs* (drugs taken), *hasContext* (having characteristic), and data type attributes corresponding to various drugs and adverse reactions.

This study used Protégé for ontology visualization. Specifically, OWL classes were used to describe the concept of knowledge; Properties were used to describe the attributes and relationships of classes connected to the basic data type, and this relationship can be connected to another class or instance; Individuals were used to describe instances and classes, attributes, domain and range were used to describe data. Figure 3 presents part of hierarchical structure of class and attributes of ADR ontology that was generated by Protégé.

Patient information can be added as instances under the patient class, and the associations with other classes are established through object attributes, which are subsequently used for reasoning and personalized drug adverse reaction detection.



(a) (b) (c)

Figure 3. The classes and attributes of diabetes ADR ontology. (a) Some classes, (b) Object attributes, (c) Data attributes.

Finally, the integrity and consistency of the ADR ontology can be tested by using the Pellet inference engine. The construction of ontology is a complex and iterative process that requires the cooperation of domain experts and knowledge engineers. The continuous improvement of domain knowledge and continuous supplement of details are important in the process of expansion of the ontology. The ontology constructed in this study was only used as an experimental model to verify the effectiveness of subsequent early warning reasoning.

4.2. ADR Ontology Rule Learning

This experiment used Microsoft SQL Server 7.0 and DBMiner 2.0 for data mining. Because of the limited pharmacogenomic data and professional knowledge about pharmacology and pathology, only two attributes (*i.e.*, gender and age) were selected in the user context for rule learning and early warning experiments to prove the effectiveness of the proposed method.

4.2.1. Data Preprocessing

This experimental data was extracted from the diabetes section of the *Ask a patient* website, and "BLOOD GLUCOSE REGULATORS" was selected as drugs for association rule mining. The hypoglycemic drug was classified by its name,

and patients were able to score and comment on the drug by their experience. The comment interface of the drug *ACTOS* is shown in Figure 4.

We chose 14 types of hypoglycemic drugs with more than 20 comments by users as the experimental object. The web crawler tool was employed to capture 13,207 comment records from the website and information about patients including user ID, gender, age and adverse reactions. The raw data was then processed and stored in a relational database. The processing task involved segmentation, removal of non-English characters and stop-words. Additionally, cleaning and filtering of noise, vacancies and inconsistencies in the data was performed, eliminating incomplete and meaningless data, and replacing the original continuity data with discrete attributes.

OLAP was used to generate a three-dimensional ADR data cube. The three dimensions were: drug dimension, patient attribute dimension and adverse reaction dimension. Table 2 presents the details of the data cube. The drug dimension "*blood sugar regulating drugs*" was chosen and capital letters were used to represent different drug names, such as A for *ACTOS*, B for *GLUCOPHAGE*, and C for *AVANDIA*. The patient attribute dimension contained patient gender, age, other drugs being taken, genome detection, etc. Two dimensions of patient attributes were selected, namely gender and age. In the gender dimension, 0 represented male and 1 represented female; in the age dimension, discrete variables of age range were used: 0 represented under 30, 1 represented 30 - 39, 2 represented 40 - 49, 3 represented 50 - 59, 4 represented 60 - 69, and 5 represented older than 70. The adverse reaction dimension involved many variables, selecting 20 common adverse reactions of blood glucose regulating drugs.

4.2.2. Results and Analysis

The Apriori_Cube algorithm was used to generate a frequent predicate set. The minimum support degree min_sup was set to 0.25, resulting in a minimum support count (min_count) of 25. Figure 5 presents the relationship between

RATING	REASON	SIDE EFFECTS FOR ACTOS	COMMENTS	SEX	AGE	DURATION/ DOSAGE	DATE ADDED
▼▲				F M	▼▲	▼▲	▼▲
4	Diabetes Type II	Severe muscle soreness especially in feet, lower legs, arms and breasts. My breasts are so tender that my poor husband has given up any chance of enjoying them anymore - can't stand even wearing a bra. Tingling and numbness in feet and fingers - shoes hurt after just a few hours so I wear sandals. Shortness of breath, heaviness of chest. Extreme blurred vision both with and without glasses on (light prescription). Always so tired even though I excersize and eat right - I take a nap almost every day. BG came down and was within normal ranges when using ACTOS, but I'm too tired to care.		F	46	3 months	9/9/2006
1	Diabetes	Shortness of breath, Rapid heart beat, TIRED, extremely tired. fainting. Increase in Glucose levels	Should be taken off the market. This is a very dangerous drug.	M		6 months	9/7/2006
2	poorly controlled diabetes	terrible edema in legs hands feet, thighs bloating in belly and increase pulse rate, tingling in legs and feet, mood swings, shortness of breath low sugar weak episodes and sweet gravings.	I have liver problems and I am afraid something will cause me increase problems as long as I stay on it. Yes it decreased the AIC but I think the	F	49	3 months	8/22/2006 Email

Figure 4. The user comment interface on the Ask a patient website.

Table 2. Dimensions of ADR data cubes.

Dimension	contents	
Drug dimension	A-ACTOS B-GLUCOPHAGE C-AVANDIA …..	
Patient attribute dimension	Gender	0-male 1-female
	Age	0-Under 30 1-30 - 39 ….
Adverse reaction dimension	1-Weight gain 2-Oedema 3-Fatigue ……	

后项	前项		前项	支持度
a_adrName=Weight increase	a_sex=0	30.52%	a_AgeRange=0	1.12
a_adrName=Weight increase	a_sex=1	27.02%	a_AgeRange=1	1.67
			a_AgeRange=2	19.05
			a_AgeRange=3	20.1
			a_AgeRange=4	17.86
			a_AgeRange=5	17.58
a_adrName=Oedema	a_sex=0	4.95%	a_AgeRange=1	1.54
a_adrName=Oedema	a_sex=1	12.32%	a_AgeRange=2	3.32
			a_AgeRange=3	4.77
			a_AgeRange=4	4.21
			a_AgeRange=5	3.43
a_adrName=Dyspnoea	a_sex=0	7.44%	a_AgeRange=2	1.48
a_adrName=Dyspnoea	a_sex=1	7.01%	a_AgeRange=3	3.56
			a_AgeRange=4	5.24
			a_AgeRange=5	4.17
a_adrName=Somnolence	a_sex=0	6.46%	a_AgeRange=0	1.22
a_adrName=Somnolence	a_sex=1	6.34%	a_AgeRange=2	3.35

Figure 5. Partial results of rules mining.

patient gender and age and the adverse reaction when patients take *ACTOS* by mining association rules from 14 blood glucose regulating drugs.

The results showed that hypoglycemic drugs were prone to hypoglycemia, allergic reaction, diarrhea, oedema, nausea, headache, somnolence (drowsiness), fatigue and other adverse reactions, when treating type II diabetes, among which hypoglycemia was the most common adverse reaction. Compared with young people, middle-aged and older people are more likely to suffer from this adverse reaction. There is literature that confirms the mining results from this study. Another result from the experiment when taking ACTOS showed that female patients are more prone to suffer from oedema than male patients, which is consistent with the drug instruction. Moreover, female patients with "arthralgia" adverse reactions were more than male patients (5.30%: 3.48%), which is con-

firmed by the report that "female patients taking ACTOS are more likely to have fractures" [22].

It is possible to obtain six groups of rules with strong association and three groups of rules with certain association However, not all rules are useful. Although insignificant associations are eliminated by setting minimum support and confidence, rules are not useful still remain. Future research needs to improve the mining algorithm in order to achieve the better results for this issue.

4.3. Personalized ADR Early Warning Based on Ontology Reasoning

SWRL was used to describe rules, combined with OWL ontology, and can infer the semantics that OWL ontology cannot display. Based on the previous conclusions of analyzing the association rules on 14 kinds of blood glucose regulating drugs, six SWRL warning rules were written into the ontology rule library. A rule example is as follows:

User(*?x*)^*hasGenderof*(*?x*, *female*)^*Drugs_used_in_Diabetes*(*?y*)^*isAlertby*(*?y*, *female*)→*hasEarlywarning*(*?x*, *Oedema*)

The inference system models used in the experiment included description logic-based reasoning and SWRL-based reasoning. In the inference engines based on description Logic, the FaCT++ inference engine using the Tableau algorithm was integrated into Protégé, because of its simplicity, ease of use and high efficiency in reasoning, and the FaCT++ inference engine was used in the experiment for description logic reasoning. In the inference engine based on rule reasoning, JESS was selected as the reasoning engine for early warning in view of its good portability, embedding and high efficiency. Figure 6 shows the operation

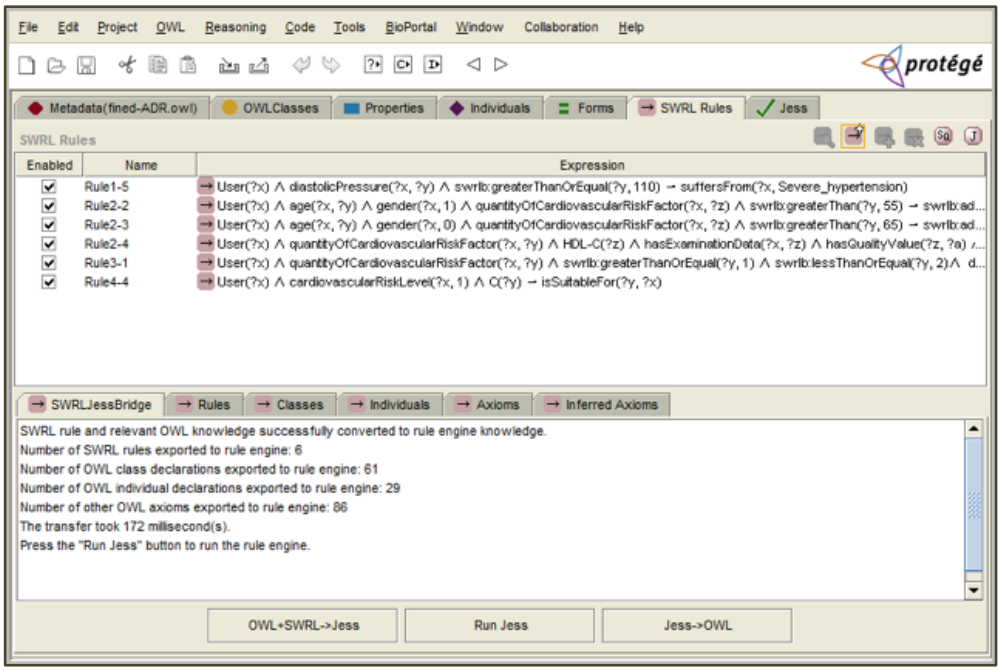


Figure 6. The inference results of OWL + SWRL → JESS.

interface of SWRL. Starting the JESS engine through the "J" button in the upper right corner of the interface, the result shows that SWRL selected six rules and 61 classes and added nine new Axioms, which are the results of JESS reasoning, coming to new conclusions.

The platform can realize the early warning of any dimension (drug, patient characteristics, possible adverse reactions, *et al.*), by editing the results of association rule mining into early warning rules into the ontology for reasoning, then constituting an early warning platform for adverse drug reactions. For example, when the gender information is input into the system as female, a warning rule will appear: *Drugs* = *ACTOS*, *Gender* = *female* → *Reaction* = *Oedema*, which shows that the female patient that takes ACTOS is likely to have the oedema reaction. The degree of support and confidence of this rule reveals the likelihood and accuracy of the reaction. Similarly, when the drug information is input, the warning information also shows that female patients are prone to develop oedema.

The traditional diabetes treatment method gradually adjusts the medication plan through the response of the patient after taking the medication for a period. This approach virtually increases the risk of treatment. The ADR early-warning scheme based on user context can provide an early warning of drug risks through the analysis and reasoning of patient and drug information before medication, and provide a reference method for personalized health and medical information services.

5. Conclusion and Future Research

This study proposes a personalized adverse drug reaction (ADR) early warning method based on user contextual ontology and rule learning. First, the semantic enhancement method was used to formalize the patient data, and the ontology modeling of the user context reflecting the characteristics of the patient population was designed. Then, based on the ADR contextual ontology, the contextual information was identified and predicted through reasoning, and a personalized ADR early warning model was established to achieve personalized medication decisions. Finally, empirical analysis was performed through the reasoning of the ADR in diabetes.

Nevertheless, some limitations are present in this study. The gene-directed personalized medications could better reflect the essence of personalized health care. Limited to the acquisition of pharmacogenomic data and knowledge, genomic detection was not included in the study of user context attributes. However, the method proposed in this study can be extended to gene-directed personalized drug research to develop a safer and more effective personalized treatment plan for patients. In addition, in the ontology rule learning process, further research needs to focus on the application of more advanced machine learning and data mining methods to improve the efficiency and accuracy of ontology inference rules generation, so that the mining results satisfy our re-

quirements.

Ethics Approval and Consent to Participate

Not applicable.

Consent for Publication

Not applicable.

Availability of Data and Material

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable requests.

Competing Interests

The authors declare no competing interests.

Authors' Contributions

Wei Wei was responsible for methodology and implementation. Liqiong Liu was responsible for formal analysis and experiments.

References

[1] Goetz, L.H. and Schork, N.J. (2018) Personalized Medicine: Motivation, Challenges, and Progress. *Fertility & Sterility*, **109**, 952-963. https://doi.org/10.1016/j.fertnstert.2018.05.006

[2] Ye, B.C. (2000) Pharmacogenomics Design of Personalized Medicines. *Journal of Chinese Biotechnology*, **20**, 53-57.

[3] Yu, Y. (2016) Adverse Drug Reaction Knowledge Integration and Application Research Based on Big Data Mining. PhD Dissertation, Jilin University, Changchun.

[4] Freifeld, C.C., Brownstein, J.S., Menone, C.M., *et al.* (2014) Digital Drug Safety Surveillance: Monitoring Pharmaceutical Products in Twitter. *Drug Safety*, **37**, 343-350. https://doi.org/10.1007/s40264-014-0155-x

[5] Van Grootheest, K., de Graaf, L., de Jong-van, *et al.* (2013) Consumer Adverse Drug Reaction Reporting—A New Step in Pharmacovigilance. *Drug Safety*, **26**, 211-217. https://doi.org/10.2165/00002018-200326040-00001

[6] Yan, P., Chen, H. and Zeng, D. (2008) Syndromic Surveillance Systems. *Annual Review of Information Science & Technology*, **42**, 425-495. https://doi.org/10.1002/aris.2008.1440420117

[7] Leaman, R., Wojtulewicz, L., Sullivan, R., *et al.* (2010) Towards Internet-Age Pharmacovigilance: Extracting Adverse Drug Reactions from User Posts to Health-Related Social Networks. *Proceedings of the* 2010 *Workshop on Biomedical Natural Language Processing*, Uppsala, 15 July 2010, 117-125.

[8] de Langen, J., van Hunsel, F., Passier, A., *et al.* (2008) Adverse Drug Reaction Reporting by Patients in the Netherlands—Three Years of Experience. *Drug Safety*, **31**, 515-524. https://doi.org/10.2165/00002018-200831060-00006

[9] Gu, J.Z. (2009) Context Aware Computing. *Journal of East China Normal University*, No. 5, 1-20.

[10] Liang, S.X. and Li, F.L. (2014) Status Quo of Context-Aware Health Information Service System and Its Prospects. *Chinese Journal of Medical Library and Information Science*, **23**, 31-36.

[11] Farion, K., Michalowski, W., Wilk, S., *et al.* (2009) Clinical Decision Support System for Point of Care Use: Ontology-Driven Design and Software Implementation. *Methods of Information in Medicine*, **4**, 381-390. https://doi.org/10.3414/ME0574

[12] Andrej, M., Pierre, J., Olivier, D., *et al.* (2015) Ontology for Assessment Studies of Human Computer Interaction in Surgery. *Artificial Intelligence in Medicine*, **63**, 73-84. https://doi.org/10.1016/j.artmed.2014.12.011

[13] Lei, H., Sow, D.M., Davis, I., *et al.* (2002) The Design and Applications of a Context Service. *ACM Sigmobile Mobile Computing and Communications Review*, **6**, 45-55. https://doi.org/10.1145/643550.643554

[14] Li, F.L. and Li, N. (2016) Research on Context Based Ontology Modeling and Rule Reasoning for Medical Information Service. *Information Studies: Theory & Application*, **39**, 120-124.

[15] Tang, X.L., Zhang, B. and Fu, W.G. (2018) Context Ontology Driven Multi-Source Knowledge Fusion Framework. *Library and Information Service*, **62**, 109-117.

[16] Mou, Y.X. and Cao, Y.C. (2017) Advances in Research on the Effects of Gender Differences on Drug Effects. *Journal of Jilin Medical University*, **38**, 69-71.

[17] Wang, Y.S., Liang, Y.J. and Qian, J. (1999) The Relationship between Plasma Concentrations and Sedative Depth of Diazepam in Patients with Different Age. *Journal of Clinical Anesthesiology*, **15**, 195-196.

[18] Jonathan, P. and Jarow, M.D. (2018) Personalized Reproductive Medicine: Regulatory Considerations. *Fertility and Sterility*, **109**, 963-967. https://doi.org/10.1016/j.fertnstert.2018.03.027

[19] Wei, W. (2008) Consideration of Research on Some Basic Problems of Clinical Pharmacology. *Chinese Journal of Clinical Pharmacology and Therapeutics*, **13**, 1081-1085.

[20] Winter, A., Stäubert, S., Ammon, D., *et al.* (2018) Smart Medical Information Technology for Healthcare (SMITH). *Methods of Information in Medicine*, **57**, 92-105. https://doi.org/10.3414/ME18-02-0004

[21] Gao, X.D., Wang, W.X. and Wu, S. (2013) Multidimensional Association Rule Mining Method Based on Data Cube. *Computer Engineering*, **29**, 74-77.

[22] Zhai, Y., You, Y.M. and Ren, X.P. (2015) The Mechanism and Main Adverse Reaction of Diabetes Drugs. *Strait Pharmaceutical Journal*, **27**, 194-197.