



Knowledge engineering for adverse drug event prevention: On the design and development of a uniform, contextualized and sustainable knowledge-based framework

Vassilis Koutkias^{a,*}, Vassilis Kilintzis^a, George Stalidis^a, Katerina Lazou^a, Julie Niès^{b,c}, Ludovic Durand-Texte^d, Peter McNair^e, Régis Beuscart^c, Nicos Maglaveras^a

^a Lab of Medical Informatics, Medical School, Aristotle University, Thessaloniki, Greece

^b MEDASYS, Gif-Sur-Yvette, France

^c Lille University Hospital, EA2694, France

^d VIDAL, Issy les Moulineaux, France

^e Region Hovedstaden & Kennedy Center, Copenhagen, Denmark

ARTICLE INFO

Article history:

Received 22 July 2011

Accepted 12 January 2012

Available online 2 February 2012

Keywords:

Adverse drug event (ADE) prevention

Patient safety

Knowledge engineering

Knowledge-based framework

Contextualization

Clinical Decision Support System (CDSS)

ABSTRACT

The primary aim of this work was the development of a uniform, contextualized and sustainable knowledge-based framework to support adverse drug event (ADE) prevention via Clinical Decision Support Systems (CDSSs). In this regard, the employed methodology involved first the systematic analysis and formalization of the knowledge sources elaborated in the scope of this work, through which an application-specific knowledge model has been defined. The entire framework architecture has been then specified and implemented by adopting Computer Interpretable Guidelines (CIGs) as the knowledge engineering formalism for its construction. The framework integrates diverse and dynamic knowledge sources in the form of rule-based ADE signals, all under a uniform Knowledge Base (KB) structure, according to the defined knowledge model. Equally important, it employs the means to contextualize the encapsulated knowledge, in order to provide appropriate support considering the specific local environment (hospital, medical department, language, etc.), as well as the mechanisms for knowledge querying, inference, sharing, and management. In this paper, we present thoroughly the establishment of the proposed knowledge framework by presenting the employed methodology and the results obtained as regards implementation, performance and validation aspects that highlight its applicability and virtue in medication safety.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

Adverse drug events (ADEs) constitute a major public health issue endangering patients' safety and causing considerable extra healthcare costs [1]. An ADE is typically defined as “an injury due to medication management rather than the underlying condition of the patient” [2]. ADEs are classified as preventable and non preventable [3]; preventable ADEs are assimilated to “medication errors” [4], while non preventable ADEs are considered adverse drug reactions (ADRs) that could not be avoided [5].

A major challenge in research on ADEs and adverse events in general involves their identification and prevention [6]. Towards this aim, the potential of Information Technology (IT) tools and techniques has been highlighted in various studies [7,8]. In particular, major focus of IT-based research on ADEs has been the

automatic or semi-automatic identification of ADEs by employing machine learning and statistical inference techniques applied to patient data repositories [3,9–13], e.g. Electronic Health Records (EHRs). Besides statistical methods, knowledge-based approaches have been also employed for the identification of ADEs, e.g. based on ontologies [14], formal concept analysis [15], intelligent agents [16], and Semantic Web technologies [17].

In this regard, studies have been initially concentrated on the development of IT tools capable of providing evidence on the origin of ADEs, following typically experts review evaluation of the obtained results [9]. These outcomes were foreseen to constitute the basis for introducing/advancing the decision support functionalities on ADEs offered by clinical information systems, such as Computerized Physician Order Entry (CPOE) systems [18]. However, the majority of the proposed approaches have not elaborated further towards the incorporation of the ADE signals identified into actual Clinical Decision Support Systems (CDSSs) capable of inter-operating with clinical information systems, e.g. CPOEs and EHRs. As more mature evidence on ADEs' prevalence is gained, the focus

* Corresponding author. Address: Lab of Medical Informatics, The Medical School, Aristotle University, Box 323, 54124 Thessaloniki, Greece. Fax: +30 2310 999263.

E-mail address: bikout@med.auth.gr (V. Koutkias).

of IT research has been attracted by the incorporation of the identified ADE signals into sophisticated knowledge-based models for automatic ADE prevention. For example, Rommers et al. presented an ADE alerting system that consolidates clinical rules (formulated by a multidisciplinary experts team based on seven risk categories) to construct a Computer Interpretable Guideline (CIG) based decision support framework [19]. A total of 121 clinical rules were defined by the experts via the analysis of the Dutch national formulary and local medical reference books to identify drugs or drug classes suitable for use. Del Fiol et al. proposed a Knowledge Base (KB) incorporating 207 rules related to drug–drug interactions [20]. Major emphasis in this work has been given on the knowledge management potential for the end-users and on connectivity aspects of the proposed system with hospital information systems. In addition, aiming to reduce medical errors within hospitals, a prototype intelligent assistant has been presented by Payne and Metzler [21], following an ontology-based approach. The ontology encapsulates hospital care concepts including activities, procedures and policies, as well as medical knowledge, and is particularly designed to track the implications of medical decisions taken by health professionals within the context of guidelines/regulations of the medical environment, and the established medical knowledge.

Although significant progress has been made in both ADE identification and prevention, the efficiency of the results obtained by the proposed methods so far is still questionable due to the following major obstacles: (a) the lack of reliable knowledge about ADEs, and (b) the poor ability of IT solutions to deliver contextualized knowledge appropriate for each case [22,23]. Moreover, some studies concluded that over alerting may result in alert fatigue and alert overriding by the end-users [24,18], with major risk important alerts be overridden along with unimportant ones, thus, compromising patient safety.

Motivated by the above challenges [25], this paper presents a knowledge engineering framework that has been constructed aiming to represent and manage various ADE signals, with major focus on novel rule-based signals obtained through knowledge discovery techniques, and validated by following a knowledge elicitation phase [26]. Knowledge engineering constitutes the discipline that elaborates on the theories, methods and tools for developing knowledge-intensive applications [27–29]. In the scope of this work, knowledge engineering tasks involved first the systematic analysis of the relevant knowledge sources, resulting in the construction of a knowledge model and the selection of the appropriate knowledge engineering formalism. The model was employed to develop a relevant KB, i.e. the core component of the framework, encapsulating the abovementioned signals that are provided in the form of rules. The framework incorporates mechanisms for knowledge sharing, exploitation and management, as well as the appropriate inference component, all integrated within a uniform and sustainable architecture. In addition, the framework has been designed to constitute the basis for the construction of contextualized CDSS modules for ADE prevention, in order to contribute in the delivery of localized support services per clinical setting (hospital, clinical department, etc.), advancing the decision support impact and eliminating potential over alerting.

In this paper, we present thoroughly the establishment of the proposed framework. Specifically, Section 2 presents the employed methodology in terms of the elaborated knowledge sources, the constructed knowledge model, the employed knowledge engineering formalism, the architecture of the framework, as well as its underlying reasoning scheme. Section 3 presents the results obtained as regards the implementation of the respective Knowledge-based System (KBS), along with performance and validation aspects. Finally, the proposed approach and future research challenges are discussed in Section 4.

2. Material and methods

2.1. Knowledge sources

The current work focuses on the construction of a rule-based knowledge framework, which is designed to support ADE prevention through effective decision support delivered via alerts and recommendations to the clinical personnel. In particular, the knowledge elaborated in the framework consists of production rules [30], which are generally expressed in the form:

$$C_1 \text{ AND } C_2 \text{ AND } \dots \text{ AND } C_n \rightarrow E, \quad (1)$$

where C_1, C_2, \dots, C_n constitute the conditions of the rule, expressed in a general atomic formulae of some accepted language (e.g. propositional logic, first order logic, etc.), and E is the conclusion, action or decision. In the scope of this work, such rules correspond to ADE signals, i.e. the E part denotes a potential ADE that typically corresponds to a diagnosis or laboratory examination result along with a recommendation for actions and information as regards the explanation of the risk. The conditions C_i correspond to: (a) groups of drug codes expressed in the ATC (Anatomical Therapeutic Chemical) classification system, (b) groups of laboratory examination results expressed in C-NPU/IUPAC (Nomenclature, Properties and Units/International Union of Pure and Applied Chemistry), (c) groups of diagnosis codes encoded in ICD-10 (International Classification of Diseases), or (d) patient parameters compared to numerical or categorical values, e.g. age and gender. Thus, the conditions C_i denote a special type of rules that we call “intermediate” (as these are the building blocks for defining the ADE signals). As an example, an intermediate rule defines the variable “Antibiotic” as the presence of any member of a set of ATC codes corresponding to individual antibiotic drugs. The exploitation of these ADE signals for decision support is initiated by a drug-related procedure, such as a new drug prescription, which triggers the rules’ assessment based on the provided patient data.

The types and origin of the elaborated ADE signals are primarily: (a) *Association or decision-tree induced rules* obtained by applying data-mining techniques on routinely collected patient records of past hospitalizations from various hospitals across Europe [26], according to a common data structure (specifically designed for this analysis) [31], and validated by clinical experts [32]. Data-mining aimed at detecting atypical hospital stays and, subsequently, at extracting associations among drugs, hospitalization parameters, patient parameters, diagnoses and observed effects. (b) *Drug interactions*, e.g. drug to drug, drug to allergy class, drug to laboratory examination result, drug to diagnosis, etc., that are already known and registered in pharmacovigilance KBs.

In addition, our research elaborated on knowledge sources such as: (a) the *literature*, i.e. obtaining evidence from either similar statistical analysis performed on clinical data repositories or focused drug-safety related studies [32]; (b) *tacit knowledge* [33], which was primarily captured in the knowledge elicitation process in which experts validated the data-mining originated rules based on their experiences and specialties, and (c) *human factors and clinical procedures analysis*, resulting in specifications as regards the logic according to which the ADE signals discovered should be applied in practice for the particular domain context, as well as in recommendations for the CDSS design and functionality [34].

Especially, for the data-mining originated rules, the importance and applicability of each rule is determined based on its statistical significance in the local context that is being triggered [22], i.e. hospital or clinical department. Thus, statistical features for each rule such as the *confidence* (probability of having the effect knowing that the conditions are met), the *support* (probability of having the effect and matching the conditions at the same time), the

Fisher test *p*-value, etc. for assessing its statistical significance constitute rules’ “meta-data” that may be particularly used for contextualization of rules [32]. In this regard, the proposed knowledge framework incorporates a context-sensitive, “meta-rule” level, which is employed to address rule ranking and determine the applicability of ADE signals per patient and clinical setting. Thus, meta-rules along with the contextualization perspective proposed in this work advance the source knowledge obtained in the discovery phase towards its effective use for decision support operation.

Besides the incorporation of statistical features, rules obtained by applying data-mining techniques are able to associate more complex patterns of conditions compared to the rules that are typically met in pharmacovigilance KBs. In the later case, the effects of drug discontinuation may be ignored and additional conditions that could specify more accurately the probability of the ADE occurrence are not taken into account [26].

As an example, a data-mining originated rule is the following:

```

dr1.diuretics_potassiumLowering = 1
AND di1.urinary_retention = 0
AND di1.cardiovasc_hbloodpressure = 1
AND di1.endoc_diabetes = 0
AND dr1.agtCon vInhib_agtAntag = 0
AND mi1.age.quant1 >= 70
AND di1.diag_hypo vol_dehydr = 0
→ bi.kidney_i,
    
```

which is translated in physical language as “Treatment with potassium lowering diuretics & NO urinary retention & high blood pressure & NO diabetes & NO treatment with angiotensin conversion enzyme inhibitor & age ≥ 70 & NO diseases at risk for hypovolemia may result in renal failure”. In this case, the rule consists of seven conditions with the first one being the condition/cause triggering the rule

(a potential drug prescription), the next three referring to conditions that are clinically associated with the effect, and the rest corresponding to conditions that, although not explicitly linked with the outcome from a clinical viewpoint, appear to significantly increase the probability of its occurrence. The effect of this rule constitutes a laboratory examination result that is associated with renal failure. All the conditions, apart from the condition that involves the age, correspond to groups of either drug codes in ATC classification or diagnosis codes in ICD-10 classification, being either present (=1) or absent (=0) from the evaluated patient data.

From a clinical perspective, the above example rule describes that there is a risk of renal failure, when a patient over 70 is prescribed a potassium lowering diuretic for high blood pressure (hypertension). Assessing this rule in three different hospitals participating in the study, we obtained different risk levels according to the considered clinical setting, as demonstrated from the statistical features provided in Table 1. Thus, healthcare professionals may not assess the risk in the same way in all hospitals, for example in cardiology or surgery, or if they have to add another drug in the treatment plan (another diuretic or an antibiotic modifying renal function). In this regard, the alert aims to facilitate the optimization of the risk-benefit analysis taken into account by the healthcare professionals for each patient, as it provides information concerning the relative risk of the ADE, which is weighed by the probability of occurrence. In our experience, this alert can have two consequences: (a) Rarely, to change the medication and (b) more often, to monitor more precisely and more adequately the patient’s medical status, e.g. by better monitoring of uremia and creatininemia, if renal failure occurs.

Fig. 1 illustrates the overall classification of the data-mining originated rules that have been elaborated in this work, according to the associated effects and causes, where the size of each circle denotes the available amount of rules per category. A detailed description of the techniques employed in the knowledge discovery phase and the categories of the obtained rules along with further rule examples are presented in the work of Chazard et al. [26].

The analysis of the above knowledge sources resulted in the definition of a relevant knowledge model, which is presented in the following subsection.

Table 1
Statistical prevalence of the example rule described in (2).

	Hospital 1	Hospital 2	Hospital 3
Ratio: Number of patients meeting both the conditions and the effect/number of patients meeting the conditions	5/74 = 6.8%	2/6 = 33.3%	2/21 = 9.5%
Fisher test <i>p</i> -value	0.0797	0.0047	0.7557

2.2. Knowledge model description

The knowledge model construction process aimed at developing a KB structure that effectively captures source knowledge on ADEs and is generic enough to capture relative knowledge from a

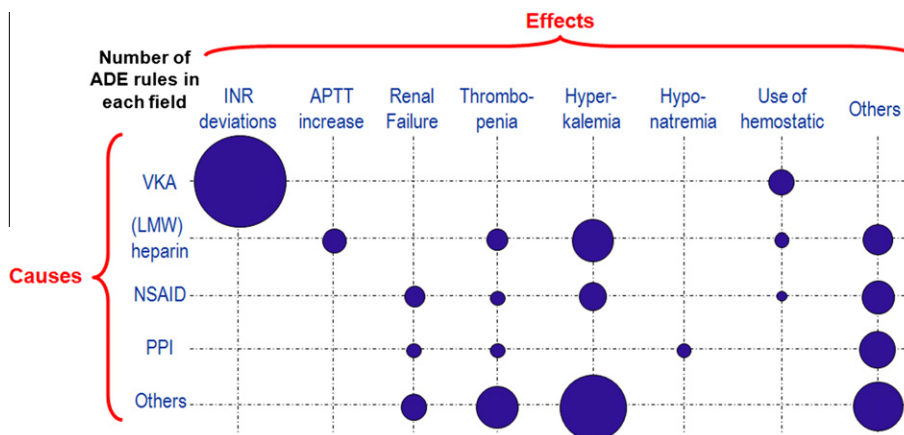


Fig. 1. Classification of ADE rules according to the associated effects and causes; the size of each circle denotes the available amount of rules per category (APTT: Activated Partial Thromboplastin Time, INR: International Normalized Ratio, LMW: Low Molecular Weight, NSAID: Non-Steroidal Anti-Inflammatory Drug, PPI: Proton-Pump Inhibitor, VKA: Vitamin K Antagonist).

wide spectrum of possible existing or future sources. In addition to basic knowledge on ADE prevention in the form of rules, the developed knowledge model is designed to express higher-level knowledge, in the form of meta-rules, as regards the effective use of ADE signals, their scope of applicability and contextualization parameters. The model encapsulates background knowledge of the application domain via the intermediate rules and the terminologies adopted, since the definition of the drug, diagnosis and laboratory examination related conditions correspond to application-specific knowledge. The conceptual schema of the model has been presented in [35].

The proposed model can be logically divided into several components that are of different type and may be considered as sub-models in the overall framework, with some of the components corresponding to knowledge as derived from the sources and others providing the background terminology to support the former. Primarily, the knowledge model encapsulates rule-based components that are defined via a set of classes and populated with ADE signals and ontology-based structures, either problem-specific or based on standard classifications. The knowledge model is illustrated in Fig. 2, via a UML (Unified Modeling Language) class diagram. In the design followed, all the elements and concepts

involved are represented as classes or groups of classes along with their associations, while the attributes defined for each class reflect the adopted approach regarding the rule structure, contextualization mechanisms and relation to standard classifications.

The domain knowledge in this work is represented by a number of classes that are used to express the administration or discontinuation/absence of a specific medication, diagnoses (information about medical conditions existing for the patient), results of laboratory examinations, other patient parameters (e.g. age, weight, etc.), as well as the predicted effects. These classes are namely:

- (a) *Drug* and *Drug_specific*, where each *Drug* instance corresponds to a set of possible medications of the same family, e.g. vitamin K antagonists, that is given a unique name taken from the namespace defined in the data-mining process for drug variables. The individual drugs that belong to the group of one *Drug* are instances of the class *Drug_specific* and are coded according to the ATC classification.
- (b) *Lab*, which is used to express results of laboratory examinations in the form of binary values that indicate a specific type of anomaly. This is expressed via several instances of

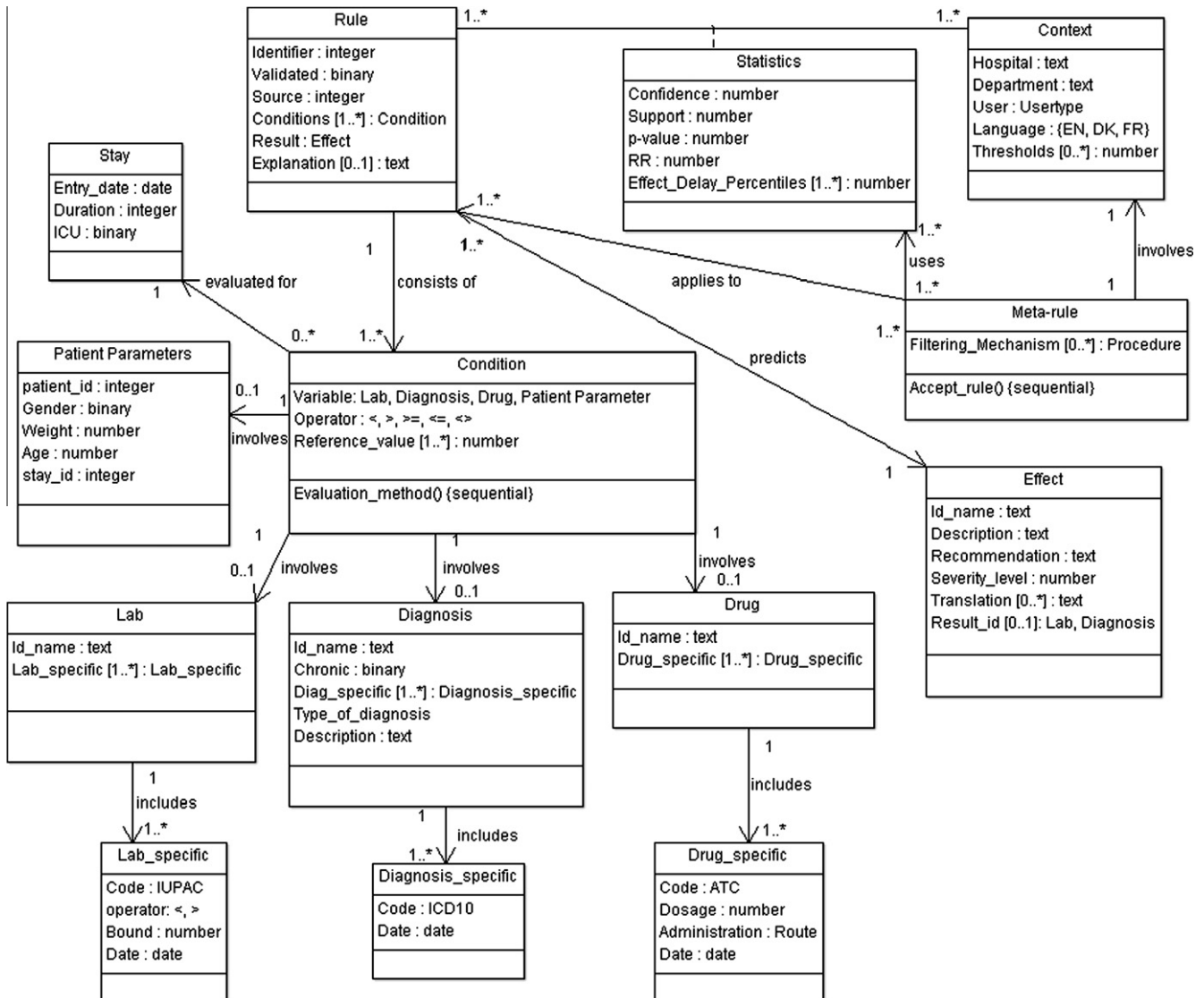


Fig. 2. Class diagram of the proposed knowledge model for ADE signals representation.

the class *Lab_specific*. In particular, the specific type of examination is identified using the relevant C-NPU/IUPAC coding, while the boundary and the corresponding arithmetic comparison operator are also defined for each examination.

- (c) *Diagnosis* that expresses a known patient condition that may play a role in a possible ADE. Chronic medical conditions are differentiated from acute diagnoses (as discriminated also in the source ruleset). A *Diagnosis* instance typically corresponds to a set of possible specific diagnoses of a similar group, e.g. cancer. Specific diagnoses are defined with the class *Diagnosis_specific* and are identified by their ICD-10 code.
- (d) *Patient Parameters* expresses information about the patient, such as age, weight, and gender. The respective attributes comply with the fields defined in the common data model [31], which has been adopted to specify the link between the present framework and hospital information systems.
- (e) *Effect* defines the predicted outcome of a rule and may contain a recommendation for actions. All possible instances of the *Effect* class are predefined in a problem-specific taxonomy that has been produced during the data-mining process and is considered as domain knowledge acquired from the data-mining knowledge source.

The main knowledge on predicting possible ADEs (inference knowledge) is represented by the classes *Rule* and *Condition*. Each *Rule* instance corresponds to a uniquely identified rule that is linked to the associated *Conditions* and *Effect*. The provenance of each *Rule* with respect to the considered knowledge sources is defined via the “Source” attribute. *Condition* is a class which defines the structure of each condition participating in the rules, containing the variable on which the condition is based (i.e. drug, diagnosis, laboratory examination or patient parameter), a comparison operator and a reference value.

In addition to the main knowledge on ADE prevention contained in the *Rules*, the knowledge model provides elements concerning the effective application of these rules. For this purpose, the classes *Meta-rule*, *Context* and *Statistics* are defined. The *Meta-rule* contains knowledge in the form of a process that indicates whether a rule should be considered or not in specific circumstances. Each *Meta-rule* instance applies to a set of *Rule* instances in relation to a specific *Context* instance. The *Meta-rule* operates as a filtering mechanism that may “de-activate” certain rules in a particular context. The *Context* class contains as attributes parameters that can be used to specify a particular local setting, such as the hospital/clinic, and the targeted user, while the logical process that is followed to evaluate whether an ADE signal should be visible or not is represented in the class diagram as the *Accept_rule()* method of class *Meta-rule*.

The class *Statistics* is defined as an association class which contains a set of statistical parameters (meta-data) related to the application of a *Rule* in a particular *Context*. These parameters are initially evaluated for each rule discovered during the data-mining phase. However, the statistical features may also be evaluated in the specific clinical epidemiology in which the KB is to be used. Hence, the thresholds for rule application may be adjusted for the hospital or clinical department where the KB is used. The value of such statistical parameters is an indication of how likely this rule will fire (sensitivity), and with what confidence the predicted event will actually happen (predictive value). Thus, their usefulness is two-fold: (a) to enable rule filtering according to their statistical significance, and (b) to adjust the KB to a specific hospital and clinical department by evaluating the parameters locally.

It is interesting to note that, although the proposed model encapsulates concepts targeting ADE prevention via rule-based signals, it may be reusable for expressing rules that are applicable in

the clinical environment in general, and especially for rules obtained by applying data-mining techniques on EHRs (as in this case, the statistical features of the rules are especially applicable). In the general case, laboratory examination results and diagnoses may participate in the conditions of rules for a particular clinical domain without requiring drug-related conditions to be present. Although the model has been designed without having drugs as a mandatory part (however, all the rules that have been elaborated in the scope of this work contain at least one drug), the exact definition of the included concepts was primarily driven by the scope of the specific knowledge discovery activities (e.g. note the attributes included in the *Patient Parameters* class). However, as the major concepts of the model are core concepts in healthcare settings, further attributes could be introduced to extend the model, if necessary. A limitation of the model that was implied by the elaborated knowledge sources involves the support of crisp rules (i.e. the comparison of data with a specific number via an arithmetic operator) via the definition of the *Condition* class. Nevertheless, the overall rationale and the design of the model may be reusable for expressing rule-based knowledge for other clinical applications besides ADE prevention.

2.3. Knowledge engineering formalism

Several knowledge engineering formalisms and methodologies were investigated for the design and development of the proposed knowledge framework [27–29]. The primary aim was to develop a KB in correspondence with the presented knowledge model, i.e. comprising of a set of ontology-based structures, either application-specific or standard classifications, as well as to include a rule-based component that is defined via a set of classes and populated with ADE signals. The above knowledge components constitute the fundamental elements for defining complex procedural logic for ADE prevention. In this regard, as unification of the former knowledge components was required, our analysis resulted in the adoption of the CIG formalism [36,37], as the basis for developing the common knowledge framework to deploy ADE prevention services. Initiated by the need to electronically encode clinical practice guidelines in order to incorporate them in decision support and monitoring applications, the knowledge modeling approaches that have been proposed in the literature resulted in the establishment of CIGs as a knowledge engineering formalism that can be used to express procedural knowledge for a variety of clinical applications [38–40]. Thus, in the scope of this work, the term CIG refers to this established knowledge engineering formalism in healthcare, which is applied in our case for describing the paths leading to an ADE forecast based on a variety of knowledge sources and reasoning steps.

The majority of CIG formalisms rely on “Task-Network Models”, while the following are considered as their common elements [40]: (a) patient states, (b) execution states, (c) eligibility criteria, (d) classification schemes, (e) goals, (f) decisions and (g) actions. Ontologies, rules and procedures that constitute the building blocks of the proposed framework are effectively supported and uniformly integrated via CIGs [36,41].

The application of the CIG formalism for constructing knowledge-based medical applications for decision support has been reported in several studies [42–45]. Especially, in the application domain of patient safety and error prevention, Grando et al. presented a generic approach for handling exceptions in workflow execution based on a flexible workflow definition schema using clinical goals at multiple hierarchical levels, and separating exception detection and handling from normal workflow execution [46]. The approach allows for modeling of exceptions that occur during CIG execution and their handling, and addresses some important barriers to CIG adoption such as the ability to provide decision-support in

the case of exceptions, and the fact that the majority of CIG-based CDSSs do not work well when the encoded knowledge is incomplete. Seyfang et al. elaborated on the detection and prevention of risks against patient safety by comparing the actual care practice against descriptions of best practice given in clinical guidelines and protocols (CGPs) [47]. In order to perform such comparisons automatically, CGPs are modeled in a computer-executable form, while the execution of the CGP model is integrated with the care process at the site of application, and with risk-assessment tools used by the hospital's risk manager to explore "what-if" scenarios.

The current work employs CIG as the main engineering formalism in the construction of a unified knowledge framework for ADE prevention. Our research involves knowledge engineering of novel ADE signals discovered by applying data-mining techniques across diverse EHRs across Europe [26], aiming to integrate them with other knowledge sources into a single KB supported by appropriate management mechanisms and fully exploitable at the point of care. In practical terms, each ADE prevention rule incorporated in the framework constitutes a CIG (in other words, a protocol or workflow). In addition, the context-sensitive strategy employed to support the delivery of localized knowledge to the end-users, involves the definition of filtering mechanisms (i.e. the "meta-rules"), which are also defined as CIGs. The overall architecture of the knowledge framework is presented in the following subsection.

2.4. Knowledge framework architecture

The major components of the knowledge framework architecture are (Fig. 3):

- (a) *Knowledge Base*: Constitutes the central part of the framework, encapsulating the source knowledge concerning the ADE signals into an exploitable structure and in an appropriate electronic format. The knowledge involves the domain ontology, i.e. standard classifications and major concepts that are used to express parts of the ADE signals, inference knowledge in terms of defining the ADE signals as a whole, as well as task knowledge, i.e. procedures and protocols defining the scope according to which the ADE prevention rules are applicable/used or not.
- (b) *KB Instantiation & Update Mechanism*: Provides the means for populating the KB with ADE signals, supporting also potential updates of the KB content. In particular, an automatic import mechanism has been developed, enabling straightforward KB population with ADE signals expressed in an XML (eXtensible Markup Language) based document structure [32], which was commonly agreed between the knowledge discovery experts and the knowledge authors.
- (c) *Knowledge Verification Mechanism*: Performs syntactic verification of source knowledge [32], upon launching the KB instantiation or update procedures. In case of errors identified, the knowledge authors are prompted with informative messages as regards the relevant inconsistencies to resolve them.
- (d) *KB Contextualization Mechanism*: The KB content is contextualized via ADE signals' meta-data, such as the statistical features per ADE signal in the particular context (i.e. hospital or clinical department) [22]. As these data are dynamic by nature, it is expected to be recalculated in periodic time intervals for the local context, as part of the knowledge discovery phase [26]. Upon the availability of the updated figures, the mechanism supports the update of the KB content as regards this part per se. In addition, it handles linguistic issues concerning the knowledge delivered to the end users (English, Danish and French languages are currently supported).

- (e) *Knowledge Export Mechanism*: Enabling knowledge sharing and reuse is a typical requirement in knowledge engineering tasks, adding value concerning the exploitation potential of knowledge. Several export mechanisms have been developed and made available in the proposed framework, either linked with implementation details and tools (e.g. Resource Description Framework (RDF)/XML), or standards defined/applicable for the medical domain (i.e. HL7/ANSI GELLO [48]).
- (f) *Inference Engine*: The inference part involves the CIG execution engine and filtering mechanisms that are particularly targeting elimination of alert fatigue, i.e. filtering ADE signals with respect to contextual information and meta-rules. Inference has its basis on a FSM (Finite State Machine) based execution engine (the typical, inherent mechanism of CIGs [36,37,41]) that executes rules upon data requests, as well as on a basic terminological reasoning mechanism developed to address query expansion on the requests posed, when applicable and according to the underlying semantics (as explained in the description of (3), in Section 2.5). Contextualization of the *Inference Engine* is achieved via the definition of various configuration parameters that control activation of the filtering mechanisms, define thresholds for statistical parameters, customize the KBS output, and so forth.
- (g) *Interface to External Knowledge Sources*: There are several knowledge sources available related to ADEs, such as drug–drug interaction databases. In this regard, the framework provides a generic database import mechanism, which is designed in a parameterized way to assure its reusability. In this viewpoint, the KB may be considered as a distributed knowledge source, which enables connection with external KBs and data, provided that there is an efficient mechanism able to exploit all knowledge components as a whole. In addition to technical aspects, organizational issues may be even more important for preferring a connectivity mechanism with external knowledge sources, instead of an import mechanism, e.g. issues including ownership of the content, responsibility/accountability, maintenance and control, thus the adopted approach.
- (h) *Interface to External Healthcare Systems & Services*: This constitutes the necessary mean for exploiting the KB content and the reasoning services provided by the knowledge framework. The interface involves both the request and the response parts of communication. It is XML-based providing this way a clear specification as regards its implementation for external hospital systems, i.e. EHR or CPOE systems. The interface relies on the mapping of each concept/attribute defined in the knowledge model with the relevant tables/fields respectively defined in the data model [31].

The functionality offered by the proposed framework in terms of its reasoning scheme and exploitation in decision support operation is formally presented in the following subsection.

2.5. Functional attributes and reasoning scheme

The KB developed aims to constitute the core part of CDSS modules for ADE prevention; thus, the *Interface to External Healthcare Systems & Services* enables querying the KB with patient data, so as to assess the case(s) of interest against the ADE signals incorporated in the KB. Let us consider a data request as the input to a relevant CDSS module that is expressed as the tuple $\langle Dr, Di, Bi, Mi \rangle$. Dr corresponds to a set of drug values, Di to a set of diagnosis values, Bi to a set of laboratory examination results and Mi to a set of medical parameter values concerning a patient (with Di , Bi and Mi being optional parts in general). An inference mechanism f_1 is

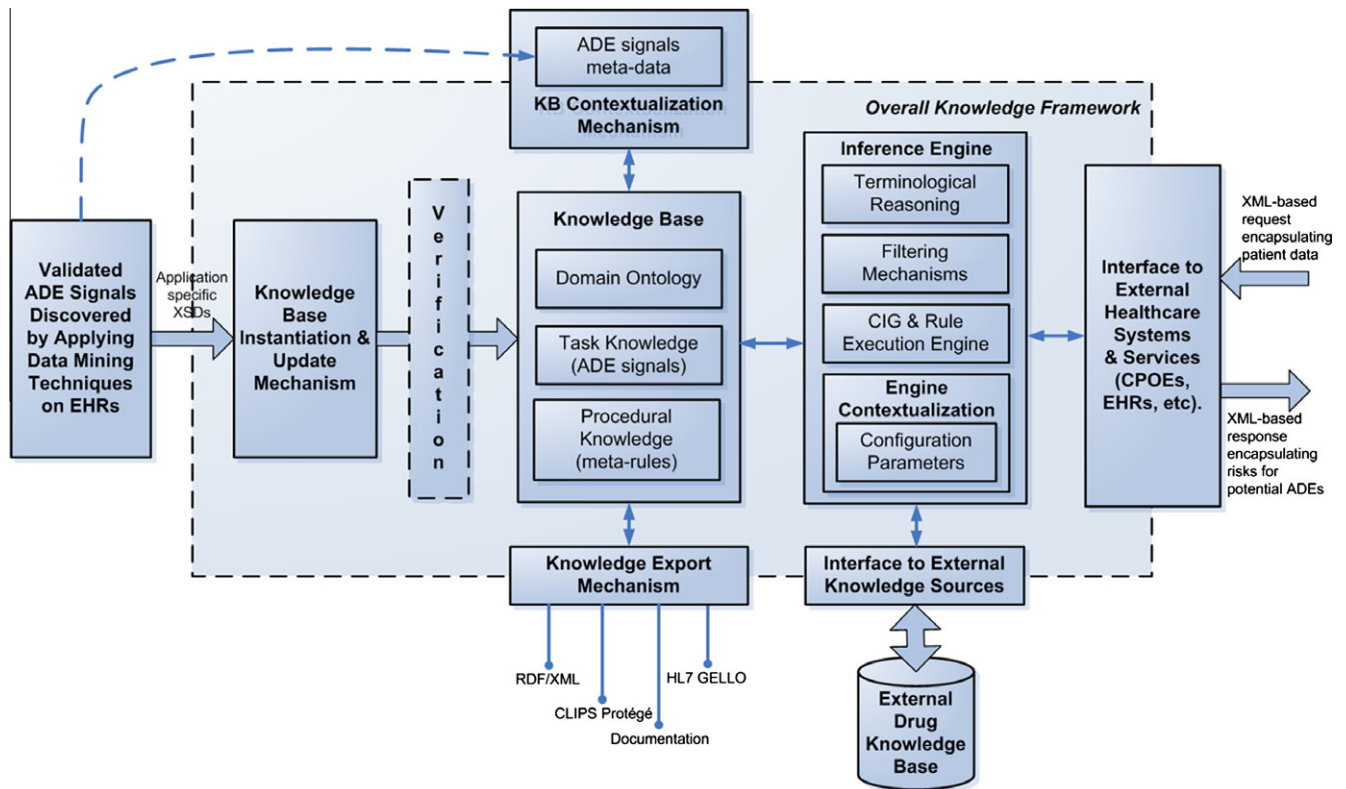


Fig. 3. Overall architecture of the proposed knowledge framework.

introduced in order to match the above tuple in the set of ADE prevention rules R that are incorporated in the KB, i.e.:

$$f_1 : \langle Dr, Di, Bi, Mi \rangle \rightarrow R. \quad (3)$$

In essence, f_1 involves the assessment of matching conditions in the entire ruleset incorporated in the KB. It also provides the means to support the dynamic filtering of input data according to temporal constraints, e.g. considering only the last laboratory measurement per C-NPU/IUPAC code, the recent drug prescriptions, etc. (configurable options). In addition, particularly for diagnosis-related data, the mechanism examines whether there are rules in R that match diagnosis conditions, in case of non-exact match in Di (encapsulated in the request), taking advantage of the *is-a* hierarchy of ICD-10 (e.g. let us consider a rule with matching condition the ICD-10 code A071A and a data request sent to the KBS that does not contain A071A, but A07; the condition is considered true, as A071A is subclass of A07). The outcome of this procedure is a new set $R_1 \subseteq R$, corresponding only to the relevant ADE signal(s).

In case of multiple applicable ADE signals, i.e. multiple rules triggered by f_1 , a major issue in the application of discourse constitutes the assessment of the significance per alert according to the case, so as to eliminate over-alerting. In this regard, the KBS is fine-tuned following a context-sensitive strategy (in both construction and runtime mode) that applies meta-rules in terms of threshold-based filtering as regards the statistical significance of the corresponding triggered rules, to determine the most significant alerts or recommendations that will reach the CDSS end-user. Thus, a new mechanism f_2 is introduced that maps R_1 into a set R_2 , with $R_2 \subseteq R_1$, according to the contextual criteria C_X :

$$f_2 : R_1/C_X \rightarrow R_2, \quad (4)$$

C_X may be a set of statistical thresholds associated with each ADE signal in the local context, such as confidence >20% and Fisher test *p-value* <0.05. It has to be noted that there are cases where it is

necessary to preserve a part of R_1 (i.e. specific rules that might be considered important) as an outcome of (4), independently of the constraints introduced by applying the criteria C_X . This is also a configurable option supported by f_2 . The final outcome of f_2 is a list of effects that correspond to R_2 , along with appropriate explanations of the respective rules, the importance of the potential ADEs, as well as the data that made the rules fire. The language of the text and the user type (clinician/nurse/patient) constitute also configurable runtime options.

Finally, the framework supports enabling/disabling interfaces for querying external pharmacovigilance databases; thus, R_2 may be potentially extended with additional signals into R_3 , with $R_2 \subseteq R_3$. In that case, the final outcome constitutes a unification of the signals generated by all the considered knowledge sources.

3. Results

3.1. Implementation platform

Implementation of the presented knowledge framework was based on Gaston [49]. The core of Gaston consists of a CIG formalism relying on a combination of knowledge representation approaches and concepts, i.e. primitives, problem solving methods (PSMs), domain and method ontologies. Domain ontologies model domain-specific knowledge in terms of entities, attributes and relations, while method ontologies model concepts such as primitives, PSMs and guidelines. The ontologies are defined in the frame-based version of Protégé [50]. Frames are used to represent: (1) knowledge related to the application domain (domain ontologies), and (2) knowledge related to the guideline's control structure (method ontologies). CIGs are represented in turn by a set of primitives or by means of PSMs. The two main advantages offered by Gaston with respect to the knowledge engineering requirements

posed in this work that lead to its adoption are: (a) support for all the identified knowledge engineering formalisms, i.e. ontologies, rules and protocols encapsulated its CIG formalism, and (b) support for various interface technologies to explore/query the underlying KB.

It is interesting to note that during the KB implementation phase, a proof-of-concept development was conducted to assess the portability and implementation of ADE rules in another KBS. In particular, the potential offered by the open-source rule engine Drools [51] was investigated. The selection of Drools was mainly because of its openness, the fact that its programming environment was familiar to the developers' team, and its wide exploitation in business-logic oriented applications, appropriate for the procedural knowledge elaborated in this work. In our case, an application was developed that automatically creates the relevant KB (intermediate and main rules) based on the source knowledge.

Both Gaston and Drools were successfully used in this development.

3.2. Knowledge-based system

Using Gaston for representing the ADE prevention rules in the form of a KB, involves instantiating and extending its underlying knowledge model. In essence, the KB comprises several ontologies related to domain knowledge, the mapping between the common data model and the domain knowledge model, the overall ADE rule component (i.e. the intermediate rules, the ADE signals and the meta-rules), along with procedures and options for exporting this knowledge. An excerpt of the KB structure is illustrated in Fig. 4 via the Protégé knowledge modeling tool.

The concrete implementation of the knowledge model presented in Section 2.2 involves mapping its items to Gaston's knowledge representation contracts and concepts. For example, each ADE signal of the framework is an instance of the class *Guideline* in Gaston,

under the hierarchy *Method_Entity* > *Complex_Method_Entity* > *Guideline*. Similarly, in the internal Gaston representation, each intermediate rule of the framework (as defined Section 2.1) is an instance of the class *Intermediate*, under the hierarchy *Method_Entity* > *Complex_Method_Entity* > *PSM* > *Common_PSMs* > *Situation-Action-Rule_Entity* > *Intermediate*. Finally, mechanisms such as meta-rules, filters, and interfaces to external knowledge sources, constitute also instances of the *Guideline* class.

The *KB Instantiation and Update Mechanism* (Fig. 3) developed enables straightforward population of the KB schema. Currently, the KB contains 55,687 classes and 33,879 instances. The above KB features correspond to the incorporation of 236 ADE rules and 403 intermediate rules in total, besides the incorporation in the domain knowledge of ICD-10, ATC and C-NPU/IUPAC standard classifications and codings, as well as rule-specific concepts (e.g. *Patient*, *Stay*, etc.). The large number of classes is due to the large domain ontology defined (as standard classifications have been included in the KB), while the similarly large number of instances is because the number of steps and conditions in the guidelines defined is also large.

Example meta-rules defined are: (a) the *Static Filter*, i.e. a rule may fire, only if certain rule meta-data parameters reach specific thresholds (*p-value*, confidence, relative risk, severity of the effect, etc.); (b) the *Temporal Filter*, i.e. rules do not generate alerts, if the period since all their conditions are met and the current time is more than a threshold (that corresponds to a rule-specific time delay parameter), and (c) handling *Rules with a Laboratory-related Effect*, i.e. when a rule involves a laboratory examination result abnormality as the effect, the output will be suppressed in case there is laboratory examination result corresponding to the effect and indicating absence of the effect after all the conditions of the rule are met.

In order to provide encapsulation of expert knowledge related to a specific context, another procedure named *Final Manual Filter* was defined. Configuring the *Final Manual Filter* for specific rules, experts

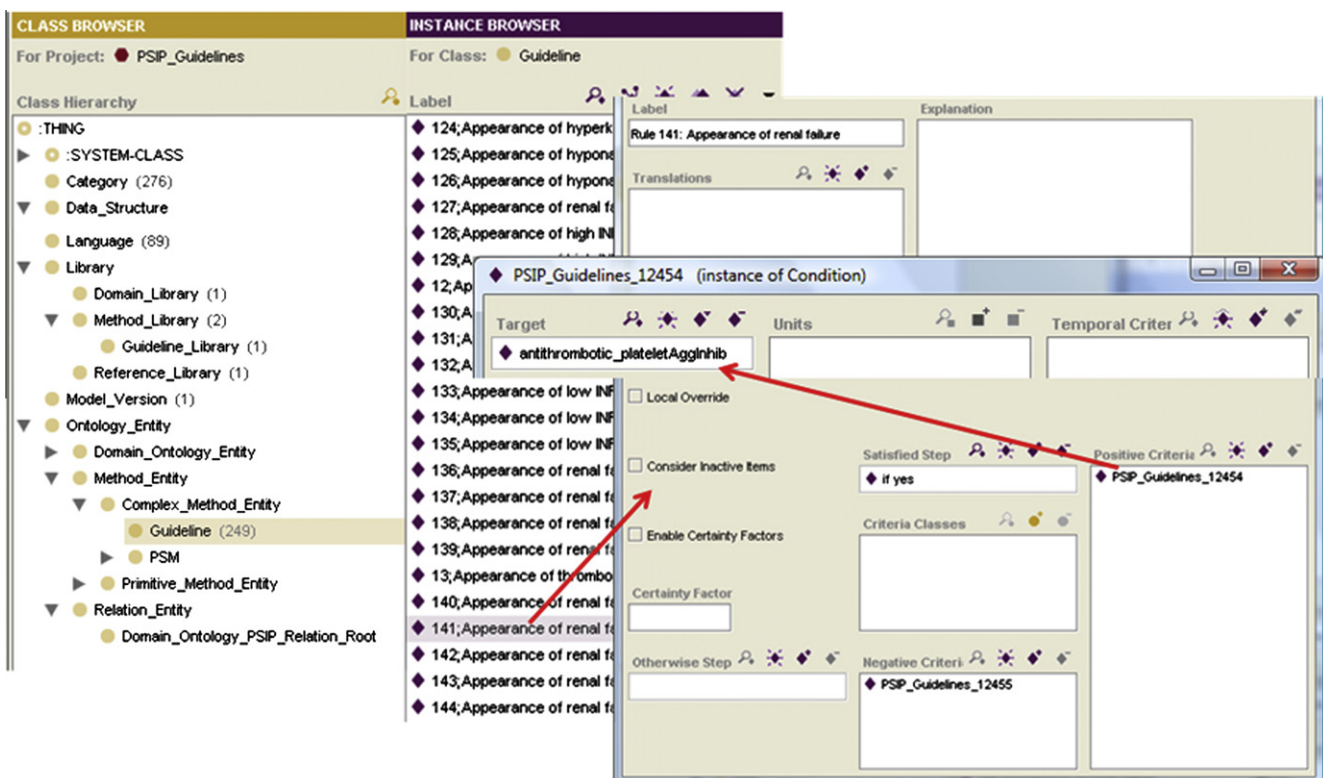


Fig. 4. Excerpt of the KB in Protégé (basic, ontology class hierarchy on the left side and example *Guideline* instance depicted on the right side).

may override the filtering mechanisms of the previous steps for these rules, e.g. for a rule that involves an effect that occurs often in the particular clinic, one could suppress the *Temporal Filter* for this rule to allow generation of alerts without temporal constraints.

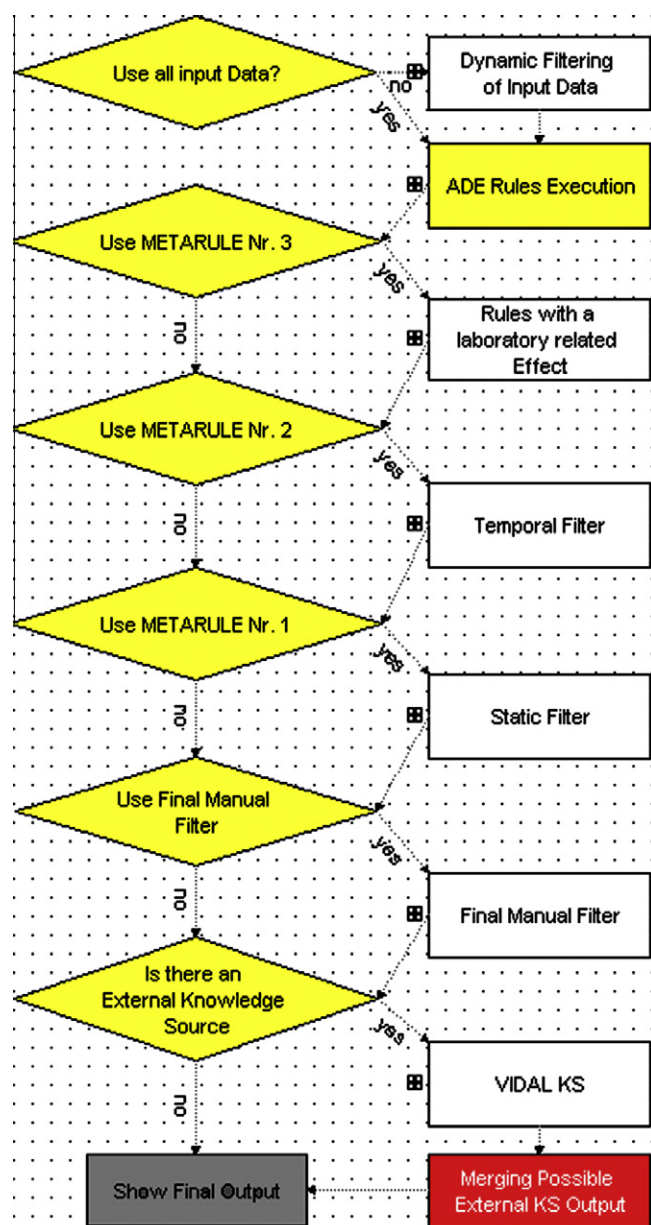


Fig. 5. The overall workflow for generating ADE signals based on the reasoning steps and knowledge components supported by the proposed framework, as illustrated in Gaston's knowledge acquisition tool.

The overall workflow encapsulating the reasoning steps for exploiting the ADE prevention knowledge provided by the proposed framework is illustrated in Fig. 5. In the depicted flowchart each procedural step is represented by a distinct shape. The decision steps correspond to configuration choices that may be set in the local environment, in order to enable or disable features and subsequently define and adapt the generated output.

3.3. Interfacing with external knowledge sources

Besides data-mining originated rules, a commercial knowledge source concerning drug interactions has been made available for

this development by Vidal (<http://www.vidal.fr/>), following also a rule-based formalism that is based on the abovementioned standard terminologies/classifications. Thus, the *Interface to External Knowledge Sources* component (Fig. 3) has been instantiated. The nature of the Vidal knowledge source involves rules that may relate two drugs regarding their interaction or one drug with a particular diagnosis, allergy, laboratory examination, or patient characteristic. The result of each rule is a description of the possible ADE and, depending on the type of rule, it can also include a degree of severity or a suggestion for action. The main sources of the abovementioned drug interactions come from international scientific publications, health agencies, guidelines and summaries of product characteristics (SPCs) provided by AFSSAPS (Agence Française de Sécurité Sanitaire des Produits de Santé), EMEA (European Medicines Agency), FDA (Food and Drug Administration), etc.

Such a knowledge source is considered as complementary to the data-mining originated rules which were introduced in this work and obtained by analyzing routinely collected patient data of past hospitalizations. There might be cases where important ADE cases may not be identified in patient data repositories through data-mining techniques, because either the conditions never occur, or when the conditions are present, the outcome never occurs. This is probably because such cases are well known and, consequently, the relevant risks are well monitored in clinical practice [26].

3.4. Knowledge sharing and reuse

The KB is available in the following representations: (a) Protégé, frame-based ontology and (b) RDF/XML structure. Although Protégé is an open-source tool and RDF/XML originates from an open specification of W3C (WWW Consortium), the above representations are specific to the adopted development platform.

Standardization constitutes a major challenge in medical knowledge modeling (and in CIG representations in particular [40]). Among the several efforts made, GELLO is the most relevant to the application discourse, as it is based on the CIG formalism. GELLO was initially conceived as a standard expression language for decision support [52], and it has evolved as an HL7/ANSI standard decision support language [48]. It aims primarily to constitute a query language for obtaining clinical information from an EHR system in a standard way. It uses an abstract "virtual medical record" (VMR), a simplified view of HL7 RIM (Reference Information Model) V3, so that the same code can run on multiple systems accessing data stored in different formats.

In this regard, aiming to reinforce knowledge interoperability and reuse of the presented KB through a knowledge representation standard applicable in the medical domain, HL7/ANSI GELLO was employed. In particular, the intermediate and main ADE rules of the KB have been represented in the GELLO.

3.5. Evaluation

Evaluation of the framework has been conducted with respect to implementation aspects, performance testing and quality of the KB content. As regards the implementation, components of the framework (e.g. the KB import/export mechanisms, the reasoning scheme and the interfaces), have been iteratively and extensively evaluated by the knowledge authors using several benchmarking tests, e.g. syntactic verification, extensive querying, exception testing, etc. As the development of the framework was performed incrementally, several technical shortcomings and errors were addressed in each development phase.

Given the large KB content and the ultimate goal of exploiting the KBS in practice via CDSSs at the point of care, performance constitutes an important issue. Thus, an analysis was conducted with a

variety of requests and runtime configuration options. Each step defined in the procedure depicted in Fig. 5 (if enabled) introduces a particular computational cost at the KBS runtime. In the particular case, when all features are enabled, the execution flow begins with the dynamic filtering of input data that corresponds to 27% of the total computational cost. The following step involves the main ADE rule execution (corresponding to the f_1 mechanism presented in Section 2.5), which is responsible for 20% of the total computational cost. The execution continues with meta-rules nr. 3, nr. 2 and nr. 1 (i.e. *Rules with a Laboratory-related Effect*, the *Temporal Filter* and the *Static Filter*, respectively, corresponding to the f_2 mechanism presented in Section 2.5). Meta-rules nr. 3 and nr. 2 introduce each 13% of the total computational cost, while meta-rule nr. 1 is responsible for the 20%. The *Final Manual Filter* and the use of an *External Knowledge Source* (in this case, the Vidal pharmacovigilance knowledge source) execution steps are each responsible for 3% of the total computational cost. As expected, the meta-rules employed in the framework introduce a significant computational cost.

Finally, as regards the completeness, correctness, relevance and comprehension of the KB content, a validation study that took place with experts and test cases of various types was proven successful [53]. Specific suggestions for enhancement were made concerning the “wording” of the generated alerts and their explanation. The recommendations made from the experts were taken into account for developing the final version of the framework for further improvement. Interestingly, the study indicated also the necessity for employing context-sensitive strategies for CDSS delivery as elaborated in this work, and highlighted the importance and potential impact of the proposed framework in medication safety along this perspective. Quantitative features concerning the overall potential of the framework are expected to be extracted during its validation in realistic clinical settings that is under development.

4. Discussion

The presented knowledge framework is the result of knowledge engineering tasks that took place in a 3-year period. The main outcome is the development of the KB, which includes the architectural design and modeling part, as well as its implementation in a fully operational KBS. The KB effectively encapsulates source knowledge as regards ADE signals into an exploitable structure and in an appropriate electronic format. The framework is complemented with import/export mechanisms from knowledge sources to external systems, and mechanisms for querying, inference and KB contextualization.

The developed KB has been subject to a large number of upgrades and refinements, which were performed in parallel to the continuous updates of source knowledge. The current version adapts well to the particular knowledge sources considered and the requirements posed, providing a sound basis for delivering effective, contextualized, clinical decision support services. A positive outcome has been the successful implementation of the KB and the interfacing of the framework with external KBs and healthcare information systems. The design followed resulted in a generic/blueprint KB that can be easily instantiated with contextual information to construct contextualized KBs specifically tailored to the local environment.

From a technical perspective, it is interesting to note that the framework has been adopted to exploit decision support services via three distinct healthcare applications, i.e. a commercial CPOE system, a commercial EHR system, and a Web prototype that is independent of any healthcare information system, as presented by Bernonville et al. [54]. Although the knowledge framework is

mature, further development may be needed as feedback will be received while using the contextualized CDSSs that will be built upon and used in clinical practice. Furthermore, several limitations have been also recognized and many challenges for future work have been identified. In this regard, issues worth discussing as well as future research directions are presented in the following.

4.1. Source knowledge

The XML-based model employed to share knowledge between the knowledge discovery and the knowledge authoring teams was proven of paramount importance [32]. In essence, the evolution of this representation into a mature, rich and highly accurate (multiple XML documents based) structure, enabled the construction of appropriate mechanisms for analyzing, verifying and importing this type of source knowledge into existing KBSs, while resolving the introduced ambiguities. In this regard, the KB implementation timeframe was reduced significantly. Whenever not feasible (e.g. in the case of meta-rules), natural language was employed as the knowledge sharing mechanism instead.

However, there were cases where the delivered source knowledge introduced limitations due to (inevitable) simplifications applied. For example, the representation of laboratory examination results via binary variables (the value of which rely on “crisp” logic and generally-defined thresholds) constitutes a quite rough estimate introducing information loss. A more informative approach and representation (e.g. fuzzy logic oriented) would enhance the granularity of this information. Nevertheless, as remarked in Section 2.2, the proposed knowledge model is applicable in expressing rule-based knowledge for other clinical applications besides ADE prevention.

Similarly, the analysis of clinical observations concerning the patients could be quite an important knowledge source that was not elaborated in the scope of this work. These observations could provide further insights, potentially realized by employing semantic mining techniques. In addition, elaborating on time-dependencies derived from, for example, trend analysis could reveal time-related developments as conditions and ADE prodromes.

4.2. Knowledge representation

A significant challenge in the application domain of ADE prevention constitutes handling time-dependent clinical data and information. Several concepts have been included in the data model [31], such as *stay*, *duration_of_stay*, and *delay_drug*, and, thus, corresponding concepts have been defined in the knowledge model and implemented in the KB. In this regard, the knowledge model does not explicitly define time-dependent information [55], but rather implicitly encapsulates this time dimension for the implementation of ADE rules. Thus, a future challenge would be to encapsulate in the KB all the “pre-rules” which are associated with time variables and are now implicitly elaborated.

With respect to the CIG formalism adopted, as it typically relies on frame-based ontologies, it is evident that to some extent the reasoning capabilities and expressiveness offered are limited compared to those available in Description Logics (DLs) based formalisms. However, in a comparative study among frame-based ontologies and OWL (Web Ontology Language) [56], the closed-world assumption (CWA), i.e. what is not currently known to be true is false (negation as failure), that frame-based ontologies are based on is particularly favorable in the application context, as knowledge has to be explicit and complete, in order to avoid over-alerting in a CDSS for ADE prevention. On the other hand, the open-world assumption (OWA), i.e. lack of knowledge does not imply falsity (negation as unsatisfiability), is favorable in cases where DL reasoning is required to ensure logical consistency of

ontologies and in cases of applications in which classification is a paradigm for reasoning. Based on the knowledge engineering rationale of this work, the requirements posed differ from a DL-oriented scope. In particular, the procedural knowledge that the meta-rules imply, would definitely require an underlying model per se to support them. Meta-rules are effectively represented under the CIG formalism, which is procedural by nature.

4.3. Standardisation – sustainability – reusability

Standardization of the KB constitutes a challenging issue, as there are no widely used standards available for representing medical knowledge in particular; on the other hand, the attributes of openness and interoperability are important in modern healthcare information systems. In this regard, the KB has been partially (excluding the meta-rules that could not be explicitly encoded) represented in GELLO [48], which is a promising option for the time being, having its origins in the CIG knowledge engineering formalism. This effort was particularly significant so as to enable the accessibility and reuse of the accumulated knowledge potentially from other, third-party systems that conform to the above standard, advancing this way the exploitation potential. In addition, the representation of the KB as a Protégé frame-based ontology enables its straightforward editing and management using Protégé, a widely known, open-source, knowledge modeling tool.

The development of the KBS with Gaston and the proof-of-concept implementation conducted with Drools revealed that a significant weakness of each particular KBS is the dependence on its specific knowledge representation scheme that is employed to express the KB, as standardization of such a representation is a major issue. Nonetheless, while Gaston was the primary tool employed, experimentation with Drools enables us to argue that the deployment of the knowledge framework is feasible with other KBS platforms as well. In addition, the knowledge management mechanisms (e.g. import/update, interfaces with external knowledge sources) that have been elaborated contribute further in the framework's sustainability and extensibility.

As highlighted in Section 2.2, the proposed knowledge model introduces a significant reusability potential for other types of clinical applications elaborating on rule-based knowledge. The overall reusability of the proposed approach is further reinforced through the design that was followed for the majority of the knowledge framework components (Fig. 3). Specifically, the interfaces of the framework rely on XML-based specifications, which enable the necessary openness for populating the KB, updating source knowledge, accessing additional knowledge sources, and providing access to the services of the framework for diverse healthcare applications. The KB structure is capable of representing production rules like those elaborated in this work, exceeding this way the current application scope. The subcomponents of the *Inference Engine* concerning terminological reasoning and filtering may be used to elaborate on other types of rule-based knowledge. The contextualization-related components may be used in cases where rules are associated with meta-data denoting their statistical significance for the local application setting.

4.4. Inference logic

The issue of over-alerting [24], which is the most challenging one that needs to be addressed by the reasoning scheme encapsulated in the presented KBS, is currently being dealt with by a number of meta-rules handling contextual parameters [22]. Meta-rules are important for tailoring the KBS functionality, according to the clinical procedures and followed practices in actual settings. Such meta-rules, for example, are able to filter-out certain ADE signals based on a set of (statistically-oriented) rule meta-data

and heuristic practices. However, it has become evident that the issue of which (fired) rules are really applicable in each case and whether it is productive to present the respective ADE alerts to the user at a specific moment, is quite sophisticated and crucial at the same time. A future development would involve the design and implementation of a distinct and more intelligent layer incorporating such meta-rules that would be applied on ADE signals for this purpose. This layer should be considered as an additional knowledge-based component that can either be based on its own KB or integrated in an expanded version of the current KB. Such an approach would advance the inference logic according to which the KBS will operate, offering more sophisticated and effective functionality for the CDSS end-users.

4.5. Human factors and clinical procedures

One of the challenges for preventing ADEs is considering not only the medical parameters available at the time of drug management, but also the human factors related to increased possibility for errors during the prescribing–ordering–dispensing–administration (PDAC) medication chain. For this purpose, one of the identified knowledge sources was related to human factors and clinical procedures. Such knowledge was initially considered as a means to associate medication issues and other medical conditions with specific steps of the clinical practice and actions. The human factor analysis until now resulted in a set of recommendations for the CDSS design and functionality [34], which would increase the effectiveness of such systems in clinical practice. In this regard, the set of meta-rules was defined providing the logic according to which the ADE rules should be applied in practice in a particular context.

A potential field for future work that is worth exploring lies in developing knowledge discovery methods and the appropriate knowledge representation methods and inference mechanisms capable of: (a) capturing clinical procedures, e.g. related to drug administration and effect monitoring, (b) associating individual steps and actors of such procedures with the main ADE rules, and (c) introducing an additional layer of advanced rules which would be able to prevent a much wider set of errors by sensing human-factor related parameters.

5. Conclusions

A uniform, exploitable and sustainable knowledge-based framework has been designed to support the implementation of context-sensitive and advanced knowledge-based CDSSs for ADE prevention. The CIG formalism that was adopted as the main knowledge engineering endeavor has proven successful in the construction of the proposed framework. The context-sensitive layer introduced enables the adaptation of the KB in the local clinical environment (hospital, department, clinic), defining the applicability of the ADE prevention signals it incorporates according to their statistical significance, so as to enhance the delivery of targeted and effective decision support services. The results obtained as regards implementation, performance and validation aspects of the framework highlight its applicability in medication safety. The potential of the framework is currently under assessment in realistic clinical scenarios and settings.

Acknowledgments

The research leading to these results has received funding from the European Community's Seventh Framework Program (FP7/2007–2013) under Grant Agreement n° 216130 – the PSIP project.

The authors would like to thank Dr. Emmanuel Chazard (CHRU Lille, France) leading the knowledge discovery phase, Dr. Marie-Catherine Beuscart-Zéphir (Evalab, CHRU Lille, France) leading the knowledge elicitation phase, as well as all the PSIP partners for their cooperation and support in this work. In addition, the authors would like to thank Dr. Paul de Clercq (Medecs BV, Eindhoven, The Netherlands) for technically supporting parts of the framework related to Gaston.

References

- [1] Schachter M. The epidemiology of medication errors: how many, how serious? *Br J Clin Pharmacol* 2009;67(6):621–3.
- [2] Institute Of Medicine. Preventing medication errors. Washington, DC: The National Academic Press; 2007.
- [3] Morimoto T, Gandhi TK, Seger AC, Hsieh TC, Bates DW. Adverse drug events and medication errors: detection and classification methods. *Qual Saf Health Care* 2004;13(4):306–14.
- [4] Aronson JK. Medication errors: definitions and classification. *Br J Clin Pharmacol* 2009;67(6):599–604.
- [5] Nebeker JR, Hurdle JF, Hoffman J, Roth B, Weir CR, Samore MH. Developing a taxonomy for research in adverse drug events: potholes and signposts. *J Am Med Inform Assoc* 2002;9(Nov-Dec suppl):S80–5.
- [6] Amalberti R, Benhamou D, Auroy Y, Degos L. Adverse events in medicine: easy to count, complicated to understand, and complex to prevent. *J Biomed Inform* 2011;44:390–4.
- [7] Bates DW et al. Detecting adverse events using Information Technology. *J Am Med Inform Assoc* 2003;10(2):115–28.
- [8] Murff HJ, Patel VL, Hripcsak G, Bates DW. Detecting adverse events for patient safety research: a review of current methodologies. *J Biomed Inform* 2003;36(1–2):131–43.
- [9] Jha AK et al. Identifying adverse drug events: development of a computer-based monitor and comparison with chart review and stimulated voluntary report. *J Am Med Inform Assoc* 1998;5(3):305–14.
- [10] Brown JS et al. Early adverse drug event signal detection within population-based health networks using sequential methods: key methodologic considerations. *Pharmacoepidemiol Drug Saf* 2009;18(3):226–34.
- [11] Honigman B, Light P, Pulling RM, Bates DW. A computerised method for identifying incidents associated with adverse drug events in outpatients. *Int J Med Inf* 2001;61(1):21–32.
- [12] Trifirò G et al. Data mining on electronic health record databases for signal detection in pharmacovigilance. which events to monitor? *Pharmacoepidemiol Drug Saf* 2009;18(12):1176–84.
- [13] Morimoto T et al. Incidence of adverse drug events and medication errors in Japan: the JADE study. *J Gen Intern Med* 2011;26(2):148–53.
- [14] Bousquet C, Henegar C, Lillo-Le Louët A, Degoulet P, Jaulet M-C. Implementation of automated signal generation in pharmacovigilance using a knowledge-based approach. *Int J Med Inform* 2005;74(7–8):563–71.
- [15] Estacio-Moreno A, Toussaint Y, Bousquet C. Mining for adverse drug events with formal concept analysis. *Stud Health Technol Inform* 2008;136:803–8.
- [16] Ji Y et al. A distributed adverse drug reaction detection system using intelligent agents with a fuzzy recognition-primed decision model. *Int J Intell Syst* 2007;22(8):827–45.
- [17] Stephens S, Morales A, Quinlan M. Applying Semantic Web technologies to drug safety determination. *IEEE Intell Syst* 2006;21(1):82–6.
- [18] Kuperman GJ et al. Medication-related clinical decision support in computerized provider order entry systems: a review. *J Am Med Inform Assoc* 2007;14(1):29–40.
- [19] Rommers MK et al. Development of a computerised alert system, ADEAS, to identify patients at risk for an adverse drug event. *Qual Saf Health Care* 2010;19(6):e35.
- [20] Del Fiol G, Rocha BHSC, Nohama P. Modeling a decision support system to prevent adverse drug events. In: Proc of the 13th IEEE symp on CBMS; 2000. Available at: <http://dx.doi.org/10.1109/CBMS.2000.856887>.
- [21] Payne VL, Metzler DP. Hospital Care Watch (HCW): an ontology and rule-based intelligent patient management assistant. In: Proc of IEEE symp on CBMS; 2005. p. 479–84. Available at: <http://dx.doi.org/10.1109/CBMS.2005.64>.
- [22] Niès J et al. Information contextualization in decision support modules for adverse drug event prevention. *Stud Health Technol Inform* 2011;166:95–104.
- [23] Bernonville S et al. Implementation of a taxonomy aiming to support the design of a contextualised clinical decision support system. *Stud Health Technol Inform* 2011;166:74–83.
- [24] van der Sijts H, van Gelder T, Vulto A, Berg M, Aarts J. Understanding handling of drug safety alerts: a simulation study. *Int J Med Inf* 2010;79(5):361–9.
- [25] Beuscart R, McNair P, Brender J. The PSIP Consortium, Patient safety through intelligent procedures in medication: the PSIP project. *Stud Health Technol Inform* 2009;148:6–13.
- [26] Chazard E, Ficheur G, Bernonville S, Luyckx M, Beuscart R. Data mining to generate Adverse Drug Events detection rules. *IEEE Trans Inf Technol Biomed* 2011;15(6):823–30.
- [27] Schreiber G. Knowledge engineering. In: van Harmelen F, Lifschitz V, Porter B, editors. Handbook of knowledge representation. Elsevier; 2008.
- [28] Studer R, Benjamins R, Fensel D. Knowledge engineering: principles and methods. *Data Knowledge Eng* 1998;25(1–2):161–97.
- [29] Menzies T, van Harmelen F. Editorial: evaluating knowledge engineering techniques. *Int J Human-Computer Stud* 1999;51(4):715–27.
- [30] Ligèza A. Logical foundations for rule-based systems. Studies in computational intelligence, vol. 11, 2nd ed. Berlin Heidelberg: Springer-Verlag; 2006.
- [31] Chazard E et al. Detection of adverse drug events: proposal of a data model. *Stud Health Technol Inform* 2009;148:63–74.
- [32] Chazard E, Ficheur G, Merlin B, Serrot E, Beuscart R. Adverse drug events prevention rules: multi-site evaluation of rules from various sources. *Stud Health Technol Inform* 2009;148:102–11.
- [33] Abidi SSR, Cheah Y-N, Curran J. A Knowledge creation info-structure to acquire and crystallize the tacit knowledge of health-care experts. *IEEE Trans Inf Technol Biomed* 2005;9(2):193–204.
- [34] Marcilly R et al. Medication related computerized decision support system (CDSS): make it a clinicians' partner! *Stud Health Technol Inform* 2011;166:84–94.
- [35] Koutkias V et al. Constructing clinical decision support systems for adverse drug event prevention: a knowledge-based approach. *AMIA annu symp proc*; 2010. p. 402–6.
- [36] de Clercq P, Kaiser K, Hasman A. Computer-interpretable guideline formalisms. *Stud Health Technol Inform* 2008;139:22–43.
- [37] Peleg M et al. Comparing computer-interpretable guideline models: a case-study approach. *J Am Med Inform Assoc* 2003;10:52–68.
- [38] Wang D et al. Representation primitives, process models and patient data in computer-interpretable clinical practice guidelines: a literature review of guideline representation models. *Int J Med Inform* 2002;68(1–3):59–70.
- [39] Kumar A, Smith B, Pisanelli DM, Gangemi A, Stefanelli M. Clinical guidelines as plans—an ontological theory. *Methods Inf Med* 2006;45(2):204–10.
- [40] Sonnenberg FA, Hagerty CG. Computer-interpretable clinical practice guidelines. Where are we and where are we going? *IMIA Yearb Med Inform* 2006;145–58.
- [41] Peleg M. Guideline and workflow models. In: Greenes RA, editor. Clinical decision support: the road ahead. Elsevier; 2007.
- [42] Musen MA, Shahar Y, Shortliffe EH. Clinical decision-support systems. In: Shortliffe EH, editor. Biomedical informatics: computer applications in health care and biomedicine. Springer-Verlag; 2006.
- [43] de Clercq PA, Blom JA, Korsten HMM, Hasman A. Approaches for creating computer-interpretable guidelines that facilitate decision support. *Artif Intell Med* 2004;31(1):1–27.
- [44] Latoszek-Berendsen A, Talmon J, de Clercq P, Hasman A. With good intentions. *Int J Med Inf* 2007;76(53):S440–6.
- [45] van Oosterhout EMW et al. Three-layer model for the design of a protocol support system. *Int J Med Inf* 2005;74(2–4):101–10.
- [46] Grando A, Peleg M, Glasspool D. A goal-oriented framework for specifying clinical guidelines and handling medical errors. *J Biomed Inform* 2010;43(2):287–99.
- [47] Seyfang A, Kaiser K, Miksch S. Modelling clinical guidelines and protocols for the prevention of risks against patient safety. *Stud Health Technol Inform* 2009;150:633–7.
- [48] HL7/ANSI V3. GELLO: A common expression language, release 2; May 2010.
- [49] de Clercq PA, Hasman A, Blom JA, Korsten HMM. Design and implementation of a framework to support the development of clinical guidelines. *Int J Med Inform* 2001;64(2–3):285–318.
- [50] The Protégé ontology editor and knowledge acquisition system. <<http://protege.stanford.edu/>>.
- [51] Drools – JBoss community. <<http://www.jboss.org/drools/>>.
- [52] Sordo M, Ogunyemi O, Boxwala AA, Greenes RA. Software specifications for GELLO: an object-oriented query and expression language for clinical decision support. Decision systems group, Brigham & Women's Hospital, Harvard Medical School, Boston, MA, DSG-TR-2003-02; 2003.
- [53] Ammenwerth E, Hackl WO, Massari P, Darmoni S. Validation of completeness, correctness, relevance and understandability of the PSIP CDSS for medication safety. *Stud Health Technol Inform* 2011;166:254–9.
- [54] Bernonville S, Niès J, Gliese Pedersen H, Guillot B, Maazi M, Berg A-L, et al. Three different cases of exploiting decision support services for adverse drug event prevention. *Stud Health Technol Inform* 2011;166:180–8.
- [55] Adlassnig K-P, Combi C, Das AK, Keravnou ET, Pozzi G. Temporal representation and reasoning in medicine: research directions and challenges. *Artif Intell Med* 2006;38(2):101–13.
- [56] Wang HH et al. Frames and OWL side by side. In: Proc of the 9th int Protégé conf. Stanford University, California, USA; July 23–26, 2006.