Journal of King Saud University – Computer and Information Sciences (2014) **26**, 69–78

King Saud University

Journal of King Saud University – Computer and Information Sciences

www.ksu.edu.sa
www.sciencedirect.com

ORIGINAL ARTICLE

A distributed clinical decision support system architecture

Shaker H. El-Sappagh [a,b], **Samir El-Masri** [c,*]

[a] *College of Science, King Saud University, Saudi Arabia*
[b] *Faculty of Computers and Information, Minia University, Egypt*
[c] *Department of Information Systems, College of Computer and Information Sciences, King Saud University, Saudi Arabia*

Received 8 September 2012; revised 22 February 2013; accepted 20 March 2013
Available online 8 April 2013

KEYWORDS

Data Mining;
Knowledge management;
Clinical decision support systems (CDSS);
Electronic health record;
Health informatics

Abstract This paper proposes an open and distributed clinical decision support system architecture. This technical architecture takes advantage of Electronic Health Record (EHR), data mining techniques, clinical databases, domain expert knowledge bases, available technologies and standards to provide decision-making support for healthcare professionals. The architecture will work extremely well in distributed EHR environments in which each hospital has its own local EHR, and it satisfies the compatibility, interoperability and scalability objectives of an EHR. The system will also have a set of distributed knowledge bases. Each knowledge base will be specialized in a specific domain (i.e., heart disease), and the model achieves cooperation, integration and interoperability between these knowledge bases. Moreover, the model ensures that all knowledge bases are up-to-date by connecting data mining engines to each local knowledge base. These data mining engines continuously mine EHR databases to extract the most recent knowledge, to standardize it and to add it to the knowledge bases. This framework is expected to improve the quality of healthcare, reducing medical errors and guaranteeing the safety of patients by helping clinicians to make correct, accurate, knowledgeable and timely decisions.

1. Introduction

Healthcare faces multiple problems, including high and rising expenditures, inconsistent quality and gaps in care and access. For this reason, healthcare services represent a major portion of government spending in most countries (Canadian Institute for Health Informatics, 2012). Healthcare information technologies, especially EHRs, have been thought to be a possible solution to healthcare problems. EHRs help administrators, physicians, nurses, researchers and healthcare personnel. An EHR provides a complete, integrated and consistent view of patient conditions. However, the volume of data is considerable and is increasing continuously. Healthcare personnel must take all of the patient medical history into consideration; these personnel also need to connect this information together and receive advice from domain experts. This large amount of data cannot benefit physicians without having an automated system. The system can analyze these data, connect it, integrate it with knowledge from a domain expert, and search for a

* Corresponding author. Tel.: +966 535376100.
E-mail addresses: sker@ksu.edu.sa (S.H. El-Sappagh), selmasri@ksu.edu.sa (S. El-Masri).

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

1319-1578 © 2013 Production and hosting by Elsevier B.V. on behalf of King Saud University.
http://dx.doi.org/10.1016/j.jksuci.2013.03.005

needed knowledge – if possible – in other connected systems. This system is a clinical decision support system (CDSS).

CDSSs are computer applications that assist practitioners and healthcare providers in decision making through timely access to electronically stored medical knowledge (Giannopoulou, 2008), to improve the practitioners' medical practice. A CDSS interacts with practitioners and electronic medical record systems to receive the patient data as input and provides reminders, alerts, or recommendations for patient diagnosis, treatment, long-term care planning, and other aspects. A clinical decision support system requires access to healthcare data and knowledge that are stored in data and knowledge bases. In this paper, we attempted to build a complete architecture for this system. The proposed model will take an order and an initial diagnosis from healthcare personnel and will provide a decision support in an understandable form based on existing knowledge. It will integrate off-line standardized knowledge bases from domain experts and Clinical Practice Guidelines (CPG) knowledge with online knowledge that is extracted continuously from EHR databases and provides applicable decision support. This paper is organized as follows. Section 2 discusses related work. Section 3 explains the research problem. In section 4, we define a CDSS. The proposed framework for a CDSS is discussed in section 5. The conclusion is shown in section 6.

2. Related studies

In this section, we go through some background about essential principal aspects of health informatics that are related to CDSSs. EHR (Yina, 2010), EHR standards, data mining and artificial intelligence, service oriented architecture and knowledge representation are strongly related to CDSSs.

2.1. EHR standards

Many organizations provide EHR standards that standardize structuring, implementation, sharing, integration and interoperability in an EHR environment. Some of the standards are ISO (International Organization for Standardization, 2012), CEN (The European Committee for Standardization, 2012), CFR (The World Health Organization, 2012), ASTM (ASTM International, 2012), HL7 (The Health Level Seven, 2012), NEMA (National Electrical Manufacturers Association (NEMA), 2012), and ONCHIT (US Department of Health and Human and Services, 2012). In addition, coding systems are critical to build a shared EHR because the new environment connects each heterogeneous system with different terminologies. Some organizations that provide these standards are Regenstrief (The Regenstrief Institute, 2012), AMA (The American Medical Association, 2012), IHTSDO (International Health Terminology Standards Development Organization, 2012), CMS (US Department of Health & Human and Services, 2012), and WHO (The World Health Organization, 2012). These organizations provide standards for encoding healthcare data and knowledge.

2.2. Data mining and artificial intelligence (AI)

Applying data mining and AI techniques on EHR data provides many opportunities for improving delivery, efficiency, and effectiveness of health care (Ramakrishnan et al., 2010; Giannopoulou, 2008), such as operations management, preventive healthcare, chronic disease treatment and prevention, association analysis, evidence-based treatment, and population tracking. If CDSS depends only on the Knowledge Base (KB) that is derived from a knowledge expert, then it will be inactive and not applicable. There are two other sources for knowledge. The first source is CPG, which publishes free text guidelines. These guidelines are created by using many methods, such as Systematic Reviews and Meta-analysis. The modeling CPG knowledge is formulated in rules that use many methodologies (Peleg et al., 2003). The other source is the application of data mining techniques on EHR data. EHRs contain a very large and historical dataset that changes continuously and contains useful hidden knowledge. As a result, data mining and AI services should be embedded into the active CDSS to continuously update its knowledge base by the most recent patterns.

2.3. Knowledge representations in the medical domain

Because there are many sources and uses for medical knowledge, many methodologies and standards for representing medical and healthcare knowledge are integrated. Clinical workflows (clinical guidelines) are used to represent human-based medical knowledge through rule-based or flow-based guideline techniques. Furthermore, mined knowledge can be automatically extracted from EHR databases through data mining and AI techniques, to be incorporated into human-generated knowledge that enhances their decision-making processes.

Both types of knowledge can be represented as logical conditions, rules (Kuo and Fuh, 2011), graphs/networks, or structural representations (Kong et al., 2008). Predictive Model Markup Language (PMML) (Data Management Group (DMG), 2012) and GLIF (Guideline Interchange Format (GLIF), 2012) are examples of knowledge representation languages that are used to acquire and integrate knowledge. Additionally, there are many tools for knowledge acquisition and representation, such as Unified Medical Language System (UMLS) Bodenreider, 2004, Protégé (Protégé Official Web Site, 2010; Chen et al., 2011), GLARE (Terenziani et al., 2004), PROforma (PROforma, 2010) and ASBRU (Open clinical, 2010).

2.4. Service oriented architecture (SOA)

A SOA has been widely adopted to solve the interoperability of the involved heterogeneous distributed EHR systems (Hahn et al., 2010; Maciel and David, 2007). This architecture plays a key role in the integration of heterogeneous systems by means of services that represent different system functionality, independent of the underlying platforms or programing languages, and interacts via message exchanges. *Web services* also play a critical role in systems' interoperability. Web services technology is defined as a systematic and extensible framework for application-to-application interactions that is built on top of existing web protocols. These protocols are based on XML (World Wide Web Consortium, 2012) and include: Web Services Description Language (WSDL) to describe the service interfaces, Simple Object Access Protocol (SOAP) for communication between web services and client applications, and

Universal Description, Discovery, and Integration (UDDI), to facilitate locating and using web services on a network (Bacon and Moody, 2002).

2.5. Knowledge representation techniques

Knowledge is a description of the world. Knowledge representation is the way that knowledge is encoded, which can be a methodology for transforming knowledge from one form to another. In the context of our research, knowledge representation means transforming the medical domain's knowledge into clinical practice guidelines from a textual form to a computerized form. Different types of knowledge require different types of representation, such as Logic, Rules, Frames or Semantic Nets (Sanders et al., 2000; Kuo and Fuh, 2011; van Harmelen and Fensel, 1995).

3. The research problem

Because health information technology, especially EHR, can allow clinicians to have a real-time access to complete patient data, the involved systems help them to make accurate clinical decisions (The Regenstrief Institute, 2012). To make an accurate decision, clinicians must consider EHRs, clinical databases and clinical knowledge bases (which could be in different domains). EHRs are distributed and are very large; they contain the complete history of patient health data. Clinicians must collect, compare and analyze data from these sources and make a timed decision. This information overflow could cause physicians to disregard vital information (EHR data, hidden knowledge in clinical databases and EHRs, and knowledge bases) and could make it take a long time to make a correct decision. Therefore, they need an automated system that helps in collecting, calculating and analyzing all of the available data and helps in making decisions. This type of system is called a CDSS.

The proposed model will solve the problem of having multiple healthcare providers; each provider has its local and large EHRs, knowledge base, and clinical databases. The model contains a set of knowledge bases (one in each site or hospital) for extracted off-line information from domain experts. If a CDSS depends only on these knowledge bases, it would be inactive and would become not applicable over time. The solution is to continually update these knowledge bases, to make the CDSS more active and useful. At each healthcare site, new knowledge will be discovered and added to the knowledge base from:

(1) New expert knowledge discovered by research,
(2) A data mining engine connected to local EHRs and clinical databases that extracts up-to-date knowledge.

This action will make CDSS more active by including the most recent knowledge from active databases. Because a knowledge base must be in a specific domain, such as heart disease, the proposed framework will be distributed with co-operative and integrated knowledge bases. Each knowledge base will be in a specific domain. When a patient visits the hospital, the CDSS will build a patient profile from the patient medical history and current diagnosis, and it will use its local knowledge base to make decisions.

If a CDSS cannot make a decision by using its local knowledge base, then it can send some data from a patient profile to another site's inference engines, to consult with its specialized knowledge base. Other sites will respond by some knowledge that could help the CDSS to make more accurate decisions about the patient case.

4. Clinical decision support system

CDSSs are interactive computer programs that are designed to assist physicians and other health professionals (Gamberger et al., 2008). They help in drug prescriptions, diagnosis and disease management, to improve services and reduce costs, risks and errors [69]. The CDSS can check for patient drug allergies, compare drug and laboratory values, evaluate the potential for drug–drug interactions, suggest drug alternatives, block duplicate orders, suggest drug doses, routes, and frequencies and provide recommendations. In addition, a CDSS can provide clinical knowledge and best practice standards and guidelines for non-expert physicians. A CDSS must be integrated with EHR and CPOE systems, which are connected to other HISs (e.g., laboratory, radiology, billing systems). The basic components of a CDSS include a medical knowledge base and an inference mechanism (usually a set of rules that are derived from experts and evidence-based medicine) and are implemented through medical logic modules based on a language such as Arden syntax (The Arden Syntax for Medical Log) or using an artificial neural network, as in Fig. 1 (Aleksovska-Stojkovska and Loskovska, 2010).

A CDSS provides recommendations that are based on the available patient-specific data (EHR) and medical facts (knowledge bases). CDSSs have 10 levels of automation, which range from L1, in which all the decisions are made by humans, to L10, in which the computer makes all the decisions.

The EHR is continuously updated; thus, the knowledge bases must be continuously updated by discovered knowledge from domain experts and discovered knowledge from EHRs and clinical databases.

5. Proposed CDSS framework

We assume that EHR architecture and connectivity exists, and we will integrate the distributed CDSS architecture with it. The interoperability problems are outside of our focus. The

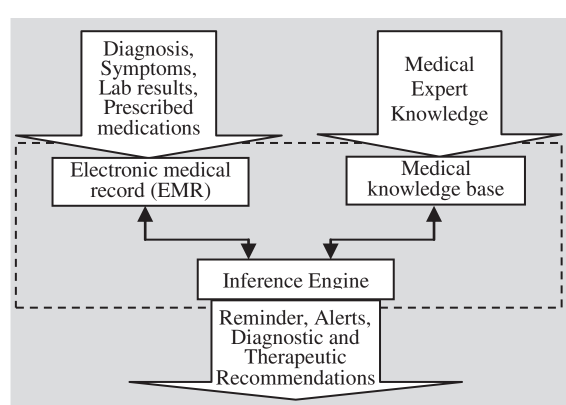

Figure 1 General model of CDSS.

interfacing of individual components of our framework is assumed to be possible. The proposed architecture of the CDSS is independent; it does not depend on and does not affect the architecture of the EHR or HIS. Moreover, the architecture is scalable. We can add any number of knowledge bases, EHRs, or clinical databases to the architecture using available standards and technologies.

Previous CDSSs were a separate system from the healthcare systems. In this way, these systems will require a physician to manually activate them, log in, and re-enter redundant patient data. This process will make the CDSS not applicable and waste time. Additionally, the CDSS will depend on the entered data, which could be inadequate or contain errors. The needed CDSS is directly integrated with the healthcare system's CPOE component. It will be activated automatically, collecting the needed data from the patient order and EHR history, asking for unknown parameters, and making recommendations on time. Fig. 2 (Kazemzadeh and Sartipi, 2005) shows the three phases in the decision-making process.

Phase 1 (knowledge preparation) uses data mining techniques to extract knowledge from electronic healthcare data and to store it in a knowledge base. Techniques such as classification, clustering and association can discover very important patterns that enrich the knowledge base. Knowledge can come from other sources, such as domain expert experience and clinical practice guideline knowledge.

Phase 2 (knowledge interoperation) takes the patient data that require decision making and translates it into standard XML form (CDA) and makes PMML encoding of the knowledge from the knowledge base (KB). In this way, both the knowledge base and the patient data have the same format, which enhances the knowledge interpretation.

Phase 2 (Knowledge interpretation) takes the previous standardized data and knowledge and makes decisions. This reasoning phase takes the patient current and historical data and searches in the knowledge base for the most appropriate recommendations. Fig. 3 shows our proposed CDSS framework, which will operate as follows:

5.1. Knowledge base building

In our framework, the first step is to build the initial KBs. This task is performed in each hospital by building a local knowledge base that is specialized in the hospital's domain. Constructing KBs of the CDSS is a crucial task that determines the success of the CDSS in general (Kim et al., 2008; Koutkias et al., 2012). The goal is to collect the medical knowledge from the relevant sources (domain experts, EHR databases, and clinical practice guidelines), systematize it and represent it in a formal human-understandable and a computer-interpretable manner (Sari et al., 2012). In this framework, to populate the standard XML KBs, we use three services, which are KEM, DME and PEM. The following sections discuss the services in detail.

1- *Knowledge Extraction Module (KEM)*, which is responsible for extracting knowledge from domain experts or formulating the knowledge that is contained in Clinical Practice Guidelines (CPG). There are many ways to represent this knowledge. This module can represent the extracted knowledge by using production rules, semantic networks/web, frames, decision tables, decision trees or conceptual graphs (Tanwar et al., 2010a,b; Chakraborty, 2008; Priss, 2010; Sittig et al., 2010). The knowledge structure can be in a relational database, an XML database or an ontology (Do, 2008; Iqbal et al., 2011; Riaño et al., 2012; Tenórioa et al., 2011; Bohada et al., 2012). Many tools might be used to extract knowledge from domain experts or CPGs, such as Protégé (Peleg et al., 2003; The Protégé Ontology Editor and Acquisition System, 2012). Knowledge bases and patient data are standardized by using available technologies such as HL7 RIM, which standardizes the patient data elements (The Health Level Seven, 2012), the medical ontologies and the terminologies such as SNOMED CT, which standardizes the data fields' semantic, names and values (The International Health Terminology

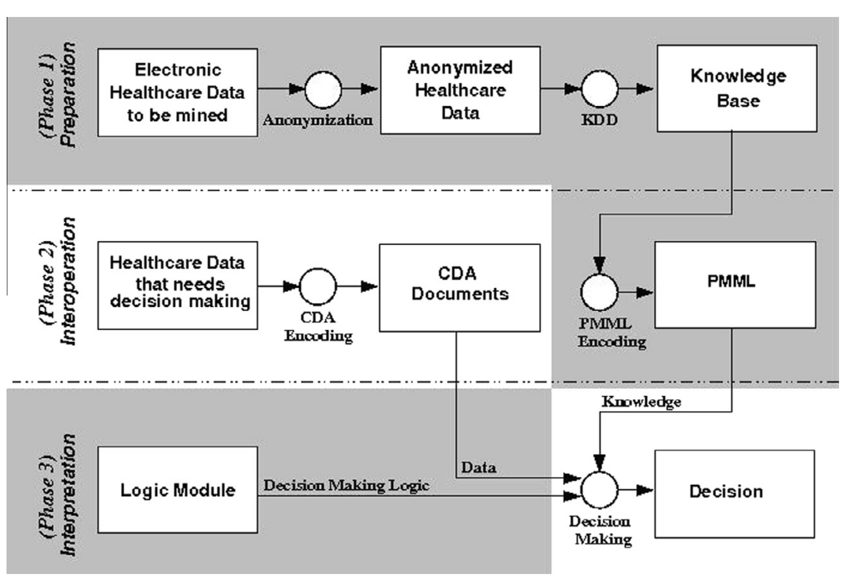

Figure 2 Health care knowledge management framework. The shaded areas designate off-line parts.

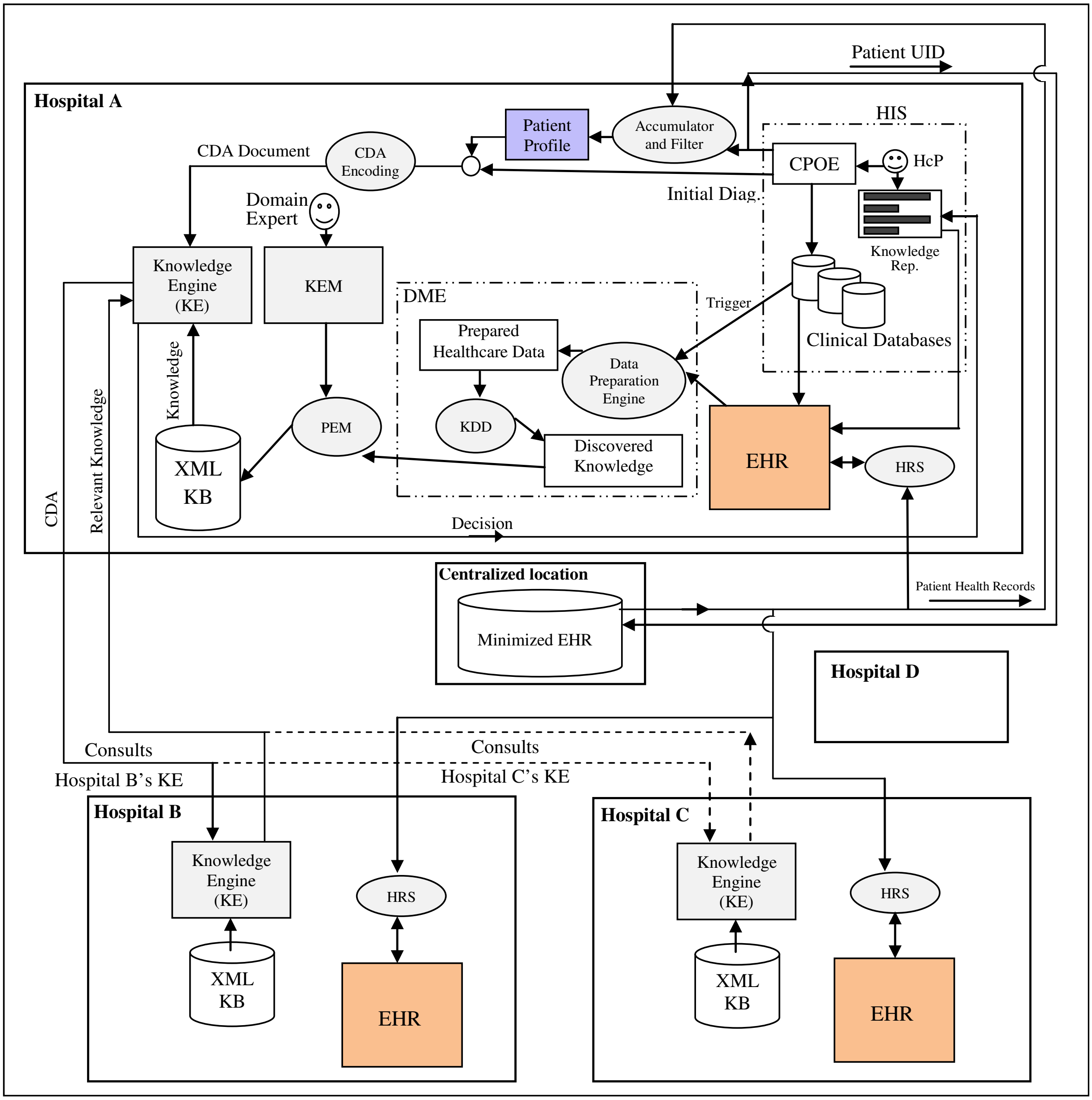

Figure 3 Proposed distributed CDSS architecture.

Standards Development Organization (IHTSDO), 2012). For knowledge representation, we will use Case Based Reasoning (CBR) because it is easy to extract knowledge from experts in the form of cases. Additionally, we will use SNOMED CT for standardizing the clinical terminologies in the EHR and the knowledge base to achieve interoperability.

2- *Data Mining Engine (DME)*, which is responsible for mining the EHR database. Data are collected, preprocessed, transformed, integrated, and mined to extract useful patterns (Ramakrishnan et al., 2010; Sartipi et al., 2007; Erraguntla et al., 2012; Liao et al., 2012; Wanga et al., 2010; Raza Abidia, 2002). The resulting mined knowledge is encoded and stored in a knowledge base, to be incorporated into the CDSS process. This module performs classical data mining operations, such as data selection, data preparation, data transformation and integration, data mining, and evaluation and interpretation, as shown in Fig. 4 (Kazemzadeh et al., 2010). Because EHR contains many types of data, such as special data, temporal (historical), social and clinical data, it is a very rich environment for the application of data mining. For example, using social data such as father, mother and family data of the patient helps in finding chronic diseases. In addition, using spatial data can help to find patterns that can help when classifying patients geographically. We will use classification, association and clustering techniques to enrich the knowledge base of CDSSs.

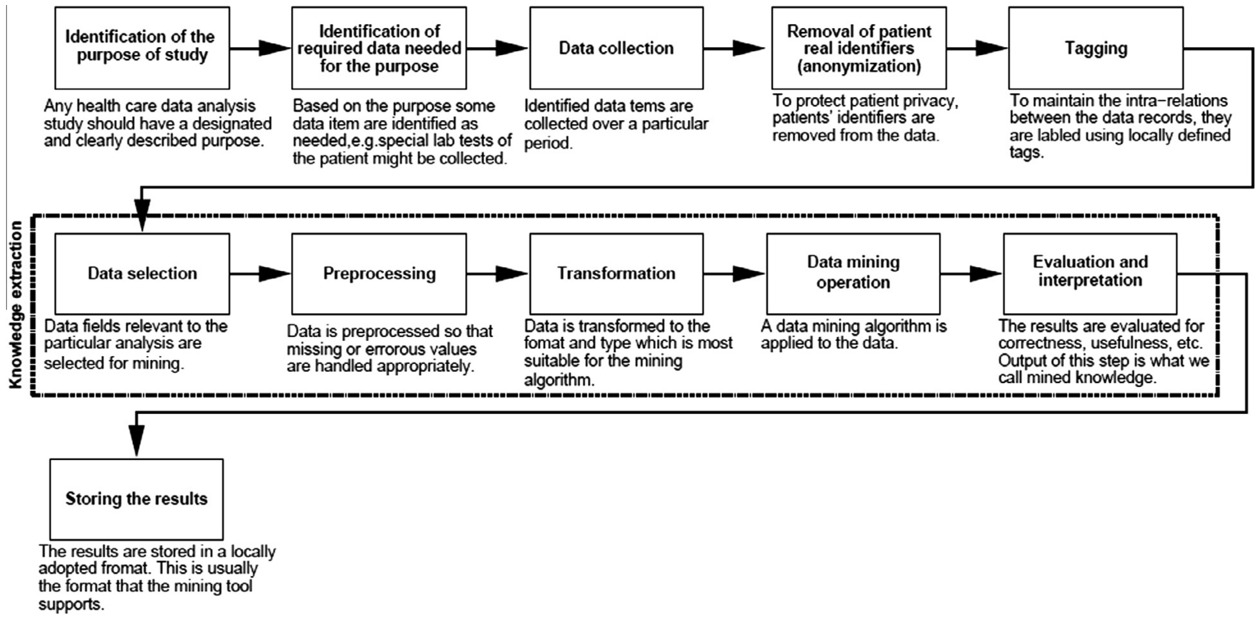

Figure 4 Different steps in KDD process (Resa et al., 2010).

3- *PMML Encoding Module (PEM)*, which employs PMML standards to encode the generated knowledge from domain experts and data mining to make standard XML-based documents. The generated knowledge will achieve the interoperability goals between the knowledge that was discovered from different HISs and the knowledge that was extracted from domain experts. It provides a language that encodes various types of data mining models, including clustering, regression, classification and association rule models. The XML schema for each document describes the input data items, data mining algorithm specific parameters, and the final mining results (Data Management Group (DMG), 2012).

There are many challenges in constructing and maintaining the knowledge bases. First, for the specific domain of one hospital, the KB is built from the domain experts' knowledge off-line. KEM can use a variety of methods and techniques to build the knowledge base (Kong et al., 2008). KEM then passes the knowledge to a PMML encoding module to translate it into XML form. In each hospital, CDSS will have specialized KBs according to the field of the hospital, and the distributed framework will make these KBs co-operative.

Assuring that KBs are up to date is critical to make CDSSs active and continuously learn and therefore applicable, and this task can be achieved by the following:

(1) Continuously update KBs by new domain expert or research knowledge,
(2) Apply data mining techniques and algorithms on local clinical databases and EHRs to discover hidden and non-trivial patterns and to update the KB with these results.

In this way, the CDSS will provide the most up to date and the most applicable knowledge. DME has two processes:

(1) The Data Preparation Engine, which identifies the task-relevant data from HER databases after triggers from their sources to apply a data mining process, removes the healthcare data attributes that can identify a patient or reveal their private data (Anonymization) and performs data selection, cleaning and transformation.
(2) The KDD, which performs the actual data mining operations. We do not differentiate between different implementations and algorithms of data mining techniques because their results are representable by the general concepts of the corresponding data mining type. For example, different association rule discovery algorithms take different approaches when extracting the frequent item sets, and they opt to choose different measures to exclude intermediary sets and hence prevent explosion in the results set. Some can refine the set based on standard constraints of support and confidence; others can apply additional constraints on the size of the rules' antecedents and consequents. Finally, the results are assessed in terms of usefulness, validity, and understandability.

An EHR is an important source for medical knowledge; it contains a longitudinal and history of patient clinical and diagnostic data. This aspect makes EHR a good place for applying data mining and AI techniques. Additionally, EHR attributes are selected carefully, which adds another advantage. This process is continuous because an EHR is updated continuously. Any update to an EHR will trigger the DME to discover new knowledge and then pass it to the PEM to standardize and store it in the KB.

Another source of knowledge is clinical databases because they contain detailed data about patients. The DME is triggered to search for new knowledge in the updated databases, as with HER. In this way, we assure that a KB contains the complete, most recent, accurate and applicable knowledge.

The domain expert knowledge and the data mining discovered knowledge are passed to the PEM module to be standardized in the XML format and stored in a knowledge base.

5.2. CDSS supporting CPOE

Healthcare personnel use the CPOE for prescriptions (Kuperman et al., 2007). Previously, Health Information Systems (HISs) depended on paper-based prescriptions and/or unstructured notes in a separate system (order entry system). In addition, the order entry system collected only administrative data and not medical, clinical or diagnostic data. The required system uses an electronic prescribing system, which allows the writing of e-prescriptions. Additionally, human errors are expected when writing the prescription. With the existence of a CDSS that is integrated with a CPOE system, the CDSS will not only provide recommendations for treatment but also can check for errors or shortages of data and notify physicians before proceeding with decision support. There are many methodologies for building a user interface for a CPOE. This interface could be a series of questions and answers (Xiao et al., 2010). Another methodology uses standard paper-based forms to build data entry templates and adds features that are relevant to decision support. Web-based order entry forms can also be used. In our system, the healthcare professional will collect all of the patient current diagnostic data and pass it to the Accumulator and Filter Service, which merges it with all of the patient historical data, as shown in Fig. 3.

5.3. Framework execution processes

After building knowledge bases, the CDSS is now ready to guide and help healthcare personnel. The system is implemented as a collection of services (SOA) that cooperate to achieve the system's functionality. We will discuss the details of the system in one hospital and the links or co-operations with other CDSS systems in other hospitals. The execution of this framework will work as follows:

1. Hospital Information System (HIS) component contains clinical databases that store patient conditions and a CPOE that allows healthcare personnel to interact with the CDSS. In an on-line operation, Healthcare personnel use the CPOE to enter a patient Universal ID (UID), which identifies the patient nation-wide. Moreover, they enter the patient current and initial diagnoses (i.e., healthcare data that require decision making).
2. UID passes to HRS (History Retrieval Service) in the local hospital and travels via a secure network channel to all of the hospitals. Each HRS in each hospital is a service that checks whether the patient has an EHR in its own hospital or not according to the UID. The HRS will perform a simple search for a patient with ID = UID. If the patient has no record, then the service returns a message that indicates that there is no record; otherwise, there are many methodologies for implementing the service to retrieve patient records. These methodologies could be implemented to retrieve the last N visits, visits within a specific period, and a specific disease's related data, according to the system requirements.
3. The process of spreading the UID to all of the hospitals is time consuming. We can solve this problem by building a centralized database that contains a minimized nation-wide EHR (the Minimized EHR component). This EHR is a small centralized database that contains only the patient UID, the visited hospitals, and the most critical patient diseases and diagnoses for all of the patients in the country.
4. The patient UID will be sent to this centralized database. This database will return the patient list of visited hospitals, if it exists, and the most critical patient medical data. These hospitals are the only hospitals that would be able to return the patient records. For example, in Fig. 3, hospital D was not visited by the patient, which means that the patient has no records in hospital D; thus, it will not be returned by the centralized database and will not be included in the process.
5. The returned records, which contain patient history, will be collected and filtered by Accumulator and Filter services. This service will produce the patient profile (the current case that requires decision support). It will remove redundancies, conflicts and inconsistencies from the patient historical data. In other words, it will make data preparation for the next step.
6. The patient profile is integrated with the current diagnosis and is entered into a CDA Encoding service. This service will standardize the patient medical and diagnoses data by encoding it into a standardized XML-based HL7 version 3.0 CDA (The Health Level Seven, 2012; Yuwen and Yang, 2009). We used standard CDA to achieve semantic data interoperability between the patient data that require clinical decisions and the PMML-encoded knowledge (Kazemzadeh et al., 2010).
7. The encoded PMML knowledge from local KB and CDA documents from patient profiles provides the interoperability of knowledge and data in our framework, in the sense that CDSS will be independent of the proprietary data format of the involved healthcare providers. Next, we have a complete view of a patient current and previous conditions.
8. The encoded patient profile enters input into the local Knowledge Engine (KE), which makes inferences to diagnose and determine the correct medicines, as discussed in section 4 (Chen et al., 2011). The inference engine will depend on the knowledge base format. For example, it can perform reasoning with logic, with rules, with frames or by using case-based reasoning. In this model, the knowledge base is structured on cases Ahmed et al., 2012; thus, an integration of case-based reasoning and rule-based reasoning is the best choice for inference.
9. KE can be programed by any AI methodology. It can access, query, and interpret data and knowledge that flow from CPOE and KB, respectively. Decision making is conducted in three main steps, retrieving the correct data fields from the data source (CDA); applying the knowledge models to the data; and eventually taking an action or a set of actions that is based on the results of the application. For example, if a module was

invoked at a decision step in a guideline, it could branch to a specific path; alternatively, it might simply display the results in the form of a reminder or an alert.

10. According to the complexity of the problem and according to the specialization of KEs, KE may need to consult another site's KE on its problem, if it has a shortage of available knowledge or if it is not specialized in this problem. It will send the CDA or specific fields from its site to all or a set of other sites that use the same technologies, interfaces, standards, services, and terminologies. All of the helping KEs determine the relevant knowledge and send it back to the requesting KE.
11. In this way, the KE will make a decision that is based on the initial physician diagnosis, the EHRs, and the knowledge from its local KB and other KBs. Additionally; it will use a KB that contains the most recent knowledge. In this way, we ease the process of developing KBs because each KB will be specialized in a specific domain and KEs will co-operate or consult with each other according to a patient profile, to make the most accurate decision.
12. The final results (decisions) of the KE will be displayed to healthcare personnel by the Knowledge Representation (KR) module. These results will be used to communicate the final results to the physician. According to the level of automation in the CDSS, the KR could:
    1. Display recommendations in the form of images, texts, sounds, and videos.
    2. Require a physician's decision about the final diagnosis and actions. The physician has the choice to refuse, alter or accept the given support. If the physician accepts the support, then CDSS will send an order to any of the following: the pharmacy to prepare the medicine and give it the treatment policy, the laboratory system to prepare for specific tests, and the radiology system to be ready for some tests.
    3. Request additional data to be entered again into the CPOE.

13. The CDSS could make many diagnoses with different probabilities, and the physician can choose the best. Additionally, data mining and machine learning can predict the likelihood of any future problem in the health of the community.
14. The CDSS's final and approved decisions will be stored in the patient local EHR for future reuse. If a patient suffers from a chronic disease, then the CDSS can generate a care plan according to the relevant CPG. A CDSS will store the patient entry point in its EHR. In the next visit, the CDSS will remember the patient previous state. It could use multiple entry points or fuzzy entry points in the CPG if the patient conditions are changed in an unexpected way.

We expect that this framework will provide the most accurate and applicable decision support and will achieve a large amount of integration between HIS and decision support processes.

In addition, the proposed model is fully automated. The physician enters only the patient UID and initial diagnoses, and the CDSS returns decision support. Moreover, the architecture is component-based. Each component of the architecture is pluggable and reusable.

6. Conclusions

A novel knowledge management framework for distributed health care systems has been introduced in this paper. This framework incorporates the knowledge that is extracted by data mining techniques from EHRs with knowledge from domain experts into health care information systems for decision-making support. The model successfully integrates a CDSS into the workflow of the HIS. This framework is fully automated and requires only the patient universal ID and the physician's initial diagnosis to make intelligent decisions. This system has an open architecture in which any number of hospitals, KBs, EHRs and data mining engines can be integrated into an existing environment. This model depends on a set of knowledge bases in different hospitals. Each KB is specialized in a specific domain, and the distributed CDSS architecture facilitates the integration and cooperation of KEs in the cases of patients who have complex medical or diagnostic problems. The KE could send the patient profile or specific data to other sites for consultation. The model also assures that the KB is continuously up to date to allow the CDSS to produce applicable recommendations and actions. To represent the final decisions or results to a physician, the framework has a module for representing the results from the KE in a meaningful way, to allow physicians to make fast and accurate decisions. Future work will seek to implement this framework.

Acknowledgement

This study is part of a research project funded by the National Plan for Science and Technology (NPST) in the Kingdom of Saudi Arabia. Grant No. 09-INF880-02.

References

Ahmed, Mobyen, Begum, Shahina, Funk, Peter, 2012. Case studies on the clinical applications using case-based reasoning. Proceedings of the IEEE Federated Conference on Computer Science and Information Systems, 3–10.

Aleksovska-Stojkovska, L., Loskovska, S., 2010. Clinical decision support systems: medical knowledge acquisition and representation methods. Electro/Information Technology (EIT) IEEE International Conference on page 1.

ASTM International, 2012 [Online]. Available: <http://www.astm.org>.

Bacon, Jean, Moody, Ken, 2002. Toward open secure widely distributed services. Magazine Communications of the ACM 45 (6).

Bodenreider, Olivier, 2004. The Unified Medical Language System (UMLS): integrating biomedical terminology. Nucleic Acids Research 32, D267–D270.

Bohada, John A., Riaño, David, López-Vallverdú, Joan A., 2012. Automatic generation of clinical algorithms within the state-decision-action model. Expert Systems with Applications 39 (12), 10709–10721.

Canadian institute for health informatics (cihi) project, 2012 [Online] Available: <http://www.cihi.ca/>.

R.C., Chakraborty, 2008. Artificial Intelligence Knowledge Representation Issues, Predicate Logics, Rules, <http://myreaders.wordpress.com/>.

Chen, Chiehfeng), Chen, Kung, Hsu, Chien-Yeh, Li, Yu-Chuan, 2011. Developing guideline-based decision support systems using protégé and jess. Journal of Computer Methods and Programs in Biomedicine 102 (3), 288–294.

Data Management Group (DMG). Predictive Model Markup Language (PMML) version 3.0 specification, 2012 [Online]. Available: <http://www.dmg.org/pmml-v3-0.html>.

Nhon, Do., 2008. An ontology for knowledge representation and applications. International Journal of Computer and Information Engineering 3 (3), 180–188.

Erraguntla, M., Gopal, B., Ramachandran S., Mayer, R., 2012. Inference of missing ICD 9 codes using text mining and nearest neighbor techniques. IEEE 45th Hawaii International Conference on System Science (HICSS), pp. 1060–1069.

Gamberger, D., Prcela, M., Jovic, A., Smuc, T., Parati, G., Valentini, M., Kawecka-Jaszcz, K., Styczkiewicz, K., Kononowicz, A., Candelieri, A., Conforti, D., Guido, R., 2008. Medical knowledge representation within Heartfaid platform. In Proc. of Biostec 2008 Int. Joint Conference on Biomedical Engineering Systems and Technologies, 205–217.

Giannopoulou, Eugenia G., 2008. Data Mining in Medical and Biological Research, ISBN 978-953-7619-30-5.

Guideline Interchange Format 3.5 – technical specification, 2012 [Online]. Available: <http://smi-web.stan ford.edu/projects/intermed-web/guidelines/ GLIF_TECH_SPEC_May_4_20>.

Hahn, C., Jacobi, S., Raber, D., 2010. Enhancing the interoperability between multiagent systems and service-oriented architectures through a model-driven approach, web intelligence and intelligent agent technology (WI-IAT). IEEE/WIC/ACM International Conf., vol. 2, p. 415.

IHTSDO: International Health Terminology Standards Development Organization, 2012 [Online]. Available: <http://www.ihtsdo.org>.

International Organization for Standardization, 2012 [Online]. Available: <http://www.iso.org>.

Iqbal, Ashraf Mohammed, Shepherd, Michael, Raza Abidi, Syed Sibte, 2011. An ontology-based electronic medical record for chronic disease management. Proceedings of the 44th Hawaii International Conference on System Sciences, pp. 1–10.

Kazemzadeh, Reza Sherafat, Sartipi, Kamran, 2005. Interoperability of data and knowledge in distributed health care systems. Proceedings of the 13th IEEE International Workshop on Software Technology and Engineering Practice (STEP'05), p. 230.

Kazemzadeh, R.S., Sartipi, K., Jayaratna, P., 2010. A framework for data and mined knowledge interoperability in clinical decision support systems. International Journal of Healthcare Information Systems and Informatics (IJHISI) 5 (1), 37–60.

Kim, Jeong Ah, Cho, InSook, Kim, Yoon, 2008. CDSS (clinical decision support system) architecture in Korea. International Conference on Convergence and Hybrid Information Technology (ICHIT), 700–703.

Kong, Guilan, Xu, Dong-Ling, Yang, Jian-Bo, 2008. Clinical decision support systems: a review on knowledge representation and inference under uncertainties. International Journal of Computational Intelligence Systems 1 (2), 159–167.

Koutkias, Vassilis, Kilintzis, Vassilis, et al, 2012. Knowledge engineering for adverse drug event prevention: on the design and development of a uniform, contextualized and sustainable knowledge-based framework. Journal of Biomedical Informatics 45 (3), 495–506.

Kuo, Kuan-Liang, Fuh, Chiou-Shann, 2011. A rule-based clinical decision model to support interpretation of multiple data in health examinations. Journal of Medical Systems 35 (6), 1359–1373.

Kuperman, Gilad J., Payne, Thomas H., et al, 2007. Medication-related clinical decision support in computerized provider order entry systems a review. Journal of the American Medical Informatics Association 14 (1).

Liao, Shu-Hsien, Chu, Pei-Hui, Hsiao, Pei-Yuan, 2012. Data mining techniques and applications – a decade review from 2000 to 2011. Journal of Expert Systems with Applications 39 (12), 11303–11311.

Maciel, R.S.P., David, J.M.N., 2007. WGWSOA: a service-oriented middleware architecture to support groupware interoperability. IEEE 11th International Conference in Computer Supported Cooperative Work in Design (CSCWD), p. 556.

National Electrical Manufacturers Association (NEMA), 2012 [Online]. Available: <http://www.nema.org>.

Open clinical: ASBRU, Feb 2010. Available: <http:// www.openclinical.org/gmm_asb ru.html>.

Peleg, M., Tu, S., Bury, J., Ciccarese, P., et al, 2003. Comparing models of decision and action for guideline-based decision support: a case-study approach. Journal of the American Medical Informatics Association (JAMIA) 10 (1), 52–68.

Uta Priss, 2010. "Predicate Logic" Set 07106 Mathematics for Software Engineering, School of Computing, Edinburgh Napier University.

Open clinical: PROforma, Feb 2010. Available: <http://www.openclinical.org/gmm_proforma.html>.

Protégé Official Web Site, Feb 2010. Available: <http://protege.stanford.edu/>.

Ramakrishnan, N., Hanauer, D., Keller, B., 2010. Mining electronic health records. IEEE Computer Society 43 (10), 77–81.

Raza Abidia, Syed Sibte, Manickamb, Selvakumar, 2002. Leveraging XML-based electronic medical records to extract experiential clinical knowledge: an automated approach to generate cases for medical case-based reasoning systems. International Journal of Medical Informatics 68 (1–3), 187–203.

Riaño, David, Real, Francis, et al, 2012. An ontology-based personalization of health-care knowledge to support clinical decisions for chronically ill patients. Journal of Biomedical Informatics 45 (3), 429–446.

Sanders, G.D., Nease, R.F., Owens, D.K., 2000. Design and pilot evaluation of a system to develop computer-based site-specific practice guidelines from decision models. Medical Decision Making 20 (2), 145–159.

Sari, Anny Kartika, Rahayu, Wenny, Bhatt, Mehul, 2012. Archetype sub-ontology: improving constraint-based clinical knowledge model in electronic health records. Journal of Knowledge-Based Systems 26, 75–85.

Sartipi, K., Yarmand, M.H., Down, D.G., 2007. Mined-Knowledge and Decision Support Services in Electronic Health, Systems Development in SOA Environments, SDSOA '07, 20–26 May 2007, ICSE Workshops 2007, pp. 10,10.

Sittig, Dean F., Wright, Adam, Simonaitis, Linas, et al, 2010. The state of the art in clinical knowledge management: an inventory of tools and techniques. International Journal of Medical Informatics 79 (1), 44–57.

Tanwar, Poonam, Prasad, T.V., Aswal, Mahendra S., 2010a. Comparative study of three declarative knowledge representation techniques. International Journal on Computer Science and Engineering 02 (07), 2274–2281.

Tanwar, Poonam, Prasad, T.V., Dutta, Kamlesh, 2010b. An effective knowledge base system architecture and issues in representation techniques. International Journal on Computer Science and Engineering 02 (07), 2274–2281.

Tenórioa, Josceli Maria, Hummela, Anderson Diniz, et al, 2011. Artificial intelligence techniques applied to the development of a decision support system for diagnosing celiac disease. International Journal of Medical Informatics 80 (11), 793–802.

Terenziani, P., Montani, S., Bottrighi, A., et al, 2004. The GLARE approach to clinical guidelines: main features. Study of Health Technology Information 101, 162–166.

The American Medical Association, 2012 [Online]. Available: <http://www.ama-assn.org>.

The Arden Syntax for Medical Logic Systems, 2012 [Online]. Available: <http://cslxinfmtcs.csmc.edu/hl7/arden/>.

The European Committee for Standardization, 2012 [Online]. Available: <http://www.cen.eu>.

The Health Level Seven, 2012 [Online]. Available: <http://www.hl7.org>.

The International Health Terminology Standards Development Organization (IHTSDO), 2012 [Online] available: www.ihtsdo.org, last seen December 2012.

The Protégé Ontology Editor and Knowledge Acquisition System, Online available: <http://protege.stanford.edu/>, last seen December 2012.

The Regenstrief Institute, Inc., 2012 [Online]. Available: <http://www.regenstrief.org>.

The World Health Organization, 2012 [Online]. Available: <http://www.who.int>.

US Department of Health & Human Services, The office of the National Coordinator for Health Information Technology, 2012 [Online]. Available: <http://www.hhs.gov/healthit>.

US Department of Health & Human Services. Centers for Medicare & Medicaid Services, 2012 [Online]. Available: <www.cms.gov>.

van Harmelen, Frank, Fensel, Dieter, 1995. Formal methods in knowledge engineering. The Knowledge Engineering Review 10 (4), 345–360.

Wanga, Xiaoyan, Chasea, Herbert, Markatoub, Marianthi, et al, 2010. Selecting information in electronic health records for knowledge acquisition. Journal of Biomedical Informatics 43 (4), 595–601.

World Wide Web Consortium, 2012 [Online]. Available: <http://www.w3.org/XML/>.

Xiao, Liang, Cousins, Gráinne, Hederman, Lucy, Fahey, Tom, Dimitrov, Borislav, 2010. The design of an EHR for clinical decision support. 3rd International Conference on Biomedical Engineering and Informatics (BMEI 2010).

Yina, Wan, 2010. Application of EHR in health care. Second International Conference on MultiMedia and Information Technology 1, 60–63.

Yuwen, Shuli, Yang, Xiaoping, 2009. Standardizing the medical data in China by CDA. In the proceeding of IEEE Fourth International Conference on Computer Sciences and Convergence Information Technology.