The transformation of surgery patient care with a clinical research information system

Nan-Chen Hsieh a,⇑, Kuo-Chen Lee b, Wei Chen c

aDepartment of Information Management, National Taipei University of Nursing and Health Sciences, No. 365, Min-Ten Road, 11257 Taipei, Taiwan, ROC
bDivision of Cardiovascular Surgery, Department of Surgery Heart Center, Cheng-Hsin General Hospital, Taiwan, ROC
cDepartment of Applied Informatics and Multimedia, Asia University, Taiwan, ROC

Abstract

Implementing a computerized clinical research information system (CRIS) can make clinical research easier and more efficient while improving patient care by providing surgeons with performance feedback. To transform the original manual patient information management system so it delivers patient care, this study developed a CRIS for cardiovascular disease to facilitate surgery treatment tracking. The CRIS tracks hundreds of pieces of data through surgical stages and converts these data into computerized registries, provides surveillance mechanisms, and generates clinical interpretive reports in a timely manner. Surgeons can use the CRIS to identify surgical-related data and interventional cardiovascular procedure risks based on specific patient characteristics, and it has increased the quality and efficiency of patient care. An intelligent data analysis (IDA) tool based on the Weka library that seamlessly integrates with the CRIS has helped provide models for clinical research.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Cardiovascular disease is a type of disease that involves problems with the heart or blood vessels. Surgery is usually a primary treatment for serious cardiovascular disease. Surgeons' decisions for cardiovascular patients' care are mainly determined by similar patients' past experiences. Retrospective research studies can extract valuable information from the gathered data, and decisions can be made based on the research results. The interventional methods, procedures, and treatments should be evident for cardiovascular disease, and modern medical practices should be recorded in electronic medical records (EMRs). A computerized CRIS tracks previous patients' outcome data completely and correctly and potentiates surgeries with substantial and accurate data for clinical research. Once the research results for decisions are stored sagaciously, queries can be performed to investigate clinical issues, treatments, procedures, and outcomes.

In this study, the original paper-based special chart for cardiovascular disease was developed based on foreign experience (STS Adult Cardiac Surgery data collection form, http://www.sts.org/sections/stsnationaldatabase/), fused with local population characteristics (Cheng-Hsin cardiac surgery registries Version 2.52.2, http://docs.google.com/View?docid=ddsmjffb_13fbxdbtgx). The data in the special chart are scattered over three surgical stages: preoperative, intraoperative, and postoperative. To facilitate the diagnostic and therapeutic interventions performed by surgeons, nurses and professional staff members, the particular diagnosis or procedure is structured as an interdisciplinary care plan and sequenced on a uniform timeline. All related personnel were collaborators concerning these clinical pathways.

Prior to the development of the CRIS, manual operation was the method for collecting, transcribing, reviewing and judging patient records for the whole surgical stage. Surgeons and medical professionals had to attempt to order multiple data sources on their own and had to wait several days for hard-copy results to be filed in patients' special charts. This manual process was not only time-consuming but also prone to error due to the time-gap required from ordering to judging, as well as the multiple steps involved. There is a great deal of information about patients that can be captured in an electronic format and subsequently used in surgical care processes. The CRIS is intended to support surgeons' care processes and offer several advantages for meeting clinical research requirements. There is also great potential for the advanced healthcare information technology (HIT) features that are integrated with the CRIS to enhance the surgeon's ability to monitor a patient's condition, track the quality of care, and provide long-term patient situation tracking.

Well-managed information is an important and demanded resource in clinical practice, as all care activities, from collecting to judging patient information, are crucial. The recent increase of works on interdisciplinary study involved ontologies, motivated by the interoperability requirements among domains and applications (Abas et al., 2011; Fernandes, Guizzardi, & Guizzardi, 2011;
Kuziemsky & Lau, 2010). Using ontologies carried practical benefits for system developers, as they could reuse and share domain knowledge with a common vocabulary base across heterogeneous platforms and languages, and it prevented developers from being distracted by implementation details. Developing clinical practices and building an information system based on a cardiovascular disease ontology using informatics are worthy goals.

The information held in a CRIS touches virtually every aspect associated with all patient care services. The interactive characteristics of a CRIS allows surgeons to generate and input various procedural and patient examination reports directly into a patient’s registries, thereby moderating the time gap inherent in the paper-based process. CRIS’s procedure-specific module also contains an IDA tool, which helps surgeons make decisions and facilitates detailed reports with reference values. Patient data can be introduced from all intranet workstations of the cardiovascular department, and the medical database structure is established to permit integrating patient-specific information and to support robust statistical analyses on populations of similar patients. This system makes it possible to obtain a comprehensive summary of the cardiovascular population using standard reports and graphics, drill-downs and roll-ups on hierarchical dimensions, and temporal indicator analyses. The system is expected to be integrated with hospital information systems and communicate with other cardiovascular centers (recovery, cardiovascular surgery) to collect complete information.

Fig. 1 shows how a cardiovascular CRIS is helpful before and during cardiovascular surgery. For preoperative risk assessment, surgeons and patients can use the collected information to evaluate whether the surgical procedure would be successful. Similarly, information on surgical mortality and the postoperative morbidity of follow-up surgery care is also important for selecting low-risk patients for an operation and counseling patients about the risks of undergoing surgical operations. A robust CRIS can thus both assist surgeons in evaluating the expected outcome for a given patient and facilitate counseling and operative decision making.

Using the CRIS improves quality of care, helps evaluate treatment effectiveness, and assists in clinical research. We developed a pilot cardiovascular CRIS for Cheng-Hsin General Hospital, which will be extended to other two medical centers’ cardiovascular departments. The objectives for this study are the following:

1. To employ an ontology-based system development approach, which allows interdisciplinary team members to elicit requirements semantically and derive implementation models that meet those requirements
2. To transform the special chart of cardiovascular diseases into electronic registries; the registries’ structures offer possibilities for performing research studies on a population of similar patients.
3. To build a web-based CRIS, whose core concept is based on electronic medical records and registries to support surgery patient care, clinical decision-making and clinical research.
4. To develop an IDA tool and obtain the results of a clinical research study concerning the diagnosis and/or therapy effectiveness for cardiovascular patients.

To better understand this study, Section 2 begins with an overview of the CRIS background. Section 3 describes the regional special chart of cardiovascular diseases and the problems encountered when maintaining it. Section 4 presents an ontology-based system development approach. Section 5 presents the detailed processes for building a CRIS. In Sections 6 and 7, to help relevant personnel who want to conduct clinical research and acquire informative knowledge from a medical database, we describe the IDA tool that was developed based on the Weka library. Section 8 offers observations about practical applications and directions for future research.

2. Clinical research information systems (CRIS)

HIT allows healthcare providers to collect, store, retrieve, and transfer medical data electronically, and it has the potential to improve healthcare quality, safety, and efficiency and prevent medical errors across the healthcare system (Kauert et al., 2001; Kohn et al., 1999). Common HIT implementations provide computerized physician order entry (CPOE), electronic medical records (EMRs), and point of care (POC), allowing access to medical data for entry and retrieval. HIT in surgery care is recognized as lagging behind the acute and ambulatory settings. Thus, the infrastructure that supports the administrative and clinical applications and the CRIS that facilitates patient data input into the care process are two recent and important topics for HIT applications (Haux, 2010; Ohmann & Kuchinke, 2009). A CRIS with EMRs can be used as a clinical passive tool to store patient information and can include IDA functions. With these functions, a CRIS can make substantial contributions to supporting the delivery of patient care, effectively storing data, reporting results, planning of care, analyzing diseases, and visualizing complex correlations among data (Kauert et al., 2001; Shortliffe et al., 2001). CRISs have thus emerged as a particularly important component in healthcare due to the heightened
The awareness of medical errors and increasing numbers of developers involved in CRIS implementation. CRISs play a vital role in quality of care for daily operations.

The emergence and evolution of longitudinal EMRs and clinical CRIS data allow researchers access to large amounts of identified longitudinal clinical data to perform primary research. Using data mining techniques to support medical decisions would increase diagnostic accuracy and provide additional knowledge to the medical staff. The increased use of these techniques provides expanded opportunities for determining the utility of medical decision-making models from longitudinal clinical data. Mining these clinical data provides additional insights through advanced analytics, including visualization analysis, clustering analysis, and classification/predictive modeling. The newly discovered insights can serve as evidence for producing accurate judgments. The predictive and interpretable descriptive models contribute greatly to handling medical data gathered through the systematic use of clinical, laboratory, and existing clinical systems (Bellazzi & Zupan, 2008). In this study, the knowledge extracted with the IDA tool may be in the form of diagrams, charts, or predictive/descriptive models. The goal of predictive data mining is to derive models that can use medical data to predict patient mortality and morbidity and thus support clinical surgical decision making. Descriptive data mining considers data as affinity granulations and aims to find interpretable patterns and associations among data. The evidence might thus be in the form of raw data or original analytical outputs.

Surgeons not only deliver care to patients but also are key personnel who generate patient data. A CRIS is built to aid the practice of evidence-based care while streamlining workflows and monitoring patient information, thus helping ensure the delivery of the consistent and safe care that patients deserve. The offered intelligent decision-support protocols and surveillance alerts ensure patient safety and accelerate time-to-therapy. Relevant patient information is seamlessly and securely transferred throughout various medical professionals, enabling smarter care decisions. The development of a CRIS allows surgeons to create large, integrated databases of patients’ clinical data, which helps surgeons begin the real-time management of populations of similar patients. Mining data in these databases may provide insights into new relationships between disease states and how to effectively manage them. In addition, a CRIS can help surgeons access data for several purposes: (1) to interpret these patient data for population-based clinical decisions, (2) to analyze the patient data for outcome studies, (3) to develop evidence-based decision-aided models, and (4) to monitor the performance of individual surgeons.

3. The regional special chart of cardiovascular diseases

Before the development of a CRIS, detailed patient and procedural information was hand-gathered on multipage forms (Fig. 2) by the cardiovascular fellows and entered into Excel spreadsheets after each care process ended. The disadvantages of this system are that the data must be typed in manually, the care process could not provide accurate information, and significant time and manpower are required to maintain patient-related data, which wastes valuable medical professionals’ time for more relevant medical work. A further disadvantage is the lack of checking the information given by the patient, and because the paper is often not read until the patient has left the hospital, there is little opportunity to correct erroneous data. Furthermore, performing clinical research on paper-based or spreadsheets forms is nearly impossible. The proprietary medical database system and IDA tool for cardiovascular disease must therefore be developed.
Without receiving proper care, high-risk patients may lengthen their hospital stays, increase their cost of care, and decrease their quality of life. A CRIS is primarily used to collect, store, analyze and interpret information from various sources (Wright, Bieniewski, Pifarre, Gunnar, & Scanlon, 1985). A CRIS should support integration with existing EMRs to provide continued monitoring and careful therapeutic decision making and enable the delivery of personalized medical care based on an individual patient’s condition. Healthcare professionals can therefore follow up on patients with cardiovascular diseases. In the developed CRIS, each patient forms an individual registry (Wright et al., 1985), and each registry contains multiple forms, including preoperative, intraoperative, and postoperative datasheets. Specifically, the registry contains data on patients’ preoperative risk factors and postoperative complication findings, from which prediction models are derived by a regular risk assessment tool or computed by the IDA tool. The information is then integrated with additional information describing treatment, follow-up, relapses, metastases, and hospital mortality to create comprehensive registries that help surgeons and nurses improve clinical outcomes.

The early stage alarm feature provides a unique tool to schedule follow-up appointments, lab tests, and radio diagnostic imaging. By incorporating hospital guidelines into the CRIS, all follow-up visits are uniform regardless of the specialist. After all relevant personnel have agreed to follow these guidelines, a centralized department organizes and contacts patients for their follow-up appointments. The CRIS also allows groups of potential patients to be selected for investigatory clinical trials based on their data. Without a CRIS, it would be difficult to anticipate the number of patients who are eligible for a given trial. From the clinical management perspective, cardiovascular departments could be managed according to the specific hospital load, without having to rely on published population-based data. The number of new cases diagnosed every year in a hospital could be obtained from the hospital-based CRIS.

As stated, building a CRIS is an interdisciplinary task. An ontology provides the opportunity to coordinate various professionals and specifies a conceptualization that domain users could separately identify and use as a self-contained way to communicate domain information. Ontologies have become effective tools for handling compound and varied information sources, which helps model data in an interdisciplinary information system. Conversely, an ontology has a similar effect as conceptual data modeling. Ontologies are intended to give details and explain objects, while conceptual data modeling is used to represent objects and their relationships for a specific application (Siricharoen, 2009). An ontology thus behaves as a richer information model than traditional conceptual models, including the UML class diagram. In this study, we regard ontology modeling as an extension of conceptual modeling, which could be a promising approach for CRIS development.

4. The ontology-based system development approach

The setting for this study was the department of cardiovascular surgery at Cheng-Hsin General Hospital, a flagship medical center in Taiwan that offers a complete range of cardiovascular surgical care. We developed a CRIS to collect prospective data on the clinical course of patients who have had cardiovascular surgery. The typical development paradigm of a CRIS is to follow surgical stages. This paradigm is intuitive to surgeries and similar to entering data into conventional paper-based forms. These forms can be filled in incrementally by different users at different locations and timepoints during the distributed surgical care process. Hundreds of state variables are attributed to the forms, and by selective, state-dependent views on forms (e.g. a list of generated reports and list of validated reports), intra- and interdepartmental document flow is supported. The developed CRIS supports not only building paper-like electronic special charts but also an IDA tool for clinical decision assistance.

For system development, the conceptual model contains the core activities of the information system design and implementation. The conceptual model represents real-world concepts and relationships among them and can be described using various notations, including UML for object modeling and IDEFIX for entity relationship modeling. The traditional conceptual modeling methodologies, however, focus primarily on representing the technological processes and neglect the communication needs within interdisciplinary teams. Most conceptual modeling methodologies provide visual communication tools, but they cannot provide an easy way for surgeons to view major surgical processes within their familiar contexts. Furthermore, traditional conceptual modeling methodologies do not allow technologists to separate the complexities of the surgical processes into a set of manageable technological solutions; their analysis methods mainly incorporate technological requirements from a broad perspective and therefore overlook the specificities of the surgical processes and requirements.

We proposed an ontology-based system development approach that allows members of the interdisciplinary teams to elicit requirements semantically and derive the implementation models that meet those requirements. The rationale of this approach is that ontologies have a broader scope than conceptual modeling in system development. Ontologies focus on the universe of discourse and are used to derive conceptual models, while conceptual modeling focuses on the components of the information system that are being modeled (Fonseca & Martin, 2007; Soaresa & Fonseca, 2007). Fig. 3 shows how a three-phase ontology-based system development approach guides the requirement analysis activities to elicit requirements and helps develop conceptual models for system modeling. Phase 1 employed requirement analysis methodology to derive user requirements and cardiovascular domain knowledge. Phase 2 used ontology methodology to derive cardiovascular domain ontologies and UML methodology to use the derived ontologies to develop conceptual models. Phase 3 implemented the CRIS, which was guided by the derived ontologies. Using this approach, we can transfer surgical stages into the CRIS design, build a sharable knowledge database of the processes and requirements among users, establish communication among interdisciplinary team members, and ensure a user-centered CRIS design.

4.1. Conceptualization Phase

Through a literature review, interviews and a field investigation, professionals familiar with medical and information technology collected medical domain knowledge. The requirement collection was accomplished by a cross-department committee, which includes the surgeon-in-charge of the cardiology division and professional system developers. This committee met to discuss implementing a CRIS for patients who require a cardiovascular surgical operation. A glossary of cardiovascular relevant terms is first collected. Analyzing special charts and scaling down the amount of collected data ensured complete data acquisition on all patients. When manpower limitations occurred, data acquisition and analysis became recurrent problems. The overall goal of the CRIS was to “make it simple” while maintaining high-quality clinical data. The data collection instrument was simplified to a special chart, which included information regarding patient data and procedural results. The committee also decided to hold physicians responsible for completing the data collection form immediately after performing the procedure.
4.2. Formalization Phase

The special charts were first re-modeled as computerized registries (Wright et al., 1985). Each patient has their individual registry or file folder, and each registry contains many forms, including demographic, surgical, clinic and laboratory datasheets. Moreover, the registry contains data on a patient’s preoperative characteristics, risk factors, surgical procedure details, physical characteristics of the heart and postoperative physiological and laboratory findings. This step aimed to develop clear and concise forms, which must conform to the surgical workflow and be acceptable to formal medical records. This process allows a user to realize how forms can be filled incrementally at different surgical stages and verify whether the system workflow fits their daily work practice. This stage next utilized the derived cardiovascular surgery ontology as the basis for implementing the database schema and functional components of the CRIS. Implementing the ontology as an aid facilitates the transition from requirement analysis to conceptual modeling and then to implementation. Integrating procedural ontologies and user views of the system introduced the development of data surveillance, forms, reports and user interfaces that allow surgeons to interact with the CRIS.

4.3. Implementation Phase

The implementation stage contains two main tasks. First, the database development task involves converting domain and ontologies into the physical database model and ensures the appropriateness of the built model. The patient’s special chart becomes relational database tables guided by ontology’s concepts. Rapid prototyping was used to develop the CRIS. A Microsoft SQL server was used to develop the prototype database with ASP.net to code the derived functionality from stage 2. Because other hospitals wanted to participate in this project and the registry content will be readjusted, the registry structure was left open. At the initial stage of the functional development task, roughly 80% of the special charts were completed as computerized registries. Second, a prototype system with provisionary forms is first prepared for users to test system functionality with test data. Key users test each surgical flow in the CRIS to ensure that no surgical flow problems have been found until the application module reaches a stable state. After testing each surgical flow, each form was completed by adding detailed fields, predefined values, and surveillance criteria. Trained personnel were required to enter patients’ data retrospectively from surgeons’ special charts. The supported IDA functions were integrated with Weka API. The supported IDA functions were integrated with Weka API.

5. Ontology perspectives to the building of the CRIS

The patient care process of cardiovascular surgery roughly contains three stages: the preoperative, perioperative, and postoperative stages. A CRIS is helpful before and after the surgery. In the preoperative stage, patients are assessed by an interdisciplinary team of cardiologists and anesthetists. Patients’ demographic data
and characteristic disease histories, as delivered by the referring cardiologist, are used for preoperative risk assessment to identify high-risk patients. Surgeons and patients can use the collected data to evaluate whether the surgical procedures would be successful and to predict operative and postoperative death and complications. Similarly, information on perioperative mortality and postoperative morbidity is also important for selecting low-risk patients for an operation and counseling patients about the risks of undergoing surgical operations. A robust CRIS can therefore not only help surgeons evaluate the expected outcomes for a given patient but also facilitate counseling and operative decision making. This CRIS allows surgeons to assess the patient’s condition, utilizing the clinical evolution and changing risk factors for each patient. The structure of the CRIS database offers the possibility to conduct research studies on the analysis of cardiovascular pathology, assessment of risk factors, and building of prognosis models to evaluate pre-operation mortality and post-operation morbidity. The provided IDA tool can be used as a preliminary tool for primary cardiovascular disease research.

A major characteristic for the patient’s prognosis is that the prognosis result is not static but may change considerably during the care processes, as factors related to the surgical intervention may have important implications on the prognosis. The prognosis of the perioperative death risk during the operation or postoperative hospital stay and the estimation of the duration of the intervention and postoperative stay at the ICU are the most relevant factors in the cardiac surgery care process. These outcomes can be seen as a substitute for the complication extent during the intervention and recovery processes and thus as a measure of the quality of the delivered care. After the intervention, patients are sent to the intensive care unit (ICU) to closely monitor their postoperative physiological condition. In a normal (i.e., uncomplicated) recovery process, a stable condition is reached within 24 hours, after which the recovery process is completed at the nursing ward. Several complications may occur during this postoperative stay at the ICU, including arrhythmia, neurological complications, and infections. These complications delay the recovery process and may lead to death.

The predicted preoperative risk for cardiovascular surgical patients must therefore be reassessed during the care process based on data from the treatment course to provide surgeons with the latest relative data for the patients. Surgeons must process numerous medical records and organize these medical records into a “special chart” that they can use to analyze and evaluate medical outcomes and for research purposes. The cardiovascular special chart is a confidential document that contains detailed and comprehensive information on an individual and the care experience related to the patient. In this study, the cardiovascular special chart is primarily based on the “STS Adult Cardiac Surgery data collection form,” offered by the Society of Thoracic Surgeons. To meet the specific-population characteristics of cardiovascular disease in Taiwan, we have made regional modifications. A special chart serves as a medical record of an individual’s clinical status, care, history, and caregiver involvement.

The specific information contained in the chart is intended to provide a record of a patient’s clinical condition by detailing diagnoses, treatments, tests, and responses to treatment, as well as any other factors that may affect the patient’s health or clinical state. The special chart, for research purposes about the effects of a CRIS on patients, has two distinct dimensions: patient-related information encoded by the relative systems and care-related processes or practices that characterize how patient-related information is or could be used. Accordingly, the constructed CRIS contains five categories of encoded patient-related entities: (1) a representation of the patient’s demographic information and medical history, (2) a representation of the patient’s current health status and cardiovascular risk factors, (3) a representation of the surgeon’s care plans, (4) a representation of the patient’s diagnosis or intervention information, and (5) a representation of the patient’s related information from other medical systems.

An ontology provides many benefits when developing an intelligent information system. The ontology’s vocabulary and taxonomy abilities provide a conceptual framework for sharing, analyzing, and retrieving data in a specific domain. An ontology thus lets system developers focus on application domain structure rather than implementation details and allows the system developers to reuse and share application domain knowledge across different software platforms and programming languages. In medical informatics, an ontology can help clinical professionals and patients utilize appropriate pieces of knowledge when facing complex medical problems and situations (Gasparic et al., 2009). We developed an ontology for modeling cardiovascular surgery concepts.

The core task of ontology development is the categorization process. Good categorization can facilitate information retrieval from database system. The development of a cardiovascular surgery ontology should include relevant concepts, attributes, constraints and instances. Fig. 4 illustrates a schematic ontology class hierarchy structure for the cardiovascular surgery ontology. The cardiovascular surgery ontology presents a formalized description of cardiovascular surgery concepts. The information presented in this ontology was extracted from cardiovascular special charts and disease guidelines, and it has been discussed with surgeons. The ontology class hierarchy structure is constructed based on three surgical operation stages: preoperative, intraoperative and postoperative, and they represent the three major sub-categories under the root of cardiovascular surgery ontology. There are three major categories under preoperative sub-ontology: demographic, health condition and medication. This sub-ontology includes all concepts and risk factors that are relevant for patients in the preoperative stage. The intraoperative sub-ontology includes important surgery information and auxiliary equipment information for cardiovascular operations. The postoperative sub-ontology includes information related to health conditions after surgery discharge for follow-up.

Using an ontology system development approach, we can intuitively design a CRIS. Fig. 5(a) shows how the sub-ontology of current patient risk factors was used to develop a conceptual database schema and design the system screen. Each concept from the current patient ontology, including demographics, health conditions, and risk supporting details, has been implemented, as shown in the database schema in Fig. 5(b) and Fig. 5(c) shows a sample patient’s registry maintenance screen. A 73-year-old female patient with atrial fibrillation was preparing for surgery. She had hypertension, peripheral vascular disease, and chronic renal insufficiency. The preoperative examination revealed a height of 149 cm and weight of 46.4 kg. The cardiac sonography and pertinent lab data included the following: LVEDD, 49 mm; IVS, 9 mm; RVSP, 90 mm; creatinine, 1.9 mg/dL; FEV1, 1.03 L/sec; and FEV1/FVC ratio, 78%. Fig. 6 shows the major medical forms and several charts from the three surgical stages.

6. The functionality of intelligent data analysis for clinical research

Patients with cardiovascular problems are usually elderly and may be considered at high risk for surgical operation. These patients are likely to have a relatively high postoperative morbidity rate with complications and will highly influence longer-term postoperative outcomes. The statistical significance parameters before a surgical operation can be estimated using information from
previous patients. Moreover, identifying the reasons for postoperative complications and risk factors for re-intervention during follow-up to maintain cardiovascular problem exclusion remain challenges for surgeons. It is essential to create reliable preoperative and postoperative risk evaluation models to help surgeons keep track of patients.

Though several risk prediction systems have been proposed for patients undergoing cardiovascular operations (Bohm et al., 2008), they basically rely on traditional statistical methods and provide scant accuracy and utility when applied to risk evaluation. The research on predictive clinical outcome models and the amount and duration of treatment is indispensable to improve the quality of care. A prediction model is a risk assessment function often used for clinical decision making before and during surgery. Classifying patients for surgical mortality and postoperative morbidity is also important when selecting low-risk patients for operation and counseling patients about the risks of undergoing a surgical operation. Prediction models can also help classify patients into affinity groups for clinical trials (Abu-Hanna & Lucas, 2001).

Data mining in healthcare is a relatively new research field. The major objective is to acquire knowledge from medical databases that can predict or classify medical data with satisfactory results. Practitioners are expected to employ data mining in medical databases to support diagnosis, prognosis, and follow-up tasks. Data mining tools (Berthold et al., 2009; Mierswa et al., 2006; Witten, Frank, & Hall, 2011) are usually run from the desktop as stand-alone application software. This approach for the use of data analysis functions is inconvenient and not easily accessible for those users without training and the installed software. Because web services provide easy-to-deliver and accessible information resources on the network, in this study, we developed an IDA tool as a web application, and any client in the network can install and access the IDA tool. The corresponding algorithms in the Weka library enabled data analysis tasks by interaction with the web service and only returned the analyzed results, thus not requiring the downloading of the entire dataset. An IDA tool is therefore seamlessly integrated with medical databases.

The IDA tool accesses the Weka library using a web service call and presently contains two function models. The Data Integration function model is integrated with the medical database in the CRIS. Users can directly connect the required fields in the medical database and do not need to import data from other format files. The

Fig. 4. The cardiovascular surgery ontology class hierarchy in supporting the CRIS.
Weka library also supports the basic data edition and descriptive statistics functions. The Model Development function model aids the development of desired models from the selected dataset through an interactive interface. The user can choose various models, including BN (Bayesian network), ANN (artificial neural network), SVM (support vector machine) and regression, to resolve medical predictive and descriptive problems as an individual model (Bellazzi & Zupan, 2008). In contrast to ordinary models that evaluate one hypothesis from the training dataset, the user can also choose an ensemble model that works to construct a set of hypotheses and combine the results, from which multiple models are trained to solve the same problem (Polikar, 2006). The designed IDA tool allows different types of users to access the functionalities in Weka without the need to learn new data mining software.

The IDA functions of the CRIS reflect three surgical stages and perform analysis functions to extract useful information from the medical datasets. The gathered medical datasets not only retrieve information from the CRIS database, but they also obtain

![Fig. 5. A cardiovascular surgery sub-ontology used to develop a patient risk factors maintenance screen.](image)

Weka library also supports the basic data edition and descriptive statistics functions. The Model Development function model aids the development of desired models from the selected dataset through an interactive interface. The user can choose various models, including BN (Bayesian network), ANN (artificial neural network), SVM (support vector machine) and regression, to resolve medical predictive and descriptive problems as an individual model (Bellazzi & Zupan, 2008). In contrast to ordinary models that evaluate one hypothesis from the training dataset, the user can also choose an ensemble model that works to construct a set of hypotheses and combine the results, from which multiple models are trained to solve the same problem (Polikar, 2006). The designed IDA tool allows different types of users to access the functionalities in Weka without the need to learn new data mining software.

The IDA functions of the CRIS reflect three surgical stages and perform analysis functions to extract useful information from the medical datasets. The gathered medical datasets not only retrieve information from the CRIS database, but they also obtain
information from other healthcare systems. A robust prediction model in medical datasets can not only help surgeons evaluate the expected outcome for a given patient but also facilitate counseling and preoperative decision making. The IDA functions in the CRIS aim to help represent the patient's characteristics as satisfactorily as possible for clinical uses. A hypothesis is established that the medical datasets can be modeled and preprocessed by statistical and/or machine learning algorithms for their ability to analyze non-linear medical data. The function of risk prediction by a priori knowledge is to apply statistical and/or machine learning methods to obtain a forecast of whether a given operation outcome will occur before and after surgery. We have retrospectively examined data from hundreds of consecutive patients who underwent a cardiovascular surgical operation. The pilot study showed that the developed CRIS could provide useful results, but it needs to gather as much evidence as possible from the subject patients to assess its utility and reliability.

7. Developing intelligent data analysis functions based on open-source tools

Machine learning techniques have gradually received attention in medical informatics. Several well-known machine learning tools have already been announced, including Weka (Witten et al., 2011), Yale/RapidMiner (Mierswa et al., 2006), and Java-ML (Abeel, Peer, & Saeyes, 2009). These open-source data mining suites provide a graphic user interface for data analysis and model visualization, but this may not be a major concern for a system developer. The IDA tool should integrate with the CRIS to support uninterrupted medical data analysis tasks, including interactive exploring interesting patterns and deriving useful models (Bellazzi & Zupan, 2008; Zupan & Demsar, 2008; Zupan, Holmes, & Bellazzi, 2006). In this study, we intended to develop custom-designed medical data analysis components and schemata that access the data mining suite’s components through Java programming and embed those components into the host CRIS. The following models were embedded into the host CRIS: (1) basic statistical functions for primary data inspection; (2) exploratory data visualization functions to interact with interesting patterns in datasets; (3) data source connection and preprocessing functions to retrieve medical data from databases, sample selection, feature ranking, and feature discretization; (4) unsupervised and supervised data analysis functions for intelligent data analysis, including statistical and machine learning algorithms, (e.g. regression analysis, clustering techniques, association rules, decision trees, support vector machines, Bayesian classifiers, and discriminant analysis); and (5) model evaluation functions to evaluate the built models’ reliability and validity (e.g. accuracy, sensitivity, specificity, ROC curves, and lift chart).

Weka (http://www.cs.waikato.ac.nz/ml/weka/) is a well-known open-source data mining tool, which includes many impressive machine learning and data mining algorithms. Weka provides a menu-based interface, called Weka Explorer, for general users, or the Weka Knowledge Flow interface to accomplish intelligent data analysis tasks. Weka readily provides algorithms and an easily extensible API for researchers. These tools facilitate a broad explo-
ration of different machine learning algorithms and allow for their integration into self-development systems. Programmers can access Weka’s core algorithms and components through Java programming. Weka’s Wiki-collected extensions (http://weka.wikispaces.com/Related±Projects) cover diverse areas, including time series mining, visualization tools, parallel processing, text
mining, and bioinformatics. Many researchers contributed their research results to this community, and Weka became a primary research tool for researchers in data mining and machine learning. Consequently, medical practitioners could obtain IDA functions through Weka.

The final IDA prototype tool was developed as a module within CRIS. Fig. 7 shows how a three-tier architecture was based on dot-Net framework with SQL server and machine learning functionalities accomplished in a Weka open source library, which was designed to leverage efficient access to clinical data and support for clinical decision making. Web services are used to achieve interoperability between the CRIS and IDA modules because web-based systems are platform- and programming language-independent, standard protocols are accessible, and they facilitate communication. IDA tool adheres to a flexible design because the open source license allowed us to freely access the library and make it available for other uses without constraints. We combined CRIS and asynchronous web service calls, making accessing immediate data analysis services from the WEKA library possible (Shetty, S.Vadivel, Weka based desktop data mining as web service, & Engineering, 2010). Fig. 8 shows the web-based IDA tool that centralizes the support prognosis task. Users can select the desired fields for building classifiers and check the performance of the predictive model results.

8. Conclusion

In this study, we developed a pilot CRIS to aid in the care of cardiovascular patients. The CRIS was integrated as a medical database system to monitor patients’ status changes following the preoperative, intraoperative, and postoperative surgical stages and facilitate communication among interdisciplinary team members in the time between preoperative assessment and follow-up tracking. We described the architecture and functional modules of the CRIS and reported the result of a partial deployment study whose goals were to demonstrate the ontology-based system development approach and technological feasibility of the CRIS.

The technological feasibility to improve patient healthcare management has shown promising results and may facilitate improved efficiency in delivering medical care. We showed that the CRIS aims to help surgeons better track patients who undergo cardiovascular surgery, such as through the use of the IDA tool that can predict and classify the progression of patients. The CRIS can determine patient status and risk probability before a surgical operation. Some illness trends, however, could be identified, which is helpful for clinical decision making. Validating the derived prognosis models remains a challenge, and more patient data must be gathered from the clinical study population.

By addressing only the utilities of the initial deployment, we cannot fully measure the system’s potential impact on the quality of care. Significant reductions of paper-based recording errors, detailed recordings of patients’ clinical information, and effective patient monitoring improvements support a positive effect of CRIS intervention. Given that the cost of full implementation is negligible and the potential benefits in increased patients care’s quality are large, further work will deploy the system in other hospitals, which will make it possible to obtain a comprehensive summary of the cardiovascular population database in Taiwan by defining uniform variables. Our study shows that the CRIS design allows for the analysis of the stored data using standard descriptive statistics and data mining techniques. These data analysis methods can indicate the reliability of stored data. We anticipate that inter-hospital systems will regularly apply this system and significantly improve patient care quality in the future.

References


