Diagnosis and Management of Cardiovascular Disease with an Intelligent Decision-Making Support System

J. Bohacik andD. N. Davis

Abstract—Cardiovascular disease is the principal cause of death in most European countries and may have a major negative impact on the patients' functional status, productivity, and quality of life. It seems an automatic decision support system could lower these negative impacts. The current development stage of a patient-centric solution for remote management and treatment of cardiovascular patients is described from the point of view of decision support. The principle of the Decision-making Support System is presented. Our prototype experimental results with Data Mining Models are also provided.

Keywords—cardiovascular disease, data mining, decision support systems, risk assessment.

1 INTRODUCTION

EUROPEANhealth care systems are facing important challenges, such as ageing populations, increase in lifestyle-related health problems and limitations of health care resources. According to [4], cardiovascular diseases have been reported as the principal cause of death in most European countries. They account for 43% of mortality among men and for 56% among women. For both men and women coronary heart disease is the most prevalent cause of cardiovascular death; while stroke is relatively more prevalent in women. In cardiovascular risk assessment, diabetes is a very important factor as diabetes patients are at high risk for cardiovascular disease. Its prevalence is still rising due to several factors; overweight being one of these factors. Other important factors are age, gender, genetic factors, clinical factors such as hypertension, and life style factors such as smoking, alcohol consumption, physical exercise and diet.

Monitoring risk factors is important for the prevention of malignant events. Three areas of prevention can be distinguished: a) prevention in the total population; b) prevention in high risk groups; and c) prevention after cardiovascular events. Prevention in the total population includes life style factors and programs targeted at various groups in diverse settings, such as schools, local communities, homes for elderly people, healthcare providers etc. Prevention in high risk groups is targeted at chronic clinical conditions, which mainly affect adults aged 55 years or over, that would otherwise increase the risk for cardiovascular events, such as hypertension and diabetes. These conditions may also have a major negative impact on the patients' functional status, productivity, and quality of life. Acute cardiovascular events such as myocardial infarction and stroke, determine mortality and, if the patient survives, define the quality of life and risk for recurrent events. Prevention in high risk groups and prevention after cardiovascular events includes both life style changes and medication.

We are focused on high risk cardiovascular patients and prevention after cardiovascular events as active participants of the Bravehealth project. The Bravehealth project proposes a patient-centric vision to cardiovascular management and treatment, providing people already diagnosed as subjects at risk with a sound solution for continuous and remote monitoring and real time prevention of malignant events. The proposed solution is made up of the following sub-systems: 1) Wearable Unit: a miniaturised sensor, able to continuously monitoring parameters needed to preform a diagnosis by means of algorithms running on it. It is possible to perform sheduled analysis of parameters and to remotely trigger the screening of signs; 2) Remote Server: it is the main interface between clinicians and the system, providing both automated support, e.g. messages to the patient generated by the system, and doctor managed supervision, allowing communication with the patients with voice, text, and chat messages; 3) Gateways: they support health care delivery and they help to monitor the general health of the patient.

The focus of this paper is decision support in the BraveHealth system. The paper is organized as follows. In section 2, sub-systems of the BraveHealth system are described from the point of view of decision support. The Decision Support System (DSS) itself is presented in section 3. Our prototype experimental results are discussed in section 4. Section 5 concludes this paper.

- *J. Bohacik is with the Department of Computer Science, Faculty of Science, University of Hull, Cottingham Road, Hull, HU6 7RX, United Kingdom and with the Department of Informatics, Faculty of Management Science and Informatics, University of Zilina, 010 26 Zilina, Slovakia. E-mail: J.Bohacik@ hull.ac.uk*
- *D.N.Davis is with the Department of Computer Science, Faculty of Science, University of Hull, Cottingham Road, Hull, HU6 7RX, United Kingdom. E-mail: D.N.Davis@hull.ac.uk.*

Manuscript received on 19June 2012 and accepted for publication on 4July 2012.

2 SUB-SYSTEMS OF THE BRAVEHEALTH SYSTEM

The BraveHealth system, for decision support, is composed of logical sub-systems as shown in Figure 1. As a result, it can be seen that decision support is required across all sub-systems: Patient Measure Devices, Mobile Patient Gateway, Remote Server, and Clinician Gateway. The communication and collaboration among these sub-systems is shown in Figure 2.Patient Measure Devices are Wearable Unit, Blood Pressure Cuff, Digital Scales, etc. The prototype of the Wearable Unit is made of discrete components. It is designed to sense biological signals of interest, amplify and filter them, convert and process them in the digital domain. It is also designed to store and transmit (un)processed data to the Mobile Patient Gateway, and to receive configuration data from the Gateway. Patient Measure Devices connect to the patient mobile via Bluetooth. The data from these devices arrives encrypted on the patient mobile (an Android device) and the data is relayed to Remote Server for processing. Alert Notifications such as Green Notification meaning the patient's condition is reverted to normal, or Episodic Measures such as the patient's systolic blood pressure at a specific time point are not processed on the patient mobile.

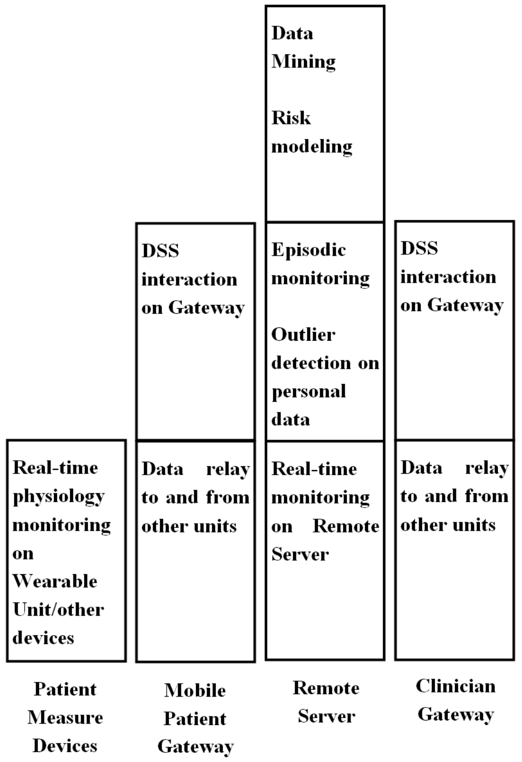

Figure 1: The BraveHealth logical architecture from the point of view of decision support.

Mobile Patient Gateway has two primary functions. The first is to act as the conduit through which Wearable Unit sends its data to Remote Server. The second is to act as a client to Remote Server so that information is displayed to provide an indication of the current health situation, prompts for medication or exercises, or requests for information, such as short surveys. Mobile Patient Gateway interfaces with Remote Server via either a desktop PC, a laptop or a tablet. From this Ethernet/Internet connection the patient mobile allows the patient to interact with the BraveHealth system. Patient configuration and information are kept on the Mobile Patient Gateway for ease of access by the patient, but this patient profile is a result of a configuration (by clinicians) that has happened on the Remote Server. In addition to the Mobile Patient Gateway, the Patient TV Gateway can be used as a client of the Remote Server and as a patient interaction device. It does not store data nor perform decision support processing other than interaction and response/query management to the Remote Server.

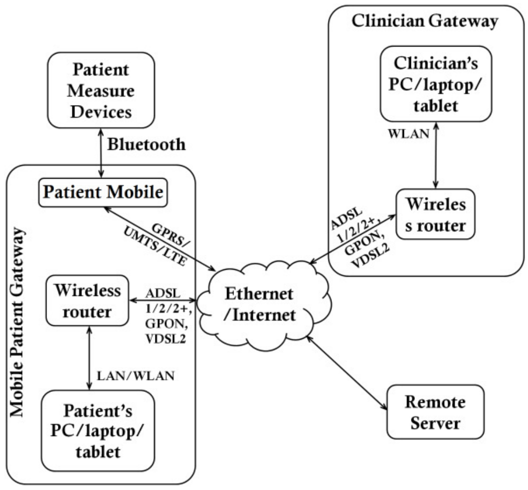

Figure 2: Working process of the BraveHealth sub-systems.

The Clinician Gateway allows the medical professionals who treat a patient to have access to the data collected from the patient's Wearable Unit, in both a trace of readings from different sensors and also a static overview of the readings. Clinicians have an indication of the status of all patients under their care, and can access the Decision Support System. Clinicians configure the BraveHealth system for a patient, including the Wearable Unit. This configuration is stored at the Remote Server. A subset of the patient data (including the Wearable Unit configuration) is stored at the Mobile Patient Gateway and it is used to configure the Wearable Unit according to the clinicians' requirements (e.g. Patient Risk Level). Episodic measures are collected according to this configuration. In normal operation, these are stored locally on the Mobile Patient Gateway and relayed to Remote Server (and stored in a database) at more over-reaching episodes. Clinician Gateway interfaces with the Remote Server via either a desktop PC, a laptop or a tablet. Any other device connects to one of these in order to interact with the BraveHealth system.

3 DECISION SUPPORT SYSTEM

The Decision Support System being built into the BraveHealth architecture is intended to provide support for both the clinician and the patient. Through the implementation of standard clinical models it ensures that routine clinical consultations are made more consistent and informative. The implementation of the BraveHealth system includes multiple Decision Support Systems that augment the functionality of the Wearable Unit (and other Patient Measure Devices) to ensure a more informative experience for both the patient and the clinician. An extended functionality to multiple Decision Support Systems augments the Telehealth functionality of the BraveHealth system and allows improved clinical risk analysis.

The BraveHealth system supports the following levels of decision support sub-systems:

- Wearable Unit Processor;
- Lightweight Decision Support System on the Mobile Patient Gateway;
- Full Decision Support System on the Remote Server;
- Decision Support for Clinicians on the Clinician Gateway.

Decision model building is done only on the Remote Server but it can be deployed across the architecture. Full Decision Support System has the capability (and models) to recommend Patient Risk Level changes. The clinicians need to decide whether they want Full Decision Support System to change Patient Risk Level or merely make a recommendation to the clinicians. Wearable Unit Processor is only capable of generating data, specific data related analyses and initiating Alert Notifications. It does not have access to the contextual information needed to determine a more appropriate status of the patient (such as the change of Patient Risk Level).

The expected functionality and knowledge models of the Decision Support System architecture are as follows:

- Clinical Models: decision support models such as alert and pathology risk models;
- Causal Models: diagnostic models that allow causal relationships to be modelled. These can be based on Clinical Models or improve on models using Bayesian networks;
- Data Mining Models: classification models deployed on the Mobile Patient Gateway and the Remote Server such as classification rules, decision trees, fuzzy rules, and neural networks.

The monitoring schedule and actions of the BraveHealth system are tailored to each patient according to the patient groupings. Thus within the BraveHealth system patients are categorized based on Patient Pathology and Patient Risk Level, with patient specific Alert Notifications. The risk profile for any patient is a combination of three factors:

- Pathology types: patients are to be assigned by clinicians as belonging to a specific pathology type, e.g. Heart Failure or Hypertension;
- Risk stratification: patients are assigned by clinicians as belonging to one of three categories: high risk, medium risk or low risk. These categories and their definitions are based on sound rules drawn from clinical practice. Movements between these levels can be monitored, or recommended. But only clinicians can change the risk status of the patient (Patient Risk Level). Full Decision Support System monitors patients, and acts as an aide to the clinicians by offering decision support (to clinicians) on a patient's current risk level by flagging up important changes. This uses expert system technologies that can be augmented (after a period of data collection within BraveHealth) through data mining;
- Alert Notifications: clinicians assign patient specific thresholds for three alert categories: Red Notification, Yellow Notification or Green Notification. Emergency Notification will be incorporated into subsequent BraveHealth releases. Full Decision Support System enhances the support provided to both the patients and the clinicians and indicates changes to the clinicians to Risk stratification based on these Alert Notifications. Green Notifications are only sent when the prior alert was Yellow Notification or Red Notification.

Alert Notifications in the BraveHealth system are defined by the clinicians as follows:

- Red Notification: a serious modification of clinical parameters requiring immediate attention;
- Yellow Notification: a non-critical but potentially dangerous modification of clinical parameters has appeared;
- Green Notification: the patient's condition is reverted to the normal status;
- Emergency Notification: a set of pathological conditions which represent a life threatening event. When one of these conditions applies, the emergency protocol is activated. Emergency Notifications are more than a special case of Red Notifications and will be introduced into BraveHealth after the initial prototype.

4 Experimental Results

At the current state we have done some prototype experiments with Alert Notifications and Data Mining Models. Experiments were carried out with our Java software tool which is being developed with the intention of its permanent integrational the Remote Server of the BraveHealth system. Some core algorithms are implemented in external libraries: Netica™[6] and Weka [7]. In this section, the results of our experiments with Data Mining Models are discussed. The performance of Data Mining Models is measured with sensitivity = tp/(tp + fn), specificity = tn/(tn + fp), and accuracy = (tp+tn)/(tp+fp+fn+tn). In the formulas, tp/ fp/ fn/ tn is the number of true positives/false positives/false negatives/true negatives. A low risk patient is considered negative and a high risk patient is considered positive. Medium risk patients are not recognized at our prototype experiments. Values tp, fp, fn and tn are computed during 10-fold cross-validation. As a dataset, a group of 839 patients is used. The patients are described by 17 attributes: *Age*, *Gender*, *Heart disease*, *Diabetes*, *Stroke*, *Side*, *Respiratory problem*, *Renal failure*, *ASA grade*, *Hypertension*, *ECG*, *Duration*, *Blood loss*, *Shunt*, *Patch*, *Coronary artery bypass surgery*, and *Consultant*. The attributes are considered equally important initially. Some of them are considered more important than others in Data Mining Models, but the decision about their importance is based on particular Data Mining Models and collected data about patients.

TABLE 1
EXPERIMENTAL RESULTS

Method	SEN (%)	SPEC (%)	ACC (%)
Bayes	7.94	97.48	84.03
C4.5	4.76	98.60	84.51
LVE	65.08	75.88	74.26
MCA-T-T	60.32	96.07	90.70
MLP	15.08	89.62	78.43
MMI	77.78	89.62	87.84
NNC	15.08	90.18	78.90
TreeBayesNet	77.78	96.63	93.80

Our experimental results are presented in Table 1. The meaning of the used symbols is as follows.Sensitivity (SEN), specificity (SPEC), accuracy (ACC) are measures of the performance of Data Mining Models in percentages(%). Bayes denotes a Bayesian network implemented in Weka as class BayesNet. The Bayesian network represents a joint probability distribution over a set of categorical attributes. Numerical attributes are discretized. It consists of a directed acyclic graph and conditional probability tables. It allows the computation of the (joint) posterior probability distribution of any subset of unobserved assignments of values to attributes, which makes it possible to use for classification. C4.5 is a decision tree classifier implemented in Weka as class J48.The decision tree has two types of nodes: a) the root and internal nodes (associated with an attribute); b) leaf nodes (associated with a class). Its creation is based on searching attributes for potential associations with nodes on the basis of information they bring.Basically, each non-leaf node has an outgoing branch for each possible categorical value/subset of numerical values of the attribute. Risk for a patient is determined using the decision tree, beginning with the root, successive internal nodes are visited until a leaf node is reached. The leaf node contains the risk. LVE is a fuzzy rule classifier based on linguistic variable elimination [1]. First it transforms attributes into variables and then it computes membership degrees on the basis of the data about patients.Then it eliminates the least important attribute in a way that leads to dividing the data into two groups with subsets of the variables and with minimal inconsistences between the membership degrees for the variables and the class variable. It continues in the groups until no further elimination is considered important and a set of fuzzy rules is formed. The formed fuzzy rules are used for risk estimation.MCA-T-T transforms attributes into variables and then it computes membership degrees. It makes a fuzzy decision tree which is transformed into fuzzy rules. It uses classification ambiguity as a criterion for association of a variable with the node of the tree and choses the variable with its lowest value.Fuzzy rules acquired from the tree are simplified and the degree of truthfulness of fuzzy rules is kept in simplification [8]. MLP is a neural network classifier using multilayer perception implemented in Weka as class MultilayerPerception.It consists of interconnected neurons. A neuron takes positive and negative numerical values from other neurons and when the weighted sum of the stimuli is greater than a given threshold value, it activates itself. MMI is similar to MCA-T-T, but it uses maximization of mutual information as a criterion for association of a variable with a node [5]. Also, there is no simplification of the acquired fuzzy rules. NNC is a nearest neighbour classifier using non-tested generalized examples [7] implemented in Weka as class NNge.It assumes known cardiovascular patients correspond to points in spaceR^n. When the risk for a new cardiovascular patient is being determined, k-nearest known cardiovascular patients to the new one are found and they are used with a weight.TreeBayesNet denotes a Bayesian network implemented with Netica™ and learnt on the basis of a C4.5 decision tree [6].

The methods in Table 1 are evaluated so that minimizing life-threatening situations, minimizing costs and maximizing interpretability are preferred. Life-threating situations appear when patients at high risk are considered low risk, which is measured by sensitivity. This risk should be minimized and so sensitivity should be maximized. Costs are increased when low risk patients are treated as if they were high risk, which is measured by specificity. Interpretability is subjective, but it is obvious that the interpretability of neural networks is worse than the interpretability of other compared methods. Also, if we consider e.g. a group of discovered fuzzy rules, a better interpretability is achieved when the number of fuzzy rules is smaller and the expressions in conditions have fewer assignments of linguistic terms to linguistic variables. Our classifier TreeBayesNet gives the highest sensitivity 77.78%, the highest specificity 96.63%, and the highest accuracy 93.80%. Its interpretability is also good as it is a Bayesian network, which is familiar to many clinicians. When only fuzzy rule classifiers are taken into consideration, the best combination of minimization of life-threatening situations, minimization of costs and maximization of interpretability is achieved by MCA-T-T.

5 CONCLUSION

A prototype design and recent research on a patient-centric solution for remote management and treatment of cardiovascular patients and its Decision Support System architecture are presented in this paper. The solution is composed of four sub-systems: Patient Measure Devices, Mobile Patient Gateway, Remote Server and Clinician Gateway. Decision support is required on all of the sub-systems. It is done by the following decision support sub-systems: Wearable Unit Processor, Lightweight Decision Support System on the Mobile Patient Gateway, Full Decision Support System capable of (re-)building knowledge models on the Remote Server, and Decision Support for clinicians on the Clinician Gateway. Cardiovascular patients are divided into three risk levels and also according to a specific pathology type. A patient is monitored in real-time and the obtained data is used by decision support sub-systems. With the support of Clinical Models, Causal Models and Data Mining Models, they can trigger

Alert Notifications meaning different levels of emergency or they can suggest a change of the patient's risk level to clinicians. Our prototype experiments with Data Mining Models are analysed further. A model using a Bayesian network learnt on the basis of a decision tree seems to have the best results. Its sensitivity measuring how unlikely life-threatening situations are is 77.78% and its specificity measuring how the costs of treatment are minimized is 96.63%. Fuzzy rule classifiers also achieve promising results. Among them, a fuzzy rule classifier based on minimization of classification ambiguity achieves the best results with sensitivity 60.32% and specificity 96.07%. Further research will include integration of Data Mining Models into alert rules for Alert Notifications.

ACKNOWLEDGMENT

This work is funded by the European Commission's 7th Framework Program: BRAVEHEALTH FP7-ICT-2009-4, Objective ICT-2009.5.1: Personal Health Systems: a) Minimally invasive systems and ICT-enabled artificial organs: b) Cardiovascular diseases.

REFERENCES

[1] J. Bohacik, "Discovering fuzzy rules in databases with linguistic variable elimination,"*Neural Network World*, vol. 20, no. 1, pp. 45-61, 2010.

[2] J. Bohacik, D. N. Davis: Estimation of cardiovascular patient risk with a Bayesian network, *Proc. Ninth European conference of young research and scientific workers (TRANSCOM 2011)*, pp. 37-40, 2011.

[3] M. Brent, *Instance-Based Learning: Nearest Neighbour with Generalization*, Master's thesis (citeseerx.ist.psu.edu): The University of Waikato, 1995.

[4] J.v.Lieshout, M.Wensing, Grol, R., *Prevention of cardiovascular diseases: The role of primary care in Europe (1st edition)*, Electronic book (www.epa-cardio.eu): Bertelsmann Stiftung and Centre for Quality of Care Research, EPA Cardio project, pp. 4-14, 2008.

[5] V., Levashenko, E. Zaitseva, "Usage of new information estimations for induction of fuzzy decision trees," *Proc. Third International Conference on Intelligent Data Engineering and Automated Learning (IDEAL '02)*, pp. 493-499, 2002.

[6] Norsys Software Corp.,"Netica™ Application,"http://www.norsys.com/netica.html, 2011.

[7] Witten, I. H., Frank, E., Hall, M. A., "Data Mining: Practical Machine Learning Tools and Techniques (3rd edition)," Burlington: Morgan Kaufman., 2011.

[8] Y. Yuan, M. J. Shaw, "Induction of fuzzy decision trees,"*Fuzzy sets and systems*, vol. 69, no. 2, pp. 125-139, 1995.

J. Bohacik received a Master's degree (Ing.) in Information and Control Systems in 2006 and a Doctoral degree (PhD.) in Applied Informatics (Fuzzy Rules and Knowledge Discovery in Databases) in 2009, both from the University of Zilina, Slovakia. During his PhD studies he was in several government and EU research stays in Finland, Portugal and Japan.After graduation he worked as Assistant Professor at the University of Zilina, Slovakia for one yearand now he has been working as Research Assistant at the University of Hull, UK for about two years. In 2012, he was conferred a professional degree of European Engineer (EurIng)by the European Federation of National Engineering Associations and The Association of Slovak Scientific and Technological Societies. He is a (co-)author of 17 published papers in journals and conference proceedings. He has been a principal holder of two faculty scientific grants, a participant in one national scientific grant and a participant in one international scientific grant. His research interests include medical systems, data mining, machine learning, artificial intelligence and applications of notions of fuzzy logic. He is a member of the European Society of Cardiology.

D. N. Davis is Director of Research in the Department of Computer Science at the University of Hull, UK. He has a B.Sc. in Psychology, M.Sc. in Knowledge Base Systems and Ph.D. in Diagnostic and Investigative Medicine (The Application of AI in Medicine). He has over 20 years experience in artificial intelligence systems. Successful applications include classification and computer vision problems in business, medicine and geology. He has over 20 years experience in data mining, first working on Decision Trees at Edinburgh and the Turing Institute, Glasgow in 1987. Medical applications of data mining and decision support include Urology, Orthodontics, Stomach Cancer, Cervical Cancer, Bowel cancer, Neuroanatomy and Cardiovascular medicine. He has published widely with over 100 publications.