

SUPPORTING CLINICAL PROFESSIONALS IN DECISION-MAKING FOR PATIENTS WITH CHRONIC DISEASES

Mitja Luštrek¹, Božidara Cvetković¹, Maurizio Bordone², Eduardo Soudah², Carlos Caveró³,
Juan Mario Rodríguez³, Aitor Moreno⁴, Alexander Brasaola⁴, Paolo Emilio Puddu⁵

¹ Jožef Stefan Institute, Jamova cesta 39, 1000 Ljubljana, Slovenia

² CIMNE, C/ Gran Capitán s/n, 08034 Barcelona, Spain

³ Atos, Albarracín 25, 28037 Madrid, Spain

⁴ Ibermática, Parque Tecnológico de Bizkaia Edif. 501, 48160 Derio, Spain

⁵ University of Rome “La Sapienza”, Viale del Policlinico 155, Rome 00161, Italy

E-mail: mitja.lustrek@ijs.si; phone: +386 1 4773380; fax: +386 1 4773131

ABSTRACT

To deal with the large amount of data produced by telemonitoring of patients with chronic diseases, a decision support system (DSS) was developed. The DSS uses sensor data and the data from a patient’s electronic health record as the input. It assesses the risk to the patient’s health using three approaches. The first approach exploits the existing medical knowledge, the second approach uses supervised machine learning, and the third approach simply detects anomalies in the values of the monitored parameters. The risk assessment can show the contribution of the individual monitored parameters to the risk, and can be tailored by the doctor to each patient. The assessed risk and the raw input data can be used to trigger alerts. Finally, following the principles of evidence-based medicine, the DSS facilitates the consultation of medical literature when needed.

1 INTRODUCTION

The amount of data produced by telemonitoring solutions can be overwhelming, so using it for clinical decision-making is difficult. When telemonitoring data is combined with data obtained by traditional means, the problem becomes even larger. In a European project aiming to integrate telemonitoring into the clinical workflow, we tackle this problem by a decision support system (DSS).

In the project, a patient is equipped with a wearable ECG, temperature, sweating and activity sensors. The data produced by the sensors is sent to the doctor, where it is combined with the data from the patient’s health record. The DSS system uses all the data to automatically assess the risk to the patient’s health, helps the doctor understand its assessment, and automatically generates alerts to warn the doctor of adverse events. It also offers personalization, allowing the doctor to tailor the risk assessment and the alert triggers to each patient. Finally, the DSS facilitates the consultation of medical literature when needed. For the most part, the DSS can support the management of any chronic disease. Our test case, however, is the congestive heart failure (CHF), and the choice of sensors and the expert knowledge contained in the DSS reflects that.

The architecture of the DSS is shown in Figure 1. The risk assessment module consists of three submodules using different approaches to risk assessment. The module receives as inputs the values of the monitored parameters from the sensors and electronic health records, and its output can be used to trigger alerts. The parameter values from the sensors and electronic health records can also trigger alerts on their own. The configuration module is used to tailor the risk assessment to each patient. The literature consultation module searches external resources (databases, ontologies ...), but it can also use text from the electronic health records to provide context for the search.

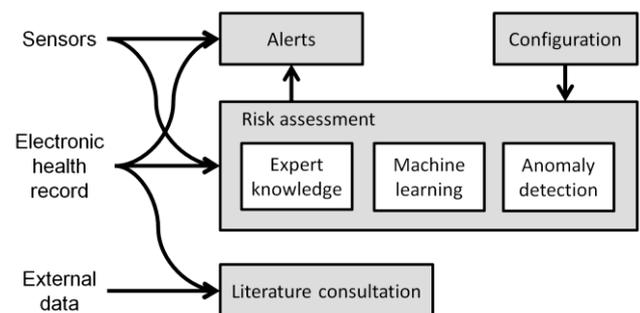


Figure 1. *The architecture of the DSS.*

The DSS supports the clinical workflow as follows:

1. The doctor starts examining the condition of a patient. He/she may do this either because of an alert or on a voluntary basis.
2. The doctor examines the patient’s current risk. The risk is computed by the risk assessment module, and its history up till the current moment may be examined.
3. If the risk is high, the doctor looks for reasons. The risk assessment module provides the monitored parameters ordered by their contribution to the risk.
4. The doctor may need further information. In this case the medical literature may be consulted using the relevant module.

The doctor may reconfigure the risk assessment module and the alerts.

2 RISK ASSESSMENT

The risk assessment is performed using three different approaches, each of which has its advantages. The first approach uses existing medical knowledge. This makes it easier for the doctors to understand and trust, but it requires the participation of experts to create it. The second approach uses supervised machine learning (ML), and is appropriate for cases where labeled data consisting of the values of the monitored parameters, and the true risk (based on hospitalizations and other adverse events) are available. The third approach observes anomalies in the values of the monitored parameters. Such anomalies are not necessarily caused by an increased risk, so the approach can be misleading, but it requires only knowledge on relevant parameters and only some data that is not anomalous.

2.1 Expert Knowledge

General Approach. We decided to construct three risk assessment models for different time horizons: long-, medium- and short-term. The following information was required on each monitored parameter:

- The minimum and maximum expected value of the parameter: p_{\min} , p_{\max} .
- The relation between the parameter value and the risk: (+) if the risk increases with the value, (-) if it decreases, or (U) if the risk is low at moderate values and high at extreme values (in this case the lowest-risk parameter value p_{mid} is also required).
- A low threshold, which separates the parameter values corresponding to low risk (green) from those corresponding to medium risk (yellow), and a high threshold, which separates the latter from those corresponding to high risk (red).
- The importance: low, medium or high.
- The frequency at which the parameter changes: low – fairly static (e.g., age), medium – measured during regular visits to the doctor (e.g., cholesterol level), or high – measured continuously by telemonitoring (e.g., heart rate).

Each parameter value first had to be transformed into a risk value. This was done by linearly scaling the parameter value from the $[p_{\min}, p_{\max}]$ interval to the $[0, 1]$ interval, applying the appropriate sign (+, -). For the U-shaped parameters, the linear scaling was performed separately on each side of the lowest-risk parameter value p_{mid} .

The risk values belonging to individual parameters were combined into an overall risk value. Since the risk values were only human estimates, we used a simple weighted sum. To obtain the thresholds between the green, yellow and red areas for the overall risk, the same summation was applied to the individual risk thresholds.

Each weight in the overall-risk weighted sum was the product of an importance weight and a model-specific weight. The importance weight was 1/3 for low-importance,

1 for medium-importance and 1.5 for high-importance parameters. In the long-term model, the model-specific weight of all the parameters was 1. In the medium-term model, the model-specific weight of the low-frequency parameters was 1/3. For the numeric medium-frequency parameters, the slope of change over time was also included. In the short-term model, the model-specific weight of the low-frequency parameters was 1/9, and of the medium-frequency parameters 1/3. For the medium-frequency parameters, the slope was again included, and for the high-frequency parameters, both the slope and standard deviation were included.

Expert Knowledge on CHF. To gather the expert knowledge on the CHF, first an extensive search of the medical literature was performed. It yielded over 60 parameters considered potential risk factors. The parameters were demographic, clinical, biological, pharmacological, angiographic, ambient and ECG-related. With the exception of the parameter importances, the information required for the risk assessment models was taken from the literature or provided by a reference expert.

To obtain the importances of the parameters found in the literature, a survey was conducted among the European opinion leaders in cardiology. The survey consisted of two parts. The first part dealt with the parameters about which the evidence was weak. The respondents were asked to express their agreement with the literature on the scale of 1–5. The second part dealt with the parameters with strong evidence, and the respondents were asked for disagreement (1) or agreement (5). Based on 32 responses, the parameters importances were assigned as follows: low for those on which the average agreement was below 3, medium for those on which the agreement was between 3 and 4, and high for those on which the agreement was above 4.

2.2 Machine Learning

We used the following five-step methodology for risk assessment using supervised machine learning (ML):

- Training data preparation: patients' data consisting of the values of the monitored parameters, and the true risk (based on actual hospitalizations and other adverse events) were gathered.
- Feature reduction: the most relevant parameters were determined, resulting in a reduced feature set. The training data were decorrelated to avoid dependencies among the features.
- ML model design: the ML architecture was optimized and the hyper-parameters set.
- Training: the ML model was trained using the appropriate ML/optimization methods.
- Validation: the ML model was validated in order to guarantee good predictive properties on unseen data.

The ML model must be able to cope with the non-linear, fuzzy and generally complex relations between the monitored parameters and the risk. ANNs are particularly

suitable for solving such problems. They are biologically inspired computational models which consist of a network composed of artificial neurons. Each neuron includes a function which returns an output based on its inputs, each of which is associated with a weight. During training, the weights are adjusted so that the final output best approximates the sought value, in our case the risk. Based on the experience with similar applications [7][8], our ANN used the multilayer perceptron architecture with three layers: the input layer consisting of one neuron for each of the relevant monitored parameters, the hidden layer, and the output layer consisting of one neuron predicting the risk.

The network was tested on a public dataset describing heart disease [4]. The network yielded good predictions, having the accuracy of 85 %.

2.3 Anomaly Detection

When a patient's health is normal, the monitored parameters usually follow recurrent patterns. For example, the heart rate is between 60 and 80 beats/minute, unless the patient exercises, in which case it is higher, but so is the activity level. Such patterns can be learned, and when a new pattern – an anomaly – is detected, the doctor is alerted. If the doctor judges the new pattern to be normal, he/she can indicate this, and the pattern will no longer be considered anomalous.

We used the Local Outlier Factor (LOF) algorithm [1] to detect anomalies. The algorithm compares the density of data instances around a given instance A with the density around A 's neighbors. If the former is low compared to the latter, it means that A is relatively isolated – that it is an outlier. Such outliers are considered anomalous. The LOF algorithm assigns a so-called LOF value to each instance, which indicates the degree of its anomalousness. We extended the algorithm to compute the degree of anomalousness of individual parameters, so that their contribution to the overall anomalousness could be observed. This was accomplished by computing the LOF values in the usual way, except that the density of instances was computed using only their distances with respect to the parameter of interest.

To use the LOF algorithm for risk assessment, it must be trained on a number of instances consisting of the parameters of a patient when his/her risk is normal. After the training data is processed, the parameters of the algorithm must be set: the number of neighbors to consider for estimating the density of instances, the low threshold separating the LOF values belonging to the green area from those belonging to the yellow area, and the high threshold between the yellow and red areas.

The anomaly detection was tested on a dataset of five test subjects performing a range of activities, and a subset of the parameters to be monitored in the project [5]. The performance of the LOF algorithm was tested with 1, 2, 3, 4 and 5 neighbors. They were compared using the area under

the ROC curve, and 1 neighbor turned out to be best. This may have been due to the modest amount of data available.

In order to set the thresholds, some anomalous data was needed. Since the test subjects were healthy, anomalous data was generated by randomly replacing some parameter values during one activity with those during another. For example, the heart rate during walking was replaced with the heart rate during running, while the other parameters remained unchanged. The low and the high thresholds were then set so that at most 5 % of the instances in the green area were considered anomalous, and at most 5 % of the instances in the red area were considered normal. As a consequence, roughly 80 % of the instances (depending on the subject) were in their proper area (normal in green, anomalous in red). That way a doctor could be fairly certain that if a patient's condition is green, the risk is low, and if it is red, the risk is high, while the yellow area represented uncertainty.

3 CONFIGURATION AND ALERTS

These two modules deal with the personalization issues, adapting the operation of the DSS to each patient as well as to the doctor's needs.

Configuration. In this module the monitored parameters used by the risk assessment module are set for each patient, since the risk to different patients' health may depend on different parameters. Furthermore, the thresholds between the green, yellow and red areas can be set. The risk assessment module provides default thresholds as described in the previous section, but the doctor may modify them if he/she wishes.

Alerts. The alerts module contains a rules engine implemented on the Drools platform [3]. It was chosen because it is open-source and allows easy implementation of complex rules. The module supports two tasks:

- Rule set configuration, which allows the doctor to configure rule sets to be applied on parameters and risks. The rules are inserted into the Drools engine to raise the corresponding actions.
- Alerts configuration, which allows the doctor to associate the rules with alerts (SMS, email etc.). The doctor can personalize the alerts depending on the rules triggered and the situation of the patient.

4 LITERATURE CONSULTATION

The literature consultation module provides the doctor with information from outside sources when making clinical decisions. In this it follows the principles of evidence-based medicine (EBM). EBM is a scientific method that seeks to assess the strength of evidence of the risks and benefits of treatments (or lack thereof) and diagnostic tests.

The goal of the literature consultation module was to offer an intelligent search engine for scientific literature, which extensively employs the tools and resources of the semantic

web and provides accurate relevance ranking of the search results. The starting point for a search is a query written either in natural language or in a structured format. The query is mapped onto the medical ontology SNOMED [9]. This ontology allows the incorporation of synonyms, antonyms and other related terms to produce a more expressive query. Furthermore, it provides the relationships between different pathologies, diagnoses and procedures, which form a context for the search. Additional context can be provided by an arbitrary natural-language text, which may be derived from a patient's electronic health record. The query is finally used to search multiple databases. Currently included are PubMed/HubMed [6] and The Cochrane Library [10], but other can be added easily.

Once the publications matching the query are obtained, they are ranked by their relevance. This is accomplished using both the frequency of the query terms, and the annotations by the users of the DSS, since the literature consultation module provides tools for the users to evaluate, tag and comment the publications. The publications are also annotated with the terms of the SNOMED ontology.

The information from the literature is presented according to the Haynes pyramid shown in Figure 2. The pyramid has individual published studies at the base. These can be summarized (automatically) into synopses. From them, syntheses are constructed, obtained from systematic reviews, for which the Cochrane Library can be used. On the next level are synopses that present a snapshot of several individual studies, such as those found in secondary journals. The summaries are the next step, obtained by integrating the best available lower levels, and represent clinical guidelines or summary documents, such as those in Clinical Evidence [2]. On top of the pyramid are the systems, with the individual characteristics of the patients linked automatically to the relevant evidence. The DSS presented in this paper is a part of such a system.

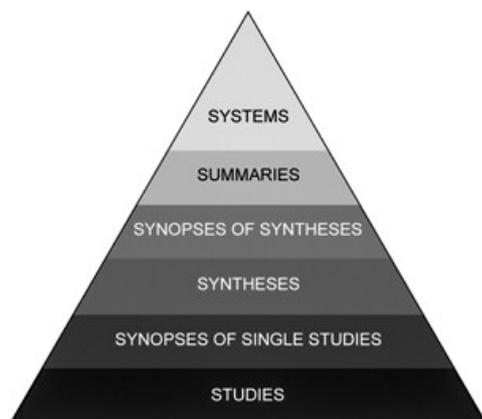


Figure 2. *Haynes pyramid.*

5 CONCLUSION

In this paper we presented a DSS for the management of chronic diseases using telemonitoring. Since the DSS was designed with CHF in mind, the sensors used for

telemonitoring (ECG, temperature, sweating and activity) are suitable for CHF patients, and the expert knowledge contained in the system also pertains to CHF. The DSS is otherwise general.

The main module of the DSS deals with risk assessment. It consists of three submodules using different approaches. The three-approach strategy was adopted because each approach has its merits, which will be evaluated during the observational study planned at the end of the project. After the evaluation, the outputs of the submodules may be merged to provide a joint decision, or one or two of the submodules may be removed if their performance proves inadequate. The risk assessment module is supported by a configuration and alerts modules. When the DSS is integrated into a clinical information system, the alerts module can be used to initiate communication with patients, order specific laboratory analyses etc.

The final module facilitates the consultation of medical literature. It follows the principles of EBM by searching external sources for published evidence supporting clinical decisions. This module provides opportunity for future work on better computer understanding of the consulted literature, saving the doctor the time needed for deep reading of the large number of found publications.

Acknowledgments

This research described in this paper was carried out in the CHIRON project, which is co-funded by the ARTEMIS Joint Undertaking (grant agreement # 2009-1-100228) and by national authorities.

References

- [1] Breunig, M. M., Kriegel, H.-P., Ng, R. T., Sander, J.: LOF: Identifying Density-based Local Outliers. In: ACM SIGMOD International Conference on Management of Data, ACM (2000)
- [2] Clinical Evidence, <http://clinicalevidence.bmj.com>
- [3] Drools – The Business Logic Integration Platform, <http://www.jboss.org/drools>
- [4] Heart Disease Data Set, <http://archive.ics.uci.edu/ml/datasets/Heart+Disease>
- [5] Luštrek, M., Cvetković, B., Kozina, S.: Energy Expenditure Estimation with Wearable Accelerometers. In: IEEE International Symposium on Circuits and Systems, IEEE (2012)
- [6] PubMed, <http://www.ncbi.nlm.nih.gov/pubmed>
- [7] RAMFLOOD project, <http://www.cimne.com/ramflood>
- [8] RAMWASS project, <http://www.cimne.com/ramwass>
- [9] Systematized Nomenclature of Medicine-Clinical Terms, <http://www.ihtsdo.org/snomed-ct>
- [10] The Cochrane Library, <http://www.thecochranelibrary.com>