Journal of Biomedical Informatics 45 (2012) 495–506

Contents lists available at SciVerse ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

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.

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 interoperating 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).

1532-0464/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.jbi.2012.01.007

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:

$$C_1\ AND\ C_2\ AND \ldots AND\ C_n \rightarrow E, \quad (1)$$

where $C_1, C_2, \ldots, C_n$ 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