COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE 123 (2016) 94–108

journal homepage: www.intl.elsevierhealth.com/journals/cmpb

Integrating HL7 RIM and ontology for unified knowledge and data representation in clinical decision support systems

Yi-Fan Zhang[a], Yu Tian[a], Tian-Shu Zhou[a], Kenji Araki[b], Jing-Song Li[a,*]

[a] *Engineering Research Center of EMR and Intelligent Expert System, Ministry of Education, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, College of Biomedical Engineering and Instrument Science, Zhejiang University, No. 38 Zheda Road, Hangzhou 310027, China*

[b] *Department of Medical Informatics, Miyazaki University Hospital, 5200 Kiyotakecho Kihara, Miyazaki-city, Miyazaki 889-1692, Japan*

ARTICLE INFO

Article history:
Received 29 April 2015
Received in revised form
21 September 2015
Accepted 23 September 2015

Keywords:
Semantic Web Technologies
Ontology
Knowledge base
CDSS
HL7 RIM

ABSTRACT

Background and objectives: The broad adoption of clinical decision support systems within clinical practice has been hampered mainly by the difficulty in expressing domain knowledge and patient data in a unified formalism. This paper presents a semantic-based approach to the unified representation of healthcare domain knowledge and patient data for practical clinical decision making applications.

Methods: A four-phase knowledge engineering cycle is implemented to develop a semantic healthcare knowledge base based on an HL7 reference information model, including an ontology to model domain knowledge and patient data and an expression repository to encode clinical decision making rules and queries. A semantic clinical decision support system is designed to provide patient-specific healthcare recommendations based on the knowledge base and patient data.

Results: The proposed solution is evaluated in the case study of type 2 diabetes mellitus inpatient management. The knowledge base is successfully instantiated with relevant domain knowledge and testing patient data. Ontology-level evaluation confirms model validity. Application-level evaluation of diagnostic accuracy reaches a sensitivity of 97.5%, a specificity of 100%, and a precision of 98%; an acceptance rate of 97.3% is given by domain experts for the recommended care plan orders.

Conclusions: The proposed solution has been successfully validated in the case study as providing clinical decision support at a high accuracy and acceptance rate. The evaluation results demonstrate the technical feasibility and application prospect of our approach.

* *Corresponding author*. Tel.: +86 571 87951564; fax: +86 571 87951564.
E-mail addresses: yifanzhang@zju.edu.cn (Y.-F. Zhang), ty.1987823@163.com (Y. Tian), zts@zju.edu.cn (T.-S. Zhou), taichan@med.miyazaki-u.ac.jp (K. Araki), ljs@zju.edu.cn (J.-S. Li).

http://dx.doi.org/10.1016/j.cmpb.2015.09.020

1. Introduction

A clinical decision support system (CDSS) is a computer-based information system developed specifically for clinical decision-making, in which the characteristics of an individual patient are matched to a computerized clinical knowledge base, and patient-specific assessments or recommendations are then presented to the clinician or the patient for a decision [1]. A large body of evidence suggests that CDSSs can be helpful in improving clinical practice [2,3]. However, to this day, CDSSs have not found wide use outside of a handful of mostly academic medical centers, and their impact on patient outcome is marginal [4]. A major impediment to their wide adoption is the lack of standard knowledge representation models and data models [5].

The knowledge in CDSS is primarily derived from healthcare process models such as clinical guidelines (CGs) and care plans (CPs). CGs are systematically developed statements to assist practitioners' decision making about appropriate healthcare for specific clinical circumstances [6], while CPs can be viewed as detailed CGs that typically include locally tailored care activities recommended by one or more CGs with explicit procedures, schedules, conditions, and goals [7]. Several process-oriented representation languages have been proposed to formulate these models, such as GLIF, GUIDE, Asbru, EON, PRODIGY, and PROforma [8]. Although these languages provide a complete set of control flow structures to describe the procedural knowledge of healthcare processes, most of them are designed with limited expressivity of institution-specific organizational knowledge and domain-specific medical knowledge, which are equally important for their practical execution where a highly complex and changeable clinical context frequently occurs.

A number of studies suggest the use of ontologies to encode CGs or CPs [9]. The ontology, originally defined as "a formal, explicit specification of a shared conceptualization" [10], can formally define the terms, relations, and constraints of commonly agreed upon concepts that constitute different aspects (procedural, organizational, and medical) of CDSS knowledge. The static knowledge in the ontology can also be dynamically updated and utilized by an inference engine when combined with rules and fact data. However, work in this area is faced with the challenge brought by the heterogeneity of healthcare knowledge and the lack of a standard paradigm for ontology engineering, leading to a proliferation of ontologies that themselves create obstacles to integration [11].

Another critical factor that impacts the successful implementation of CDSSs is their seamless integration into the organizational workflow. Despite the plentiful patient data standards available to date, including standard models such as openEHR [12] and CEN/ISO 13606 [13] and standard terminologies such as those compiled in UMLS [14], there is little consensus on their usage within CDSSs. A recent review from Ahmadian et al. [15] reveals that the diversified data models and terminologies adopted by CDSSs and healthcare information systems lead to problems in semantic interoperability. Therefore, a generic data model to align these standards is required.

The HL7 version 3 Reference Information model (RIM) is such a model that covers all aspects of healthcare information [16]; it is compatible with existing data standards and knowledge models, and thus can serve as the foundation for information integration across platforms and systems. There have been some ongoing efforts in RIM modeling and application, most of which focus on ontological engineering of the RIM [17,18,19], clinical data interoperability [20,21] and domain knowledge representation [22,23], while few seek to develop and validate the RIM for CDSS knowledge and data integration [24].

To meet the abovementioned challenges to CDSS implementation, this paper proposes a novel approach to the unified representation of CDSS knowledge and data based on Semantic Web Technologies and HL7 RIM. A semantic healthcare knowledge base has been developed, whose core components include:

- A generic ontology constructed based on HL7 RIM to represent healthcare domain knowledge and clinical data.
- A repository of semantically encoded rules and queries for dynamic and personalized clinical decision making.

We validate our solution in a prototype semantic CDSS to provide patient-specific recommendations on the management of inpatients with type 2 diabetes mellitus (T2DM). Ontology-level and application-level evaluations are conducted to demonstrate its technical feasibility and application prospect.

Our work advances the state of the art by (1) developing an ontological framework for multi-faceted CDSS knowledge integration, sharing, and reuse, (2) designing a semantic CDSS environment for dynamic and personalized clinical decision making, and (3) bridging the gap between CDSSs and heterogeneous healthcare information systems to support the practical use of CDSSs.

2. Methods

We designed a four-phase knowledge engineering cycle to develop the semantic healthcare knowledge base (SHKB), as shown in Fig. 1. The knowledge being modeled is under continual maintenance and kept up to date through systematic capturing, modeling, operationalizing and evaluation. We describe each phase in detail in the following Sections 2.1–2.4.

2.1. *Knowledge acquisition*

We studied the method of cognitive analysis by Patel et al. [25] and summarized a three-step process for knowledge acquisition: knowledge identification, knowledge extraction, and knowledge authentication. In knowledge identification, we generalize the healthcare process knowledge from three types of knowledge sources:

- Published medical literature on CGs and CPs, standard terminologies and clinical information models.
- Patient data from healthcare information systems, such as the electronic medical record (EMR).