AN INTELLIGENT RISK DETECTION FRAMEWORK USING KNOWLEDGE DISCOVERY TO IMPROVE DECISION EFFICIENCY IN HEALTHCARE CONTEXTS: THE CASE OF PAEDIATRIC CONGENITAL HEART DISEASE

Fatemeh Hoda Moghimi, School of Business IT & Logistics, RMIT University, Australia, fatemeh.moghimi@rmit.edu.au

Dr. Hossein Seif Zadeh, School of Business IT & Logistics, RMIT University, Australia, hossein.zadeh@rmit.edu.au

Professor Nilmini Wickramasinghe, School of Business IT & Logistics, RMIT University, Australia, nilmini.wickramasinghe@rmit.edu.au

Abstract

Healthcare professionals, especially surgeons must make complex decisions with far reaching consequences and associated risks. As has been shown in other industries, the ability to drill down into pertinent data to explore knowledge behind the data greatly facilitates superior, informed decisions to ensue. This proposal proffers an Intelligent Risk Detection (IRD) Model using data mining techniques followed by Knowledge Discovery in order to detect the dominant risk factors across a complex surgical decision making process and thereby to predict the surgery results and hence support superior decision making. To illustrate the benefits of this model, the case of the Congenital Heart Disease (CHD) is presented.

Key words: Knowledge Discovery, Data Mining, Risk detection, Decision making, Congenital Heart Disease (CHD).

1 The materials of this current proposal is adapted from the author’s confirmed PhD research proposal in RMIT university and also an under review paper that it’s sent to PACIS 2011.
1 SUMMARY OF PROPOSED RESEARCH STUDY

In this research, an intelligent risk detection framework using a knowledge discovery model to improve decision efficiency in healthcare contexts is proposed. Unfortunately, for many patients, surgery is not the curative answer, as the surgery might impose its own side effects or by products such as some disabilities, cancers, diabetes and even sudden death. Given, that quality of life is very important for patients and their family; decision making in this type of surgeries, is a complex issue for surgeons, patients and their families. To design a solution regarding this issue, the proposed research will focus on detecting some surgery risk factors to predict the surgery results based on these risk factors to improve surgical decision efficiency. Moreover, from the intelligence continuum model (Wickramasinghe and Schaffer, 2006), data mining followed by knowledge discovery techniques is defined as a sound approach to detect risk factors, discover the relationships between risk factors and also predict the surgery results based on these risks factors.

The aim of this research is to reduce the burden of the surgeries on the patients, their families and society by developing an intelligent risk detection framework to improve the surgery decision making process. Additionally, a risk assessment process using KPIs (key performance indicators) and an expert group’s viewpoints, beside the surgery decision making framework will be developed across this research. Based on an extensive literature review it appears that this is the first study that directly examines the benefits of integrating risk detection and prediction, to improve a decision making process in healthcare contexts. Mixture of qualitative and quantitative methods will be used to conduct the study. Key aspects of the research design include use of the interviews and questionnaire for primary and secondary data collecting in order to create a list of risk factors and evaluate them to provide KPIs to test the model and for future assessments by the proposed intelligent model. Finally, the results from the model will be judged by an expert group. This research study is designed based on a single case study approach. To illustrate the benefits of the incorporation of intelligent risk detection to improve decision efficacy in healthcare contexts, the case of Congenital Heart Disease (CHD) has been chosen, an area that is not only of significance but also involves multiple risks and critical decision making processes and hence an appropriate environment to demonstrate the benefits of the research approach.

2 BACKGROUND

This study contends that an Intelligent Risk Detection (IRD) model in support of better treatment decision making during and after surgery can provide superior healthcare outcomes for the patients and their families. In developing such a solution, it is necessary to combine three key areas of knowledge discovery, risk detection with decision support systems (figure 1). This is an important contribution of this research to both theory and practice in healthcare context.

It is important to note that the proposed IRD model draws primarily from three significant bodies of literature, RD, KD and DSS and its foundations are derived from the IC model.

Effective decision making is vital in all healthcare activities. While this decision making is typically complex and unstructured, it requires the decision maker to gather multi-spectral data and information in order to make an effective choice when faced with numerous options (Wickramasinghe et al., 2009). Unstructured decision making in dynamic and complex environments is challenging and in almost every situation the decision maker is undoubtedly faced with information inferiority. Recognizing this, Wickramasinghe and Schaffer developed the Intelligent Continuum (IC) model (Wickramasinghe and
Schaffer, 2006), a systematic approach that enables the application of knowledge management (KM) principles and tools necessary for improving the decision making processes in healthcare and to ensure that the healthcare decision making process outcomes are optimized for maximal patient benefit. The following research in progress attempts to extend this idea.

3 RESEARCH GOALS & QUESTIONS

Three primary research goals are presented below; in addition it’s noted that reducing the burden of surgery on the patients, their families and society is a strategic benefit of this research.

- To provide superior decision support in the healthcare contexts.
- To discover and extracting hidden knowledge (patterns and relationships) associated with surgeries and other treatments from historical data to detect the surgery risk factors.
- To predict the surgery results to identify patients at risk doing surgery.

This research will answer the following key question:
- How can an intelligent risk detection application for healthcare contexts be developed?

To answer this question, a conceptual model will be developed and then this model will be tested using extensive literature synthesis, contextual interviews with an expert group and a single case study methodology. In doing so it will also be necessary to answer the following sub question:

- Is this model valid, reliable and applicable in practice?

This research will use a deductive single case study method to further test and define the model, to ensure its validity and reliability, and to gain a deep understanding of the underlying logic in the model. The applicability to practice of this final model will be justified through analytical generalization. In this study, the proposed data site will be a hospital in Melbourne.

4 LITERATURE REVIEW

This section outlines the major issues pertaining to the key areas of decision support systems, risk detection and knowledge discovery as well as CHD context (as a research case).

4.1 Congenital Heart Disease/Defects (CHD)

Congenital Heart Disease (CHD) is a common health problem affecting many children around the world. CHD context has been chosen as a case study in this research because it’s involved the multiple risks and critical surgical decision making processes. The term “congenital heart disease” refers to “disorders of heart or central blood vessels present at birth” (Larrazabal et al., 2007a). “CHD is one of the biggest killers of infants less than one year old” and the risk of death remains significantly higher than normal for these patients throughout their life (Mavroudis and Jacobs, 2002). Unfortunately surgery is not always considered a final cure, as it can result in a considerably high rate of disabilities, as well as the possibility of co-morbidities for example, types of cancer and even the development of bowel disease (Amitay et al., 2006). Further, there is the direct adverse impact on the patients and their families. And finally, CHD also carries significant societal and economic implications (Gayet, 2002). A multi-faceted set of considerations including the immediate medical result, the ongoing increased risk of sudden death, exercise intolerance, neurodevelopmental and psychological problems as well as long-term impacts on the family unit as a whole is necessary when evaluating CHD surgery (Mavroudis and Jacobs, 2002). This multi-faceted consideration is important because of the far reaching consequences that can result post surgery:

- The risk of sudden death for patients surviving operation for common types of CHD is 25 to 100 times greater than an age-matched control population (Silka et al., 1998).
- More than 50% of CHD patients after surgery demonstrate abnormalities in neurodevelopmental testing (Gayet and Jacobs, 2005).
- Patient's parents have attachment difficulties with their CHD-affected infants compared to those of healthy infants (Goldberg and Rock, 1992). Parents of children with CHD are also found to be overprotective, overindulgent and inconsistent in disciplining their children (Harrison et al., 1995).
- Families of children with CHD experience more financial strain and greater familial/social stress compared to control groups (Casey et al., 1996).
As can be seen, the decisions relating to treatment strategies for children suffering from CHD are both complex and high risk.

4.2 Decision Support Systems (DSS)

Healthcare organizations are recognizing the need to incorporate the power of a decision efficiency approach. This approach in medical diagnosis and clinical practice is set to increase 10-fold within the next decades (Miller 1994). Fundamentally this research covers clinical and medical aspects typically focusing on how information technology emulates and improves decision-making effectiveness for individual physicians (Fieschi et al. 2003). In particular, computer based decision support systems focus on any software designed to directly aid in clinical decision making in which characteristics of individual patients are matched to a computerized knowledge base for the purpose of generating patient-specific assessments or recommendations that are then presented to clinicians for consideration (Hunt et al. 1998). In addition, the computer-based patient record, the Internet, shared decision-making processes and current regulations also facilitate medical decision support systems (Fieschi et al. 2003).

Decision-making regarding surgery for patients with some disease is multi-faceted and complex, including for example the clinical condition of the patients, their age and weight (Karlson et al. 2003). The decision to treatments with either drugs, or surgery, or a combination of both depends on a large number of factors (Roy and Brunton 2008). The decision making process in the context of the complex surgeries can be divided into three broad phases. In the first phase, or pre-operation phase, the surgeon, having received much information about the patient and their medical conditions, needs to make a decision relating to whether surgery is the best medical option. Once this decision is made but before surgery, the patients must then decide whether to accept or reject the surgeon’s decision in consideration of the predicted outcomes. However, patients will often have met many medical staff before they meet the cardiac surgeon. Thus already at stage one, two key decisions must be made. Once patients and surgeons have agreed to proceed, in phase two, ad-hoc decisions pertaining to the unique situations that may arise during the surgery must be addressed. Finally, in the post operating phase, or phase three, decision making is primarily done at two levels; a) strategies to ensure a sustained successful result for the patient during aftercare and beyond, and b) record of lessons learnt for use by clinicians in future similar cases. To capture this complexity, in the conceptual model, two steps of decision making are defined in three different and key phases of the decision making process for complex surgeries. The first type of decision making is called “surgical decision making” as it is primarily associated with the surgeons. The second type is called “parental/personal decision making” because some surgery outcomes (such as “quality of life”) directly affects or at least must be determined by the patients or their parents and therefore they have a critical say regarding whether (or not) to proceed with the surgery.

4.3 An Intelligent Risk Detection Framework for Healthcare Contexts

From the literature review, it has been noted that surgery-driven validated risk detection outcome analysis can indeed lead to improvements in performance by both individual surgeons and surgery centers (Mavroudis and Jacobs 2002). Surgical performance is usually indirectly measured by postoperative outcome of the initial hospital stay by means of risk-adjusted audits (Gayet 2002). Although risk adjustment is important to assess performance and compare outcomes amongst individuals or institutions (Kang et al. 2004b), statistical inferences alone cannot be used to determine what is considered acceptable performance (Gayet and Jacobs 2005). Risk adjuster systems have been under development since the 1980s and have been implemented by Medicare Choice program, numerous states, employer coalitions, and health plans (Keenan et al. 2001). Such systems have been based on many factors, including diagnosis, prior utilization, demographics, persistent diseases, and self-assessments of health and/or functional status. An analytical review on these systems shows some limitations, most important of which are listed below:

- Lack of a dynamic risk assessment system.
- Lack of a multidimensional risk detection model/algorithmm.
- Focusing on cost management and financial risks rather than clinical risk factors.

Therefore, regarding these limitation through current systems, with an in depth review of risk detection in healthcare area, it’s found that applying some IT based techniques such as knowledge discovery followed by data mining would increase the performance of current risk adjustment methods significantly.
4.4 Knowledge Discovery & Data Mining

The focus of Knowledge Discovery (KD) is usually data mining and how to achieve data mining tasks (Cios et al. 2007). Data mining is a computerized technology that uses complicated algorithms to find relationships and trends in large databases, real or perceived, previously unknown to the retailer, to promote decision support. Based on Intelligence Continuum (IC) model (Wickramasinghe and Schaffer 2006), using data mining technique followed by KD is an approach to extracting patterns from large data sets and deducting knowledge insights from those patterns (Desouza 2002). The IC Model is including but not limited to data mining, business intelligence/business analysis (BI/BA) and knowledge management (KM) (Wickramasinghe and Schaffer 2006). In order to maximize the value/utility of the IRD model and because the combined techniques of DM, BA/BI and KM are essential in the present context, the IC model is used as the foundation for the IRD model as shown in figure 1.

5 CONCEPTUAL MODEL

In order to realise the proposed IRD model, it is necessary to develop a conceptual model of the decision making stages and risk assessment (figure 3). The left-most block in figure 2 depicts the first stage of risk assessment. The output of the risk assessment process will help in determining important surgery risk factors and also predicting anticipated outcomes (in risk detection block) based on the specific risk factors. The anticipated results enable the surgeons to then make better informed decisions regarding whether (or not) to proceed with the surgery in first phase as well as second and third surgical phase.

If the decision is indeed to proceed with the surgery, all relevant information then needs to be passed on to the patients/parents, in pre-operative phase, in order to allow them to make their final decision regarding the surgery. Depending on their decision one move to second phase or this concludes the process. Any such conflicts are feedbacks into the system for future risk assessments for the same or other similar patients.

5.1 Risk Assessment

Detecting the risk factors based on a risk assessment process using knowledge discovery techniques is a useful way to assess improvements in surgery (Larrazabal et al., 2007a). Therefore, after first identifying important risk factors in the literature, we will seek expert input at two distinct stages to address this subject. The specific stages involved in the risk assessment process are shown in figure 3. In the first stage, the specialists through an expert group of surgeons are presented with risk factors identified from the literature. The experts will then nominate (or introduce) some main risk categories or dimensions as well as risk factors to be used in the surgical decision making process. In the next
stage, in order to designing a scoring mechanism, the expert group will be asked to evaluate the risk factors and also define the relationships between these factors or between these factors and some actual or anticipated surgery results.

<table>
<thead>
<tr>
<th>Finding Risk Categories</th>
<th>Finding Risk Factors</th>
<th>Scoring mechanism</th>
<th>Defining relevant KPIs</th>
<th>Risk Auditing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 3. The risk assessment**

In the Scoring Mechanism, the surgeons can achieve consistent strategy execution and monitor performance by tracking patient situation in surgery procedures, and enhancing their relevant data, the Balanced Scorecard can provide a complete view of the risk factors, their level and their impact on surgical results via a risk detection process (in section 5.2). Then, a value range of the risk factors will be extracted by the expert group, in order to define the relevant key performance indicators (KPIs). Finally, risk auditing, the last step of the risk assessment process, will serve to keep the model up to date by using results of the report made by comparing actual and anticipated results (in section 5.3).

### 5.2 Risk Detection Using Knowledge Discovery

To incorporate an intelligent technology into the proposed risk assessment process, this research suggests a data mining process followed by knowledge discovery. In the research case, the data types have a significant impact on the data mining tasks. Hence, after finishing the data collecting phase, the suitable tasks will extract such as neural networks and association rules. To apply the necessary data mining techniques, developing and then implementing the model, after the risk assessment process, this research will design a small database that included the patients’ data and also some data to show risk factors. Then it moves through step 1 to step 6 (below). The steps are:

**Step 1.** Understand all clinical requirements, dataset structure and data mining tasks and designing a dimensional data mart.

**Step 2.** Prepare target datasets: select and transform relevant features; data cleaning; data integration. Communicate any findings during data preparation to domain experts.

**Step 3.** Train multiple data mining models in randomly sampled partitions using Clementine\(^2\) or Rapid miner\(^3\).

**Step 4.** Evaluate data mining models using a set of performance metrics.

**Step 5.** Discuss the data mining results with domain experts. Explore potential patterns from data mining results. If identify new risk factors or patterns, communicate the rule(s) with decision makers and determine the appropriate actions.

**Step 6.** Go back to Step 1 if some new questions are raised during the process or new KPIs or risk factors are discovered. Otherwise, finish and exit the process.

### 5.3 Applying Anticipated & Actual Results

In the proposed conceptual model, to evaluate a risk detection process, the actual results will be compared with the anticipated results. This is because sometimes actual results present some new risk factors or new measurements to assess the risk factors. This comparison would be the best solution to create a final report to show some important items, and finally apply them to the risk assessment process, for next iterations of evaluations.

### 6 RESEARCH DESIGN & METHODOLOGY

This section describes the methodological design with which this research will be accomplished. This proposed research will be conducted using a mixed method approach as presented in table 1 while the experimental research design is illustrated in figure 4.

\(^2\) A commercial software for data mining

\(^3\) An open source software for data mining
Study design ▪ A literature search and environmental scan is conducted to inform the design of the study and guide the development of the qualitative and quantitative research tools.

Qualitative phase ▪ Key informant interviews will conduct with two selectively sampled groups of participants 1) an expert group (pediatric cardiologists and cardiac surgeons). 2) patients’ parents in the research case.

Quantitative phase ▪ Online quantitative questionnaire, informed by the exploratory phase, will distribute to the expert group.

Analysis & Validation ▪ Analysis, interpretation and conclusions based on multidisciplinary team collaboration and input.

Table 1. Mixed method approach in this research

<table>
<thead>
<tr>
<th>Phase 1: Study design</th>
<th>Phase 2: Qualitative data collection</th>
<th>Phase 3: Quantitative data collection</th>
<th>Phase 4: Data analysis &amp; validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Literature review</td>
<td>Semi structural interviews to collect data regarding all steps in risk assessment</td>
<td>Findings Risk Categories</td>
<td>Data Validation</td>
</tr>
<tr>
<td>Research Design</td>
<td></td>
<td>Findings Risk Factors</td>
<td>N</td>
</tr>
<tr>
<td>Finalizing the Proposal</td>
<td></td>
<td>Scoring mechanism</td>
<td>Y</td>
</tr>
<tr>
<td>Ethic approval</td>
<td></td>
<td>Defining relevant KPIs</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Risk Auditing</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Validation</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4. Experimental research design

7 RESEARCH OUTCOMES & LIMITATIONS

In this current research, the lack of interaction between healthcare industry practitioners and academic researchers makes it hard to discover the surgery risk factors, and limits opportunities for the application of data mining techniques, and hence weakens the value that knowledge discovery and data mining methods may bring to healthcare risk detection. Risk detection across the surgery for any disease, in itself, is challenging due to the great diversity of the patient population in terms of the diagnoses, indications for operation (Kang et al. 2004a). Based on extensive literature review it appears that this is the first study that directly examines the benefits of real time risk detection and outcome prediction in order to augment decision making process in healthcare contexts. Also, using KPIs (key performance indicators) as a set of metrics is a novel idea to control the risk factors, finding its level and defining their relationships together, also with the other factors in such healthcare context. Furthers, KPIs will be so effective to monitor some key items during surgery for surgeons. Additionally, it should reduce the burden of the surgery on its patients, their family and society is the other strategic benefits that will examined. Another benefit of the proposed IRD model is the incorporation of continual improvement. An important feature of the IRD model is the integration of the three IT solutions to solve
a clinical issue in the definition and assessment of “outcomes”, combined by some assessment measures. One of the steps that include the theoretical framework developed, needs to be tasted in the research. However, empirical testing of the framework is likely to face a number of challenges such as identifying common metrics for measuring the risk factors (Garg et al., 2005). Moreover, regarding the case study, this research will be faced with some issues. For example, The CHD surgery risk context has many dimensions and detecting the risk factors in all of these dimensions is not easy but with contribution of the relevant expert group, this research will try to cover some main dimensions.

8 PROPOSED RESEARCH SCHEDULE

This current research started in March 2010. As shown in table 2, the research proposal is confirmed on 4th March 2011 and now, ethics approval is under process to take from both a hospital in Melbourne and RMIT University.

<table>
<thead>
<tr>
<th>Activity of milestone</th>
<th>2010 - every cell is 2 months</th>
<th>2011 - every cell is 2 months</th>
<th>2012 - every cell is 2 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Literature review</td>
<td>March 2010</td>
<td>Dec 2010</td>
<td>Feb 2011</td>
</tr>
<tr>
<td>Finalizing the research proposal</td>
<td>Dec 2010</td>
<td>Feb 2011</td>
<td>Dec 2011</td>
</tr>
<tr>
<td>Ethic approval</td>
<td>Feb 2011</td>
<td>Dec 2011</td>
<td>Feb 2012</td>
</tr>
<tr>
<td>Qualitative data collection</td>
<td>Dec 2011</td>
<td>Feb 2012</td>
<td>Dec 2012</td>
</tr>
<tr>
<td>Analysis and validation of the qualitative data</td>
<td>Feb 2012</td>
<td>Dec 2012</td>
<td>Feb 2013</td>
</tr>
<tr>
<td>Test the model</td>
<td>Feb 2013</td>
<td>Dec 2013</td>
<td>Feb 2014</td>
</tr>
<tr>
<td>Quantitative data collection</td>
<td>Dec 2013</td>
<td>Feb 2014</td>
<td>Dec 2015</td>
</tr>
<tr>
<td>Analysis and validation of the quantitative data</td>
<td>Feb 2014</td>
<td>Dec 2015</td>
<td>Feb 2016</td>
</tr>
<tr>
<td>Complete the first draft</td>
<td>Dec 2015</td>
<td>Feb 2016</td>
<td>Dec 2017</td>
</tr>
<tr>
<td>Review the final draft and submit the final draft</td>
<td>Feb 2016</td>
<td>Dec 2017</td>
<td>Feb 2018</td>
</tr>
</tbody>
</table>

Table 2. Proposed Research Schedule

9 CONCLUSIONS

In this proposal, an intelligent risk detection model using knowledge discovery methods is proposed. Intelligent risk detection is a particularly challenging area for the healthcare industry while relatively common for fraud detection in finance, diagnosis in industry, and affect analysis in chemistry. In this proposed research, the application of knowledge discovery to high-level surgery risk detection and outcome prediction is presented. The model designed is based on two steps of decision making process (surgical and personal/parental) and, includes a decision support system which is suitable for high concentration prediction. This study confirms that the selection of the risk detection, prediction by knowledge discovery and then decision making are also very important for the surgical decision making process. This research has a mixed method approach to collect the data in two data collection phases. To present the benefits of the model and test it, the case of Congenital Heart Disease (CHD) has been chosen. In closing, this study contends that real time intelligent risk detection appears to be critical for many areas in healthcare where complex, high risk decision must be made.

* In addition the conceptual model was presented at HICSS, Jan 2011.
References


Silka, G., Hardy, D., MENASHE & MORRIS D. (1998) A population-based prospective evaluation of risk of sudden cardiac death after operation for common congenital heart defects Pediatric Cardiology and Medical Informatics and Clinical Outcomes. USA Published by the American College of Cardiology Foundation.


