European Journal of Operational Research 219 (2012) 564–573

Contents lists available at SciVerse ScienceDirect

European Journal of Operational Research

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

A belief rule-based decision support system for clinical risk assessment of cardiac chest pain

Guilan Kong [a,*], Dong-Ling Xu [b], Richard Body [c], Jian-Bo Yang [b], Kevin Mackway-Jones [d], Simon Carley [d]

[a] Medical Informatics Center, Peking University, No. 38 Xueyuan Road, Beijing 100191, China
[b] Manchester Business School, University of Manchester, Manchester M15 6PB, UK
[c] Cardiovascular Sciences Research Group, Core Technology Facility, University of Manchester, Manchester M13 9WL, UK
[d] Emergency Department, Manchester Royal Infirmary, Oxford Road, Manchester M13 9WL, UK

ARTICLE INFO

Article history:
Available online 10 November 2011

Keywords:
Decision support systems
OR in medicine
Uncertainty modeling
Belief rule base
Evidential reasoning approach
Clinical risk assessment

ABSTRACT

This paper describes a prototype clinical decision support system (CDSS) for risk stratification of patients with cardiac chest pain. A newly developed belief rule-based inference methodology-RIMER was employed for developing the prototype. Based on the belief rule-based inference methodology, the prototype CDSS can deal with uncertainties in both clinical domain knowledge and clinical data. Moreover, the prototype can automatically update its knowledge base via a belief rule base (BRB) learning module which can adjust BRB through accumulated historical clinical cases. The domain specific knowledge used to construct the knowledge base of the prototype was learned from real patient data. We simulated a set of 1000 patients in cardiac chest pain to validate the prototype. The belief rule-based prototype CDSS has been found to perform extremely well. Firstly, the system can provide more reliable and informative diagnosis recommendations than manual diagnosis using traditional rules when there are clinical uncertainties. Secondly, the diagnostic performance of the system can be significantly improved after training the BRB through accumulated clinical cases.

1. Introduction

Human health has improved significantly in the last 60 years, and it results in increased life expectancy in most countries. Lengthened life span causes today's governments and health care providers to face a big challenge of delivering a most effective and efficient health care to citizens (Brandeau et al., 2005). However, the quality of current health care in most countries is not satisfactory. In the UK, patient safety incidents or adverse events, which are unintended or unexpected incidents that could have or did lead to harm for one or more patients receiving National Health Service (NHS)-funded health care, represent a serious public health problem and pose a threat to patient safety (Thomas and Brennan, 2001). Research shows that around 10% of patients admitted to NHS hospitals have experienced patient safety incidents, and that up to half of these incidents could have been prevented (Department of Health, 2004). Patient safety incidents cause great harm to not only patients and their families, but also involved clinicians and host hospitals. For example, it is estimated that patient safety incidents cost NHS £2 billion a year in addition to hospital stays, without taking account of human or wider economic costs (Department of Health, 2004).

An important reason for high rate of patient safety incidents is medical errors that are mostly caused by human factors (Reason, 2001). Appropriate increase in the use of information technology (IT) in health care, particularly the introduction of clinical decision support system (CDSS), can help simplify the health care process and substantially facilitate clinical practice and reduce medical errors (Bates et al., 2001; Menachemi et al., 2007; Sim et al., 2001; Kawamoto et al., 2005). There are numerous examples of CDSSs in health care such as (De Dombal et al., 1972; Jonsbu et al., 1993; Lin et al., 2006) which have successfully improved the quality of health care.

Though CDSSs are promising in helping facilitate evidence-based medicine and reducing patient safety incidents, there are challenges in this research area that have made few CDSSs widely applied in practice. Among others, uncertainties in clinical signs, clinical symptoms and clinical domain knowledge, the complexity of involved inference mechanism, and difficulties with domain selection and knowledge base construction and maintenance may impede development and implementation of CDSSs (Miller and Geissbuhler, 1999). Therefore, representation of and reasoning with uncertain medical knowledge are critical areas that require refined methodologies and techniques (Musen et al., 2006; Lin et al., 2006). In the literature, different methods have been proposed to model

* Corresponding author.
E-mail addresses: guilan.kong@gmail.com (G. Kong), ling.xu@mbs.ac.uk (D.-L. Xu), richard.body@manchester.ac.uk (R. Body), jian-bo.yang@mbs.ac.uk (J.-B. Yang), kevin.c.mackway-jones@manchester.ac.uk (K. Mackway-Jones), s.carley@btinternet.com (S. Carley).

0377-2217/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2011.10.044

uncertainties in CDSSs using expert system based approaches. Certainty factors were used in MYCIN (Shortliffe, 1976), which was an early rule-based expert system developed in the early 1970s at Stanford University to identify causes of severe infections and to recommend antibiotics. The name MYCIN was derived from antibiotics, as many antibiotics have the suffix "-mycin". MYCIN uses backward chaining to search its rule base after acquiring a series of inputs from users, and it will abandon further searching on one rule if calculated certainty factor of the rule is 0.2 or less. Bayesian theory was employed in Iliad (Warner, 1989), which was a CDSS developed at the University of Utah to teach medical decision-making and to provide assistance in differential diagnosis across the domain of internal medicine. When clinical findings are entered into Iliad, a sequential Bayesian inference algorithm generates a ranked list of diagnoses and assigns a posterior probability to each diagnosis. Fuzzy logic was used in CDSSs such as an expert system for renal transplantation assignment developed by Yuan et al. (2002), where three steps are needed in the fuzzy inference process: fuzzification, fuzzy rule inference, and defuzzification. Fuzzification is for interpreting a crisp numerical input as a fuzzy set with the membership function, and defuzzification is for defuzzifying fuzzy conclusions to crisp values. However, these methods have their deficiencies in representing uncertainties or reasoning with uncertainties. For example, certainty factor model and Bayesian inference model require complete knowledge about all parameters in one specific decision making process which need significant estimation efforts by domain experts, and thus often create a bottleneck in the already tedious and time-consuming knowledge acquisition process. Reasoning in most fuzzy logic based experts systems is controversial, as there is usually a fuzzification and de-fuzzification procedure in the inference process.

To support modeling and reasoning with clinical domain knowledge under uncertainties, we propose to employ a recently developed belief rule-base inference methodology using the evidential reasoning approach (RIMER) (Yang et al., 2006) for developing an intelligent CDSS (Kong et al., 2009). In RIMER, belief rule base (BRB) is employed to model clinical domain knowledge under uncertainties. Such a rule base is capable of capturing vagueness, incompleteness, and nonlinear causal relationships, while traditional 'IF-THEN' rules can be represented as a special case. Inference with BRB is implemented using the evidential reasoning approach (ER) (Yang and Xu, 2002; Wang et al., 2006), which can combine triggered belief rules with uncertain information. Compared with uncertainty handling methods in existent CDSSs as discussed above, firstly, the parameters of BRB in a belief rule-based CDSS can be initially acquired from domain experts or assigned randomly if domain experts are unavailable, as the clinical belief rules in the system can be fine-tuned by accumulated patient data, and thus can help circumvent the bottleneck of estimating all parameters in initial knowledge acquisition process; secondly, the initial BRB system can contain different types of uncertainties such as incompleteness, and the ER based inference mechanism can help rationally preserve these uncertainties and represent their effects in the final conclusion. The clinical area that we chose as a target area for developing the CDSS is cardiac chest pain (CCP), and the specific computerized decision support is for clinical risk assessment of CCP in emergency department (ED).

In this paper, an overview of CCP and specific CDSSs in the area is discussed in Section 2. A brief introduction to RIMER is provided in Section 3. A belief rule-based CDSS prototype is presented in Section 4, where the web-based system architecture is discussed in Section 4.1, and the details of system components are discussed in Section 4.2. The validation of the developed CDSS prototype is discussed in Section 5, where a set of 1000 simulated patients with CCP is used for the validation and is discussed in Section 5.1, and the validation of the inference engine and the BRB training module are discussed in Sections 5.2 and 5.3 respectively. Conclusions about the research are presented in Section 6.

2. Overview of CCP and specific CDSSs in the area

In the study, we chose suspected CCP as a target clinical area for providing clinical decision support, as ischaemic heart disease remains the leading cause of death in the western world (Allender et al., 2008). Further, suspected CCP may account for up to 6% of all patient attendances at ED and for 27.8% of all acute medical admissions to hospital (Goodacre et al., 2005). On suspecting CCP, a clinician will admit most patients to hospital for investigation to determine whether they have an acute coronary syndrome (ACS), which actually represents a spectrum of disease ranging from acute myocardial infarction (AMI) with irreversible damage to the heart muscle to unstable angina, where there is no damage although the patient remains at risk of AMI occurring in the near future (Body, 2008).

When a patient presents to ED with chest pain, a clinician will enquire about the patient's symptoms and previous medical history and will examine the patient. Using this clinical information alone, it is usually impossible for a clinician to determine whether the symptoms are indeed caused by ACS (Body et al., 2008; Body et al., 2010a), meaning that the majority of patients are admitted to hospital for investigation. However, only a minority of the patients admitted will ultimately be diagnosed with ACS. Conversely, up to 6% of patients who are discharged after presenting to ED with chest pain actually have AMI that was unrecognized, which has significant prognostic implications for the patient (Collinson et al., 2000). Clearly, there is a need to improve diagnostic techniques in order to reduce the number of unnecessary hospital admissions and to reduce the number of patients who are misdiagnosed and inappropriately discharged.

In response to these problems, researchers have been trying to develop specific CDSSs to aid the risk stratification of patients with suspected CCP since the late 1970s. We provide a review of those specific CDSSs targeting automatic clinical risk assessment of CCP from an uncertainty handling perspective as follows.

Hudson and Cohen (1987, 1988, 2002) developed a modified rule-based CDSS -EMERGE to aid the diagnosis of chest pain. In EMERGE, uncertainty in medical rules and medical data is represented using antecedent weights and degrees of presence of clinical signs and symptoms, and approximate reasoning techniques are incorporated in the system to do inference with the uncertainties. However, since the overall reasoning method used in the system is rule chaining, and the system uses threshold values to decide whether a conjoined antecedent of one rule should be substantiated or not, the system cannot take all possible clinical situations into consideration in final diagnosis, and as a result it cannot provide comprehensive diagnosis recommendations. Assanelli et al. (1993) developed a CDSS which also uses rule-based reasoning for patient risk profile estimation. In the system, traditional rules are used for medical domain knowledge representation, and a mixed forward and backward rule chaining method is used for reasoning. The disadvantage of the system is that uncertainties are not considered at all. Aase (1999, 1993) developed a CDSS for risk assessment of acute coronary heart disease based on Bayes' theorem. Although Bayes' theorem is based on a well-established field of mathematics and it has the ability to represent and reason with uncertain knowledge, a complete set of precise prior probabilities as required by Bayes' reasoning are hard to get as discussed in Section 1.

Since uncertainties in clinical decision making are inevitable (Szolovits, 1995), to help clinicians in ED to improve performance in stratifying patients of CCP with uncertain clinical information,

we propose to use RIMER for development of a CDSS which can meet the needs. In the system, belief rules are used for representing uncertain knowledge about clinical risk assessment of CCP, and inference with the clinical rule base and observed uncertain clinical data for a specific patient are implemented using the ER approach. A brief introduction to RIMER is provided in the next section.

3. Brief introduction to RIMER

In RIMER, BRB is used to model domain specific knowledge under uncertainty, and the ER approach is employed for inference with BRB. In this section, we provide a brief introduction to BRB first, then we outline the process of inference with BRB using the ER approach, and finally we discuss how to train BRB using accumulated historical data.

3.1. Modeling domain knowledge using BRB

BRB is extended from traditional rule base by adding a belief structure, in which knowledge representation parameters including rule weights, antecedent attribute weights and belief degrees in consequents are embedded.

Conventionally, in a rule base, the kth rule in an 'IF-THEN' format can be described as

$$R_k: \text{ If } A_1^k \wedge A_2^k \wedge \cdots \wedge A_{T_k}^k, \quad \text{then } D_k \tag{1}$$

where $A_i^k (i = 1,\ldots,T_k)$ is a referential value of the ith antecedent attribute in the kth rule, T_k is the number of the antecedent attributes used in the kth rule, and D_k is the consequent of the kth rule.

If rule weights, antecedent attribute weights, and belief degrees associated with all possible consequents are taken into account, rule (1) can be extended to a packet rule using a belief structure, which is referred to as a belief rule and can be described as

$$R_k: \begin{pmatrix} \text{If } A_1^k \wedge A_2^k \wedge \cdots \wedge A_{T_k}^k, \text{ Then } \{(D_1,\beta_{1k}),(D_2,\beta_{2k}),\ldots,(D_N,\beta_{Nk})\} \\ \left(\beta_{jk} \geqslant 0, \sum_{j=1}^{N} \beta_{jk} \leqslant 1\right), \text{ with a rule weight } \theta_k \text{ and attribute weights} \\ \delta_{k1},\delta_{k2},\ldots,\delta_{kT_k} \quad k \in \{1,\ldots,L\} \end{pmatrix} \tag{2}$$

where $\beta_{jk} (j = 1,\ldots,N; k = 1,\ldots,L)$ is the belief degree originally given by experts or randomly assigned to D_j which is believed to be the consequent if the input satisfies the packet antecedents $A^k = \left(A_1^k, A_2^k, \ldots, A_{T_k}^k\right)$ in the kth belief rule, the attribute weight δ_{ki} $(k = 1,\ldots,L; i = 1,\ldots,T_k)$ represents the relative importance of the ith antecedent attribute in the kth rule, and the rule weight θ_k represents the relative importance of the kth rule in the rule base. L is the number of all belief rules in the rule base. T_k is the number of all antecedent attributes used in the kth belief rule. N is the number of all possible consequents in the rule base.

BRB is a collection of belief rules as described by (2). Inference with BRB is implemented using the ER approach, and knowledge representation parameters including rule weights θ_k, antecedent attribute weights δ_{ki} and consequent belief degrees β_{jk} can be learned from past experience or data.

3.2. Inference with BRB using the ER approach

The ER approach was originally proposed to deal with multiple attribute decision analysis (MADA) problems having both qualitative and quantitative attributes under uncertainty (Yang and Singh, 1994). The Kernel of the ER approach is an ER algorithm which was developed for aggregating multiple attributes based on a belief decision matrix, decision theory and the evidence combination rule of the Dempster–Shafer (D–S) theory (Shafer, 1976).

Given an input $U = (U_i, i = 1,\ldots,T)$ together with its corresponding belief degree $\varepsilon = (\varepsilon_i,\ i = 1,\ldots,T)$, where T is the total number of antecedent attributes in the rule base, U_i $(i = 1,\ldots,T)$ is the input value of the ith antecedent attribute, and ε_i $(i = 1,\ldots,T)$ represents the degree of belief assigned to the input value U_i of the ith antecedent attribute, which reflects the uncertainty of the input data. We use the rule or utility-based equivalence transformation techniques (Yang, 2001) to transform the input data to a distribution on referential values of each antecedent attribute using belief degrees as

$$S(U_i,\varepsilon_i) = \left\{(A_{ij},\alpha_{ij}); j = 1,\ldots,J_i\right\}, \quad i = 1,\ldots,T \tag{3}$$

where A_{ij} is the jth referential value of the ith antecedent attribute, α_{ij} the degree to which the input U_i with belief degree ε_i belongs to the referential value A_{ij} with $\alpha_{ij} \geqslant 0$ and $\sum_{j=1}^{J_i} \alpha_{ij} \leqslant 1 (i = 1,2,\ldots,T)$, and J_i is the number of all referential values of the ith antecedent attribute.

The analytical format of the ER algorithm (Wang et al., 2006) used for inference with the BRB and the input data under uncertainties can be described as follows.

$$\beta_j = \frac{\mu \times \left[\prod_{k=1}^{L}\left(\omega_k\beta_{jk} + 1 - \omega_k \sum_{j=1}^{N}\beta_{jk}\right) - \prod_{k=1}^{L}\left(1 - \omega_k \sum_{j=1}^{N}\beta_{jk}\right)\right]}{1 - \mu \times \left[\prod_{k=1}^{L}(1-\omega_k)\right]}, \quad j = 1,\ldots,N \tag{4}$$

with

$$\mu = \left[\sum_{j=1}^{N}\prod_{k=1}^{L}\left(\omega_k\beta_{jk} + 1 - \omega_k\sum_{j=1}^{N}\beta_{jk}\right) - (N-1)\times\prod_{k=1}^{L}\left(1-\omega_k\sum_{j=1}^{N}\beta_{jk}\right)\right]^{-1}$$

where $\beta_j (j = 1,\ldots,N)$ is the final belief degree attached to the jth consequent D_j after combining all activated rules in BRB, $\beta_{jk} (j = 1,\ldots,N; k = 1,\ldots,L)$ is the original belief degree assigned to D_j in the kth belief rule, and ω_k is the kth rule's activation weight which can be calculated by

$$\omega_k = \frac{\theta_k\alpha_k}{\sum_{j=1}^{L}\theta_j\alpha_j} = \frac{\theta_k\prod_{i=1}^{T_k}(\alpha_i^k)^{\bar{\delta}_{ki}}}{\sum_{j=1}^{L}\left[\theta_j\prod_{l=1}^{T_k}(\alpha_l^j)^{\delta_{jl}}\right]} \quad (k = 1,\ldots,L) \tag{5}$$

$\bar{\delta}_{ki} = \frac{\delta_{ki}}{\max_{i=1,\ldots,T_k}\{\delta_{ki}\}}$ $(0 \leqslant \bar{\delta}_{ki} \leqslant 1)$ is transformed from antecedent attribute weight $\delta_{ki} (k = 1,\ldots,L; i = 1,\ldots,T_k)$. $\alpha_i^k (i = 1,\ldots,T_k)$ is the individual matching degree to which the input U_i $(i = 1,\ldots,T_k)$ belongs to $A_i^k (i = 1,\ldots,T_k; k = 1,\ldots,L)$ that is the referential value of the ith antecedent attribute used in the kth rule, and it is generated from the input transformation as described by (3), with $\alpha_i^k \geqslant 0$ and $\sum_{i=1}^{T_k}\alpha_i^k \leqslant 1$. $\alpha_k = \prod_{i=1}^{T_k}(\alpha_i^k)^{\bar{\delta}_{ki}}$ $(k = 1,\ldots,L)$ is called the combined matching degree to which the input vector U matches the packet antecedent A^k in the kth rule. T_k is the total number of antecedents in the kth belief rule.

The final combined result generated by ER is represented by $\{(D_1,\beta_1),((D_2,\beta_2),\ldots,(D_j,\beta_j),\ldots,(D_N,\beta_N))\}$, where $\beta_j (j \in \{1,\ldots,N\})$ is the final belief degree attached to the jth consequent D_j after combining all activated rules in the BRB.

3.3. BRB Training

The initial belief rules and knowledge representation parameters including rule weights, attribute weights and consequent belief degrees in a BRB can be given by domain experts or randomly generated, and they may not be 100% accurate. An initial BRB can be trained using historical data to improve its ability for representing clinical domain knowledge (Yang et al., 2007).

The aim of BRB training is to find a set of parameters $(\theta_k, \delta_i, \beta_{jk})$ of a BRB that can help it accurately represent domain specific knowledge. The training process is implemented through minimizing the discrepancy between BRB results and sampled data. Assuming there are M cases in a training sample, and the input–output pairs of the M cases are $(\hat{x}_m, \hat{y}_m)$ $(m = 1, \ldots, M)$, the process of learning from these M datasets can be depicted as in Fig. 1, where y_m is produced by the BRB system, the real output $(\hat{y}_m)$ is observed by experts or acquired by instruments, and $\xi(P)$ represents the difference between the real output and the system output. In the BRB optimization model, the objective function is to minimize $\xi(P)$, and the constraints define what the knowledge representation parameters of a BRB system should follow. As a result of the training process, there will be a new set of $(\theta_k, \delta_i, \beta_{jk})$ for BRB.

4. A belief rule-based CDSS prototype

4.1. System architecture

Taking advantages of web technologies, a web-based decision support system (DSS) can deliver suggestions or recommendations generated from the system to a much broader audience who is geographically separated (Bhargava et al., 2007). As to web-based DSSs in clinical areas, a web-based CDSS has the following two advantages. Firstly, it can provide easy access to computerized decision support for clinicians in geographically different places. Secondly, it can provide easy dissemination of clinical domain knowledge and patient data among different clinical application systems which are linked through Internet or Intranet. Motivated by the web technologies, we adopt the web-based three-layer architecture (Sommerville, 2007) in the prototype design of the CDSS.

In the three-layer architecture of a web-based intelligent CDSS, in general there should be at least four system components, namely friendly web-based user interfaces, inference engine, knowledge base, and database. To meet the needs of evidence-based medicine, we propose to integrate automatic knowledge learning functionality into the belief rule-based CDSS. As a result, core components implemented in the belief rule-based system prototype include web-based user interfaces, database, inference engine, knowledge base, and BRB training module. Fig. 2 shows the actual implementation of the above mentioned core components in the prototype and their interrelationships.

It is important to note that in the prototype CDSS as depicted in Fig. 2, BRB model is uniquely stored and manipulated in back-end relational database. This makes physical knowledge base construction flexible and portable and helps the knowledge sharing between different clinical systems free of technology barriers due to mature database technologies.

We used Visual Studio 2003. NET and MATLAB on platform Windows XP Professional to develop the prototype. The data-base management system (DBMS) used for design and development of back-end relational database is SQL Server 2000. Development of core system components including knowledge base, inference engine, and BRB training module is presented in the following section.

4.2. System components

4.2.1. Knowledge base constructed using BRB

The domain specific knowledge we used to construct BRB in our prototype for assessing clinical risk of CCP is from one of our co-authors, Dr. Richard Body at Manchester Royal Infirmary (MRI). A decision tree for clinical risk assessment of CCP as shown in Fig. 3 is an outcome from his recent research (Body, 2009), and it provides initial rules for our BRB construction.

The decision tree was derived from data of 796 patients who presented to the ED of MRI with suspected CCP between 16th January 2006 and 3rd February 2007, and the method used to generate the tree is a form of binary recursive partitioning known as classification and regression tree (CART). From the decision tree, we can see that factors which determine clinical risk of CCP include 'ECG showing STEMI or not', 'EVaMACS score', 'having angina or not', 'having diabetes or not', 'smoking or not', 'gender' and 'age', and there are four different risk levels: 'very high risk', 'high risk', 'low risk', and 'no risk' that can be used to describe one patient's risk. Here ECG stands for 'electrocardiography', STEMI stands for 'ST segment elevation myocardial infarction', and EVaMACS represents 'Early vascular markers of acute coronary syndromes' and its score depends on ECG status, H-FABP, and TnI.

Regarding how to determine which risk factor should be taken into account in the decision tree growing process, plenty of statistical analyses need to be done first to test which clinical factor or variable is both reliable and making significant contribution to clinical risk, and then only those factors which can meet the criteria are entered into the CART analysis. As to the four terminal nodes in the decision tree, each terminal node is assigned with a different risk level according to its position in the decision tree and the proportion of patients in each terminal node who had positive outcome. These four risk levels can facilitate triage of patients to an appropriate level of in-patient care. More specifically, patients with 'no risk' can be safely discharged from the ED without further investigation, patients with 'low risk' can be appropriately investigated in a low dependency environment such as an ED observation ward or clinical decision unit (CDU), patients with 'high risk' should be sent for investigation in an acute medical ward or medical admissions unit (MAU) and patients with 'very high risk' should be triaged to a high dependency environment such as the Coronary Care Unit (CCU).

Based on the above decision tree, we can derive several traditional 'IF-THEN' rules for assessing clinical risk of CCP. For example, if we use 'ECG' and 'EVaMACS score' as two risk factors for predicting clinical risk, we can get the following rules.

R^1: IF ECG shows STEMI, THEN 'very high risk';
R^2: IF ECG shows no STEMI and EVaMACS score is >2, THEN 'very high risk';
R^3: IF ECG shows no STEMI and EVaMACS score is {1,2}, THEN 'high risk';
R^4: IF ECG shows no STEMI and EVaMACS score is 0, THEN $\cdots$(Check the decision tree for further information).

The traditional rules as shown above contain no uncertain information, and based on these rules, if antecedent values can be

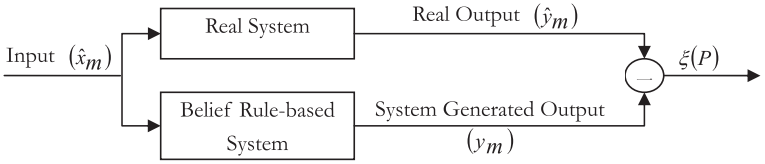


Fig. 1. Training process (Yang et al., 2007).

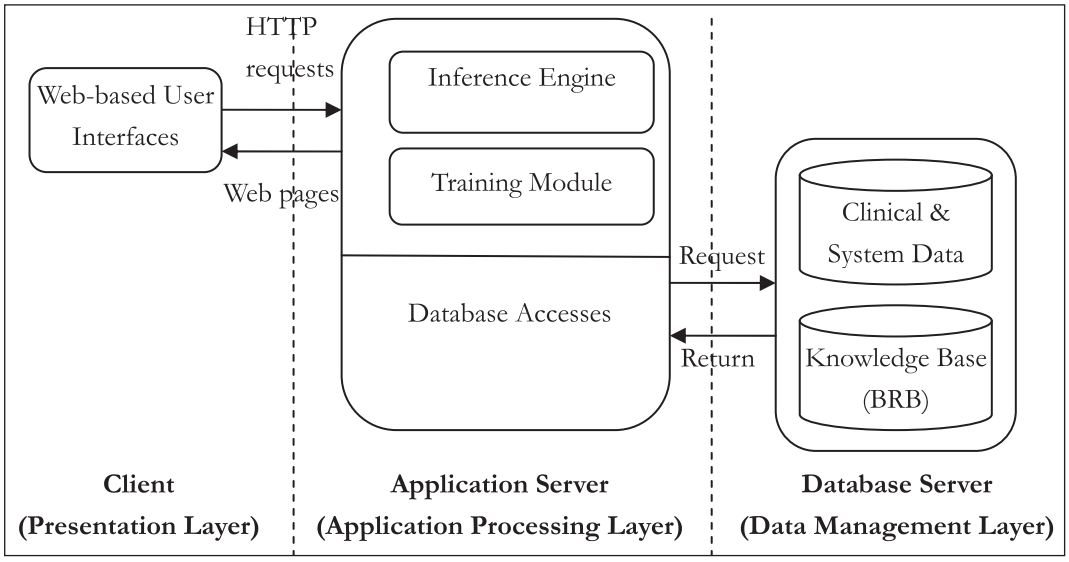
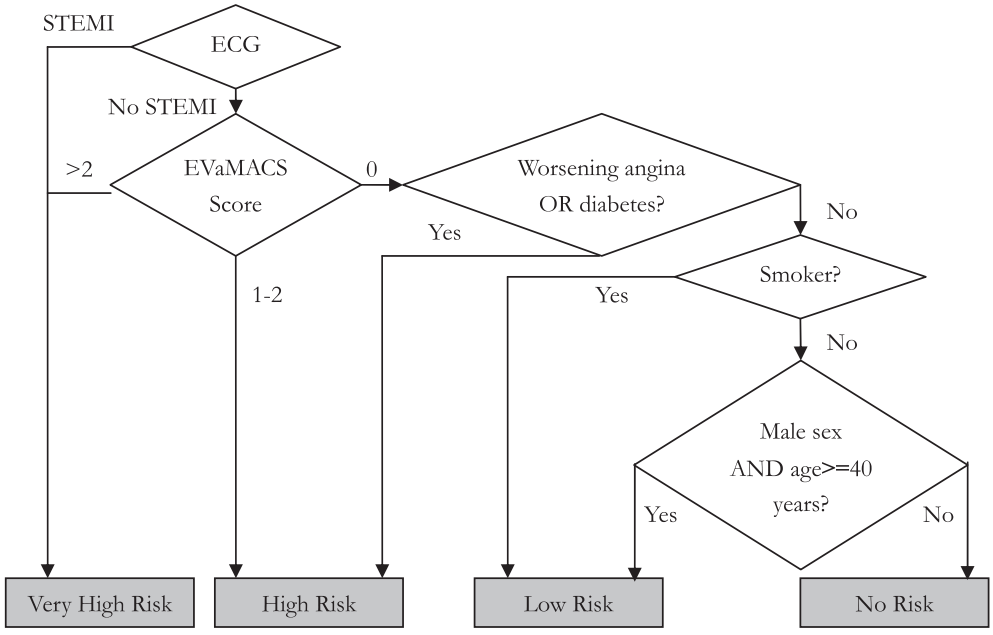


Fig. 2. System architecture of the prototype CDSS.



The EVaMACS score:

H-FABP>58ng/ml	=2
TnI>0.055ng/ml	=1
Acute ischaemic ECG features	=2

Score>2 – Very high risk

Score 1-2 – High risk

Fig. 3. A decision tree In the decision tree, the constituents of the EVaMACS score include: (a) H-FABP, which is heart fatty acid binding protein – a biomarker (blood test), levels of which are known to rise early after the onset of AMI; (b) TnI, troponin I, another biomarker, which is the biomarker of choice for diagnosing acute myocardial infarction – but we normally have to wait at least 12 hours from the time of symptom onset to do the test (levels rise late after the onset of AMI – so they are often undetectable when patients first attend the ED). for assessing clinical risk of CCP (Body, 2009).

gained with certainty, then a certain clinical risk level can be derived. However, uncertainties in both clinical knowledge and clinical data are unavoidable. As to uncertainties in clinical knowledge, take the above rule R^1 for example. For some experts, they may think that a patient will be at 'very high risk' with only 90% possibility if his/her ECG shows STEMI, and they have no exact idea

about the remaining 10% possibility. Regarding uncertainties in clinical data, take ECG for example. Most clinicians cannot always be 100% sure about their judgments on whether one ECG definitely shows STEMI or not, and they may use 'probably present', 'not sure', or 'probably not present' to express their judgments, which obviously contain uncertainties. Reasoning with traditional rules without consideration of uncertainties has its limitations, as only certain clinical data can be employed in a rule chaining process, which can only produce clinical conclusion without any uncertain information. Such a process requires clinicians to make assumptions when certain clinical information is not available.

To better handle uncertainties, we propose to use belief rules to represent clinical domain knowledge in the prototype. Since belief rules are an extension of traditional rules by addition of rule weights, antecedent attribute weights, and belief degrees as discussed in Section 3.1, we extended the traditional rules that are derived from the decision tree to 48 belief rules, where the five risk factors in the tree are considered as antecedents and the four risk levels are taken as possible consequents in each belief rule. In the constructed BRB, we assume that all belief rules have equal rule weight, all antecedent risk factors have equal weight, and the initial belief degree assigned to each possible consequent risk level is based on the statistics of available accumulated patient data. An example belief rule is as follows.

R': IF 'ECG shows no STEMI' and 'EVaMACS score is >2' and 'having angina or diabetes' and 'being a smoker' and 'being male with age > = 40', THEN {('very high risk', 95%), ('high risk', 0%), ('low risk', 0%), ('no risk', 0%)}.

In the above belief rule, the belief degrees attached to the four risk levels are based on available patient data. For example, if we have 100 patients whose clinical information fully satisfies the antecedents of the rule, and if there are only 95% patients who are known to be judged at 'very high risk', with the remaining 5% patients either having no record about their real clinical status or their status unclassified, then we can assign the initial belief degrees to the consequents (four risk levels) in the rule as shown above.

Reasoning with belief rules is more informative than reasoning with traditional rules because a distributed clinical conclusion provides a panoramic view about a patient's risk levels as judged from different aspects whilst taking a certain risk judgment as a special case. For example, we can get a risk assessment result for one patient as {(very high risk, 0.85), (high risk, 0.05), (low risk, 0.00), (no risk, 0.05)} from our system if the rules activated in the inference process are incomplete or the input patient data is incomplete. A distributed risk assessment result can help better triage patients to appropriate in-patient treatment. For example, patients may be with different disease severities even they are assessed to the same risk level. By assigning a risk score to each different risk level, an overall risk score or risk score interval can be generated for each patient based on the distributed risk assessment result, and then patients can be treated in an order according to their overall risk scores or score intervals even they are in the same risk group.

4.2.2. Inference engine

In a belief rule-based system, the inference engine is implemented using the ER algorithm. As described in Section 3.2, the ER based inference engine is used to aggregate multiple activated belief rules, each of which represents a non-linear mapping between multiple inputs and an output. In each belief rule, the inputs include parameters like rule weights, antecedent attribute weights, and matching degrees between the inputs and the rule's antecedents, and the output includes parameters like consequent belief degrees and the quantified values of consequents. The aggregated result is also represented by a belief distribution with the aggregated belief degrees associated to all consequents. The inference engine works in the following sequence: reading data for the set of parameters, calculating activation weights of all rules using (4), calculating combined belief degrees to all consequents using (5), and finally presenting the inference results to other system components.

We developed the ER-based inference engine in MATLAB environment, and it is packaged as component object model (COM) which can be called freely by main system programs developed in the. NET environment.

4.2.3. BRB training module

As discussed in Section 3.3, the aim of BRB training is to find an optimal set of parameters $(\theta_k, \delta_i, \beta_{jk})$ by minimizing the discrepancy between the system results and the sampled data. The core of the training module is a BRB optimization model. We developed the BRB training model in MATLAB environment using the *fmincon* function provided by its optimization toolbox, and we packaged the developed training model as COM. The core steps of the training process include (1) constructing objective function, (2) setting constraints for the training parameters, and (3) calling *fmincon* to search for optimal parameter set.

Regarding the objective function of the training model, we used the total mean squared error $(1/M)\sum_{m=1}^{M}(y_m - \hat{y}_m)^2$ to represent $\xi(P)$ as shown in Fig. 1. As described in Section 3.2, the diagnosis result generated by the ER based inference engine for one patient is a set of belief degrees attached to all consequents as $(D_j, \beta_j)(j = 1,\ldots,N)$ other than a single numerical value. To simplify the training process, the inferred distributed result is transformed to a single numerical value (Yang and Xu, 2002), and the transformation is implemented using $y_m = \sum_{j=1}^{N}\mu(D_j)\beta_j(m = 1,\ldots,M)$, where $\beta_j(m)$ $(j = 1,\ldots,N; m = 1,\ldots,M)$ is generated by the inference engine, and $\mu(D_j)$ $(j = 1,\ldots,N)$ is the utility value or risk score set for the jth consequent D_j. As to the observed results of the training sample $\hat{y}_m(m = 1,\ldots,M)$, the same transformation technique can be used to get numerical values if necessary.

The construction of constraints in a BRB training model depends on model requirements (Yang et al., 2007) as well as specific domain knowledge and domain experts' judgements. In the training process, the initial values of the parameter can be set by experts or randomly if no prior knowledge is available.

5. System validation

The CDSS needs to be validated using sampled or simulated clinical data. The diagnostic performances of the system using BRB and manual assessment using traditional rules need to be analyzed in the validation.

Ideally, real patient data should be used for system validation. However, due to the strict data protection regulations in the UK, we were unable to get the ethical approval for using the real patient data that had been used to develop the traditional rules for clinic risk assessment by Dr. Richard Body, as discussed in section 4.2.1. Instead, Dr. Richard Body generated a similar type of 1000 simulated datasets for patients having CCP, and the simulated datasets were used for the system validation.

As to the diagnostic test result analysis, usually, we use *sensitivity*, which is determined by the proportion of patients with the disease who were correctly identified, and *specificity*, which is determined by the proportion of negatives which are correctly identified, to assess the diagnostic test performance. However, in many diagnostic situations, when the diagnostic test results are ordinal or continuous, we cannot make a binary decision, and we need a diagnostic cut-off to discriminate positive from negative. In this case, a single pair of sensitivity and specificity values is insufficient to describe the full range of diagnostic performance

of a test, and the receiver operating characteristics (ROC) curve can then be used for evaluation of diagnostic performance. The ROC curve is a plot of sensitivity as *y* coordinate versus 1-specificity as the *x* coordinate based on different diagnostic cut-off values (Metz, 1978). An advantage of the ROC curve is that it can be used to summarize the accuracy of a diagnostic test with a single number by calculating the size of the area under the curve (AUC) (Body, 2009). The closer AUC is to 1, the better the overall diagnostic performance of the test. The AUC has been widely used for comparing diagnostic performance of different diagnostic tests and machine learning algorithms (Bradley, 1997; Jin and Ling, 2005).

In the study, we used ROC curve to evaluate performances of different clinical risk assessments conducted automatically by our belief rule-based CDSS or manually using traditional rules, and we used SPSS software (version 16.0) for plotting ROC curves. To test the statistical significance of the difference between AUCs, some specific software has been developed by researchers for AUC comparison. Frequently mentioned software in the literature include MedCalc,[1] ROCKIT,[2] and StAR[3] (Vergara et al., 2008). In the research, we chose StAR to do AUC comparison as it is online software which can be accessed freely and can meet our requirements for comparison of paired data. StAR was designed for the ROC analysis of paired data and the core of the software is a non-parametric test for the difference of the AUC that accounts for the correlation of the ROC curves. Here, paired data are data generated from those diagnostic tests in which each case in the studied sample has been tested (Metz et al., 1998), and non-parametric test means that we have no need to make assumption about the distribution of the diagnostic test result in the analysis.

In this section, simulated clinical data is discussed first in Section 5.1, and then validation of two core system components: the inference engine and the training module are discussed in Sections 5.2 and 5.3 respectively.

5.1. Data

Based on clinical experience and the collected dataset of patients with CCP, we simulated 1000 patients having CCP with different clinical risks.

The data were simulated using the following methodology. We created a database using SPSS software. For each variable, we examined prevalence data from an existing dataset of real patient data, basic details of which have been published (Body, 2009) and (Body et al., 2008, 2010a,). We used random numbers between 0 and 1, generated by SPSS, to assign values to each variable such that they were consistent with the real prevalence data. For example, from the existing dataset, approximately 68% of the population were male. Thus, a similar prevalence was reproduced in the simulated dataset by assigning a value of 'male sex' to values below 0.68 and 'female sex' to those above 0.68. Thus, while random, each variable in the simulated dataset had the same probability of a positive response to a real-life cohort of patients. In order to take account of clinical uncertainty for variables that require a degree of subjective interpretation (such as the presence or absence of ischaemic ECG changes), we stratified the data into five groups (definitely present, probably present, not sure, probably not present, definitely not present) rather than two.

The dataset has two important features that are necessary for the research. First, the dataset is close to reality. All of the variables including clinical signs and symptoms and clinical risk status in the dataset have similar positive response rates to reality and the relationship of each individual variable (clinical signs or symptoms) to the patient risk outcome is in similar manner. Second, uncertainties in doctor judgment are reflected in the dataset. For example, doctor judgment about one patient's ECG showing STEMI or not can be 'definitely present', 'probably present', 'not sure', 'probably not present', and 'definitely not present'.

In the dataset, variables used to represent one patient's clinical signs and symptoms, and outcome include '*STEMI*', '*Worsening_Angina*', '*Diabetes*', '*Smoking*', '*Sex*', '*Age*', '*EVaMACS_Score*', and '*Outcome*'. '*Outcome*' is the dependant variable, and it is used to record the composite outcomes of one patient. Other variables are independent variables that contribute to the clinical risk status and the outcomes of patients, and they are used to record status of different clinical signs or symptoms that their names stand for. In '*outcome*' column, there are only numerical values, where 1 represents that the patient had AMI or he/she died, had AMI or needed urgent coronary revascularisation within six months, and 0 represents that the patient had no real clinical risk. For '*STEMI*', various subjective judgments including 'definitely present', 'probably present', 'not sure', 'probably not present', and 'definitely not present' are used to simulate whether an ECG shows STEMI or not. For variables '*Worsening_Angina*', '*diabetes*', and '*smoking*', 1 represents 'yes' and 0 represents 'no'. For variable '*Sex*', 1 represents male and 0 represents female. In the validation study, we used both traditional and belief rules as described in Section 4.2.1 to derive clinical risk status of simulated patients based on the data recorded as independent variables in the dataset. As to observed or real clinical risk status of the simulated patients, we treated patients with outcome 1 as in 'very high risk' category and patients with outcome 0 as in 'no risk' category. Table 1 displays 5 example patient data extracted from the simulated dataset.

When we use belief rules in our prototype CDSS to make inference with the above simulated patient data which contain uncertain judgments, we need to preprocess the data as we have only certain antecedents in belief rules like rule *R*' discussed in Section 4.2.1. In the prototype implementation, we used rule-based method to transform uncertain judgments in the simulated data as follows: 'definitely present' means 100% 'present', 'probably present' means 80% 'present', 'not sure' means 50% 'present' and 50% 'not present', 'probably not present' means 80% 'not present', and 'definitely not present' means 100% 'not present'. The belief degrees here can then be used to calculate activation weight for each rule in the inference process as discussed in Section 3.2.

5.2. Inference engine validation

5.2.1. Method

Validation of the inference engine is basically composed of three main procedures. Firstly, manually assess risks for the 1000 patients using the traditional clinical rules and evaluate the overall diagnostic performance of the manual assessment. Secondly, use the simulated patient data as inputs to the system and trigger the system to do risk assessment, and evaluate the system's overall diagnostic performance. Thirdly, compare the system's diagnostic performance to the manual one. The recorded outcomes of the simulated patients are used as benchmark in the diagnostic performance analysis.

We used the ROC curve to analyze diagnostic performances of different risk assessment tests, and we used patient overall risk scores to plot ROC curves. The overall risk score for each patient is generated on the basis of (1) risk assessment result, and (2) the risk scores assigned to 'very high risk', 'high risk', 'low risk', and 'no risk'. In the study, we set a risk score of 1 to 'very high risk', 0.67 to 'high risk', 0.33 to 'low risk', and 0 to 'no risk'. For example, we can estimate one patient's overall risk score as 0.9668 if the risk assessment result for the patient is {(very high risk, 0.94), (high risk, 0.04), (low risk, 0), (no risk, 0)}.

[1] http://www.Medcalc.Be/.
[2] http://www.Radiology.Uchicago.Edu/Krl/Krl_Roc/Software_Index6.Htm.
[3] http://www.Protein.Bio.Puc.Cl/Cardex/Servers/Roc/Home.Php.

Table 1
Example patient data in the simulated dataset.

No.	Outcome	STEMI	Worsening_Angina	Diabetes	Smoking	Sex	Age	EVaMACS_Score
1	0	Probably not present	0	0	0	0	87	0
2	1	Definitely present	0	0	0	0	81	1
3	1	Probably present	0	0	1	0	80	1
4	0	Definitely not present	0	0	0	0	39	0
5	1	Not sure	1	0	0	1	61	1

5.2.2. Results

After conducting the risk assessment tests as designed in Section 5.2.1, we got two sets of overall risk scores for the 1000 patients, where one set was automatically generated by the system using the belief rules and the other set was manually produced using the traditional rules. In the ROC analysis, we used the recorded outcomes of those 1000 patients as benchmark, and we got the following two ROC curves as shown in Fig. 4 that represent the diagnostic performances of the system and the manual diagnosis using the traditional rules. The ROC curve as represented by the green line in Fig. 4 is plotted from the score set generated by the system, and the AUC is 0.7921 (95% confidence intervals 0.7586–0.8257), while the ROC curve as represented by the blue line in Fig. 4 is plotted from the score set manually produced using the traditional rules, and its AUC is 0.7525 (95% confidence intervals 0.7177–0.7873). From the two ROC curves as displayed, we can find that the diagnostic performance of the system is better than the manual diagnosis using the traditional rules. We then used StAR to do comparison of AUC for these two ROC curves, and we found that the AUC difference between these two ROC curves is statistically significant (p-value < 0.0001). Therefore, we can conclude that the diagnostic performance of the belief rule-based prototype system is better than the manual diagnosis using the traditional rules.

5.3. BRB training module validation

5.3.1. Method

The core of the training module is the BRB optimization model as discussed in Section 4.2.3. In the validation, we tried five different sets of training parameters. To reduce the complexity in conducting the training module validation, by drawing on the idea of test set cross-validation (Stone, 1974), we designed the training module validation as follows.

Firstly, split the simulated 1000 cases into training set and test set. Secondly, train the system using the training set and test the system's diagnostic performance using the test set. To find out a most suitable set of training parameters for the system, we tried the following five rounds of training: (1) T^1, training with all knowledge representation parameters and risk scores of the four consequent risk levels, (2) T^2, training with all knowledge representation parameters, (3) T^3, training with antecedent attribute weights and belief degrees, (4) T^4, training with rule weights and belief degrees, and (5) T^5, training with belief degrees. Finally, we conducted analysis of the system's diagnostic performance based on the risk assessment results generated in the second step.

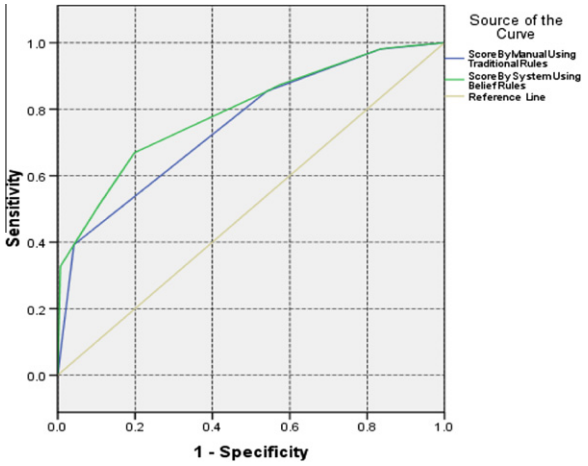


Fig. 4. ROC curves demonstrating the diagnostic performances of the system using belief rules and manual assessment using traditional rules.

To ensure that most clinical rules in the BRB can be trained in the training process and also tested in the test process, we split the simulated data into two similar sets instead of randomly splitting available data into training set and test set as in traditional test set cross-validation. It should also be noted that we used the system's diagnostic performance as an index to measure the fitness of the BRB model to clinical data, while in traditional test set cross-validation, mean squared error is a most commonly used measure for evaluating training model's performance.

5.3.2. Results

In the training process, the training module resulted in different changes to the BRB for different sets of training parameters. In the test process, based on the trained BRB after different training rounds, the system generated different sets of overall risk scores for patients in the test set. The details of the system diagnostic performance in assessing clinical risks of patients in the test set before and after each training round are presented as follows.

Before training the system using the patient data in the training set, we used the system to conduct clinical risk assessment of patients in the test set based on the initial BRB as described in Section 4.2.1, and generated a set of risk scores, denoted as 'preBRBTraining' score in the following discussion. After BRB training with different parameter sets as represented by T^1, T^2, T^3, T^4, and T^5, another five sets of risk scores were generated by the system for patients in the test set, and these scores are denoted as 'AfterT^1', 'AfterT^2', 'AfterT^3', 'AfterT^4', and 'AfterT^5' scores. Based on the system generated risk scores before and after BRB training, six ROC curves are plotted using SPSS as shown in Fig. 5. The risk score source of each curve is described in the figure. The corresponding AUC values for all the six ROC curves are shown in Table 2.

As described in Table 2, the AUC of the ROC curves representing system diagnostic performances after training are all larger than the AUC of the ROC curve before training, and the AUC of the ROC curve after training T^2 is the largest among all AUC values. However, we cannot simply conclude that training T^2 is the best training which can help the system achieve the best diagnostic performance just based on the data as shown in Table 2. To measure the significance level of the system diagnostic performance improvement after each training round, we used StAR to compare the AUC of the ROC curves as plotted in Fig. 5 and generated the AUC comparison results as shown in Table 3.

Based on the p-values as shown in Table 3, we can find that the training round of T^1(p-value = 0.0323), T^2 (p-value = 0.0198), and T^3 (p-value = 0.0076) all brought significant diagnostic performance improvement to the system, while T^3 brought the most significant performance improvement to the system. Thus we can conclude

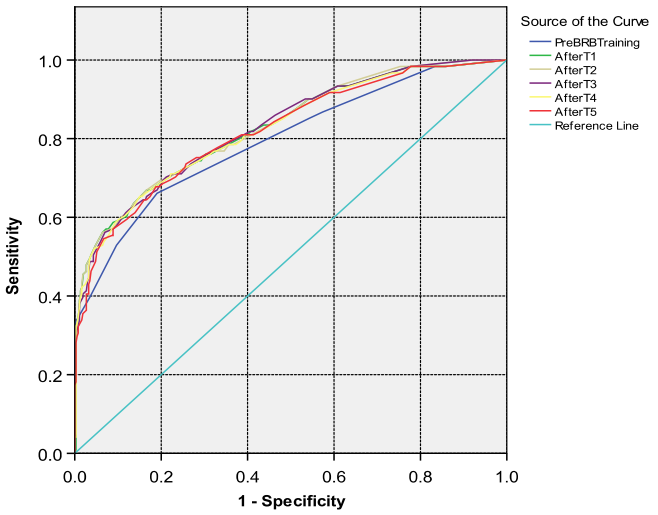


Fig. 5. ROC curves demonstrating the diagnostic performance of the system before and after BRB training.

Table 2
AUC Values of the six ROC curves.

Risk assessments	Area	Std. error	Asymptotic sig.	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
PreBRBTraining	0.7956	0.0251	0.0000	0.7464	0.8449
AfterT^1	0.8280	0.0229	0.0000	0.7830	0.8729
AfterT^2	0.8299	0.0224	0.0000	0.7860	0.8739
AfterT^3	0.8290	0.0224	0.0000	0.7851	0.8729
AfterT^4	0.8245	0.0232	0.0000	0.7791	0.8699
AfterT^5	0.8208	0.0233	0.0000	0.7752	0.8664

Table 3
AUC differences for the ROC curves (upper triangle: AUC differences; lower triangle: *p*-values).

Risk assessments	AfterT^1	AfterT^2	AfterT^3	AfterT^4	AfterT^5	PreBRBTraining
AfterT^1	AfterT^1	0.4530	0.8863	0.0034	0.0072	0.0323
AfterT^2	0.0020	AfterT^2	0.0009	0.0054	0.0092	0.0343
AfterT^3	0.0010	0.8881	AfterT^3	0.0045	0.0082	0.0334
AfterT^4	0.4984	0.2038	0.5639	AfterT^4	0.0038	0.0289
AfterT^5	0.3890	0.2326	0.0984	0.5596	AfterT^5	0.0251
PreBRBTraining	0.0323*	0.0198*	0.0076**	0.0764	0.0836	PreBRBTraining

that the BRB training module can invariably help improve the system's diagnostic performance and T^3 (training with antecedent attribute weights and belief degrees) can help improve the system's diagnostic performance most significantly.

6. Conclusion

As the delivery of health care becomes more and more demanding due to lengthened life span, today's governments and health care providers have a strong need for automatic tools such as CDSSs which can combine advances in operations research and IT to help reduce medical errors and patient safety incidents, and thus can help reduce the health care service costs caused by patient safety incidents.

In this paper, we proposed to employ a newly developed belief rule-based inference methodology – RIMER for implementation of an intelligent CDSS on risk assessment of CCP, and presented a belief rule-based CDSS prototype with knowledge learning functionality. The validation of the prototype CDSS was conducted using a set of 1000 simulated patients having CCP. Based on the developed CDSS prototype and the system validation results, we can conclude that (1) it is feasible, viable, and practical to use RIMER for implementing a CDSS, (2) the developed CDSS can provide reliable and more informative diagnosis recommendations than manual diagnosis using traditional rules when there are clinical uncertainties, and (3) the BRB system can be updated automatically by learning through accumulated clinical cases, and the trained BRB can help improve the system diagnostic performance significantly. The research partners, who are senior consultant clinicians of MRI, provided really positive judgments about the prototype from the perspectives of both knowledge owners and potential system users, and the next step of the study is to deploy the system at MRI for further clinical research.

However, the CDSS prototype presented in the paper has its limitations. Firstly, due to the fact that the clinical status of a specific patient may keep changing during the diagnosis process and the same risk factor may play different roles in predicating the patient's clinical risk after several hours or days, a dynamic clinical risk assessment model should be considered in the system. Secondly, since the structure of BRB is not always fixed and may change over time with the environment in which it is applied, not only knowledge representation parameters but also BRB structure should be considered in BRB training.

For deployment of the current CDSS prototype, however, the above limitations may not be barriers. The current prototype CDSS is designed specifically to fit in ED, where clinical risk stratification is for triage of patients to appropriate levels of in-patient care, and the belief rule-based inference methodology implemented in the system can help to handle uncertainties in patient data which may be unavailable or incomplete. To provide a real-time clinical risk status track, we need a CDSS designed for both ED triage and in-patient care, and thus we need clinical domain knowledge for assessing patient risk at different time points to construct different BRBs to support clinical risk assessment at various decision making points, and this requires domain knowledge contributions from doctors in both emergency and inpatient departments. For future research on intelligent belief rule-based CDSSs, the above mentioned areas can be considered as possible directions.

Acknowledgements

This work was supported by the UK Engineering and Physical Science Research Council (EPSRC) under the Grant No: EP/F024606/1 and EPSRC Knowledge Transfer Account scheme under the Grant No: KTA 031 of The University of Manchester. Kevin and Simon would like to acknowledge the support of the Biomedical Research Centre at Manchester Royal Infirmary. The work was also supported by Natural Science Foundation of China under the Grant No: 60736026 and Grant No: 70925004.

References

Aase, O., 1999. Clinical experience with a decision support computer program using bayes' theorem to diagnose chest pain patients. Cardiology 92, 128–134.
Aase, O., Jonsbu, J., Liestol, K., Rollag, A., Erikssen, J., 1993. Decision support by computer analysis of selected case history variables in the emergency room among patients with acute chest pain. Eur. Heart J. 14, 433–440.
Allender, S., Peto, V., Scarborough, P., Kaur, A., Rayner, M., 2008. Coronary Heart Disease Statistics. British Heart Foundation, London.
Assanelli, D., Cazzamalli, L., Stambini, M., Poeta, M.L. 1993. Correct diagnosis of chest pain by an integrated expert system. In: Proceedings Computers in Cardiology.
Bates, D.W., Cohen, M., Leape, L.L., Overhage, J.M., Shabot, M.M., Sheridan, T., 2001. Reducing the frequency of errors in medicine using information technology. J. Am. Med. Inform. Assoc. 8, 299–308.

Bhargava, H.K., Power, D.J., Sun, D., 2007. Progress in Web-based decision support technologies. Decis. Support Syst. 43, 1083–1095.
Body, R., 2008. Emergent diagnosis of acute coronary syndromes: today's challenges and tomorrow's possibilities. Resuscitation 78, 13–20.
Body, R. 2009. Clinical decision rules to enable exclusion of acute coronary syndromes in the emergency department. Faculty of Health, Psychology and Social Care. Manchester, Manchester Metropolitan University.
Body, R., Mcdowell, G., Carley, S., Mackway-Jones, K., 2008. Do risk factors for chronic coronary heart disease help diagnose acute myocardial infarction in the emergency department? Resuscitation 79, 41–45.
Body, R., Carley, S., Wibberley, C., Mcdowell, G., Ferguson, J., Mackway-Jones, K., 2010a. The value of symptoms and signs in the emergent diagnosis of acute coronary syndromes. Resuscitation 81, 281–286.
Body, R., Mcdowell, G., Carley, S., Ferguson, J., Mackway-Jones, K., 2010b. Diagnosing acute myocardial infarction with troponins: how low can you go? Emerg. Med. J. 27, 292–296.
Bradley, A.P., 1997. The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recognit. 30, 1145–1159.
Brandeau, M.L., Sainfort, F., Pierskalla, W.P., 2005. Health care delivery: current problems and future challenges. In: Brandeau, M.L., Sainfort, F., Pierskalla, W.P. (Eds.), Operations Research and Health care: A Handbook of Methods and Application. Kluwer Academic Publishers, Boston.
Collinson, P.O., Premachandram, S., Hashemi, K., Kennedy, R.L., 2000. Prospective audit of incidence of prognostically important myocardial damage in patients discharged from emergency department/ Commentary: Time for improved diagnosis and management of patients presenting with acute chest pain. BMJ 320, 1702–1705.
De Dombal, F.T., Leaper, D.J., Staniland, J.R., Mccann, A.P., Horrocks, J.C., 1972. Computer-aided diagnosis of acute abdominal pain. Brit. Med. J. 2 (5804), 5–9.
Department of Health, 2004. Seven steps to patient safety. In: National Patient Safety Agency, D.O.H. (Ed.).
Goodacre, S., Cross, E., Arnold, J., Angelini, K., Capewell, S., Nicholl, J., 2005. The health care burden of acute chest pain. Heart 91, 229–230.
Hudson, D., Cohen, M., 1987. Management of uncertainty in a medical expert system. In: Bouchon, B., Yager, R.R. (Eds.), Uncertainty in Knowledge-Based Systems: International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems. Springer, Berlin.
Hudson, D., Cohen, M., 1988. A method for determination of evidential weighting factors in a medical expert system. Uncertain. Intell. Syst.
Hudson, D.L. Cohen, M.E. 2002 Use of intelligent agents in the diagnosis of cardiac disorders. In: Bencivelli, W., Porenta, G. (Eds.) Computers in Cardiology, Tennessee IEEE.
Jin, H., Ling, C.X., 2005. Using AUC and accuracy in evaluating learning algorithms. IEEE Trans. Knowled. Data Eng. 17, 299–310.
Jonsbu, J., Aase, O., Rollag, A., Liestol, K., Erikssen, J., 1993. Prospective evaluation of an EDB-based diagnostic program to be used in patients admitted to hospital with acute chest pain. Eur. Heart J. 14, 441–446.
Kawamoto, K., Houlihan, C.A., Balas, E.A., Lobach, D.F., 2005. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ 330, 765–772.
Kong, G.L., Xu, D.L. Yang, J.B. 2009 An evidence-adaptive belief rule-based decision support system for clinical risk assessment in emergency care. In: 23rd European Conference on Operational Research, Bonn, Germany.
Lin, L., Hu, P.J.-H., Liu Sheng, O.R., 2006. A decision support system for lower back pain diagnosis: uncertainty management and clinical evaluations. Decis. Support Syst. 42, 1152–1169.
Menachemi, N., Saunders, C., Chukmaitov, A., Matthews, M.C., Brooks, R.G., 2007. Hospital adoption of information technologies and improved patient safety: a study of 98 hospitals in Florida. J. Healthcare Manag. 52, 398–410.
Metz, C.E., 1978. Basic principles of ROC analysis. Seminar. Nuclear Med. 8, 283–298.
Metz, C.E., Herman, B.A., Roe, C.A., 1998. Statistical comparison of two ROC-curve estimates from partially-paired datasets. Med. Decis. Mak. 18, 110–121.
Miller, R.A., Geissbuhler, A., 1999. Clinical diagnostic decision support systems: an overview. In: Berner, E.S. (Ed.), Clinical Decision Support Systems: Theory and Practice. Springer, New York-Verlag.
Musen, M., Shahar, Y. Shortliffe, E. 2006. Clinical Decision-Support Systems. In: Shortliffe, E.H. Cimino, J.J. (Eds.) Biomedical Informatics.
Reason, J., 2001. Understanding adverse events: the human factor. In: Charles, V. (Ed.), Clinical Risk Management: Enhancing Patient Safety, Second ed. BMJ publishing house., London.
Shafer, G., 1976. A Mathematical Theory of Evidence. Princeton University Press, Princeton.
Shortliffe, E.H., 1976. Computer-Based Medical Consultations: MYCIN. Elsevier, New York.
Sim, I., Gorman, P., Greenes, R.A., Haynes, R.B., Kaplan, B., Lehmann, H., Tang, P.C., 2001. Clinical decision support systems for the practice of evidence-based medicine. J. Am. Med. Inform. Assoc. 8, 527–534.
Sommerville, I. 2007. Software engineering, essex, Person Education.
Stone, M., 1974. Cross-validatory choice and assessment of statistical predictions. J. Royal Stat. Soc. Ser. B(Methodological) 36, 111–147.
Szolovits, P., 1995. Uncertainty and decisions in medical informatics. Methods Inform. Med. 34, 111–121.
Thomas, E.J., Brennan, T.A., 2001. Errors and Adverse Events in Medicine: An Overview. In: Charles, V. (Ed.), Clinical Risk Management: Enhancing Patient Safety. BMJ publishing house, London, Second ed.
Vergara, I., Norambuena, T., Ferrada, E., Slater, A., Melo, F., 2008. StAR: a simple tool for the statistical comparison of ROC curves. BMC Bioinform. 9, 265.
Wang, Y.M., Yang, J.B., Xu, D.L., 2006. Environmental impact assessment using the evidential reasoning approach. Eur. J. Oper. Res. 174, 1885–1913.
Warner, H.R.J., 1989. Iliad: moving medical decision-making into new frontiers. Methods Inform. Med. 28, 370–372.
Yang, J.B., 2001. Rule and utility based evidential reasoning approach for multiattribute decision analysis under uncertainties. Euro. J. Oper. Res. 131, 31–61.
Yang, J.B., Singh, M.G., 1994. An evidential reasoning approach for multiple-attribute decision making with uncertainty. IEEE T. Syst. Man Cybernet. 24, 1–18.
Yang, J.B., Xu, D.L., 2002. On the evidential reasoning algorithm for multiple attribute decision analysis under uncertainty. IEEE T. Syst. Man Cybernet. Part A–Syst. Humans 32, 289–304.
Yang, J.B., Liu, J., Wang, J., Sii, H.S., Wang, H.W., 2006. Belief rule-base inference methodology using the evidential reasoning approach – RIMER. IEEE T. Syst. Man Cybernet. Part A–Syst. Humans 36, 266–285.
Yang, J.B., Liu, J., Xu, D.L., Wang, J., Wang, H., 2007. Optimization models for training belief-rule-based systems. IEEE T. Syst. Man Cybernet. Part A–Syst. Humans 37, 569–585.
Yuan, Y., Feldhamer, S., Gafni, A., Fyfe, F., Ludwin, D., 2002. The development and evaluation of a fuzzy logic expert system for renal transplantation assignment: is this a useful tool? Euro. J. Oper. Res. 142, 152–173.