Remote Health Monitoring of heart failure with data mining via CART method on HRV features
Leandro Pecchia, Paolo Melillo and Marcello Bracale

The published version of this paper can be found at: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5638128

Copyright © IEEE. Self-archiving by authors on their own personal servers or the servers of their institutions or employers is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org.

Abstract—Disease Management Programs (DMPs), which use no advanced ICT, are as effective as telemedicine but more efficient because less costly. We proposed a platform to enhance effectiveness and efficiency of home monitoring using data mining for early detection of any worsening in patient’s condition. These worsening could require more complex and expensive care if not recognized. In this paper, we briefly describe the Remote Health Monitoring (RHM) platform we designed and realized, which supports Heart Failure (HF) severity assessment offering functions of data mining based on Classification and Regression Tree (CART) method. The system developed achieved accuracy and a precision respectively of 96.39% and 100.00% in detecting HF and of 79.31% and 82.35% in distinguish severe versus mild HF. These preliminary results were achieved on public databases of signals to improve their reproducibility. Clinical trials involving local patients are still running and will require longer experimentation.

Index Terms— Home Monitoring, HRV, Heart Failure, data mining, CART.

I. INTRODUCTION

Given the rapidly growing aging population, the increased burden of chronic diseases and the increasing healthcare costs, there is an urgent need for the development, implementation and deployment, in everyday medical practice, of new models of healthcare services. In this scenario ICT, and especially Home Monitoring (HM) [1] and Data Mining (DM) [2], play an important role. DM is the computer-assisted process of digging through and analyzing a large quantity of data [3] in order to extract meaningful knowledge and to identify phenomena faster and better than human experts [4]. As regards HM, although a wide literature describes technical solutions, the evidence of ICT cost-effectiveness is limited [5] and only a few studies compare HM with other models of DMPs [6]. DMPs are more cost-effective [7] than ambulatory follow-up, which is the gold standard [8], without using costly technologies, which are not familiar to the elderly. Also HM is reported to be more effective [9] than follow-up. Nonetheless, HM is equally effective than DMPs, but less efficient because about five times more costly than DMPs and about 20 times more costly than ambulatory follow-up [10]. This led us to search for new models of HM, which incorporates further intelligent and automatic systems/services to exceed DMPs in effectiveness, offering advanced functionalities for early detection of any worsening in patient’s condition, which could otherwise require more complex and expensive care.

Among cardiovascular pathologies, HF is one of the most studied both for HM and for DM, perhaps because it has a considerable impact on healthcare costs [11], being chronic, degenerative, age related [12], and a leading cause of the elderly hospitalization [13]. Its severity can be measured with the symptomatic classification scale of the New York Heart Association (NYHA) that is widely used and hotly debated [14]. One of the most promising methods to study HF is the Heart Rate Variability (HRV), a non-invasive measure, which reflects the variation over time of the interval between consecutive heartbeats [14]. Previous studies showed that patients affected by HF present a depressed HRV [15, 16,17]. Many studies applied data mining to HRV measures for the prognosis of HF, in particular as a predictor of the risk of mortality [18]. Fewer studies used such methods to detect HF [19, 20]. In previous studies, we investigated how short-term HRV features vary according to HF severity [21] and their power in detecting HF patients [22]. In the former, we used statistic methods, while in the latter CART. In this paper, we...
presented two CARTs we integrated in a telemedicine platform to detect HF and assess its severity. The CART [23] method iteratively splits the dataset, according to a criterion that maximizes the separation of the classes, producing a tree-like decision structure. We chose this method because it requires no assumptions regarding the underlying distribution of features’ values and can easily be expressed as logical “if...then” rules. This is important because in medical applications the intelligibility of the method is needed [24], whilst other powerful methods of DM are not easy for humans to understand.

In this paper, we present the system we developed for Remote Health Monitoring (RHM) of patients suffering from HF, which includes advanced functionalities of DM for continuous patient monitoring. The clinical goal was the early detection of any worsening in patient’s condition, with automatic “HF severity assessment” using DM via CART classifiers, assuming that during worsening, patients will gradually show characteristic of a more severe HF. The system, developed in the last three years is at this moment undergoing randomized controlled trials involving real patients enrolled ad hoc. The aim of this paper was to describe briefly the platform, to present methods employed and to present the preliminary results of the data mining for HF detection and severity assessment. The results here described were obtained testing the system with biomedical signals from public databases, in order to allow other scientists to reproduce them, and because clinical trials involving local patients were still running.

II. PLATFORM DESCRIPTION

The system designing followed the so-called “Three Tier architecture”. Functionally, the platform consisted of three parts, called “Areas”: “Client Area” (CA) acting as presentation tier, “Server Area” (SA) as business tier and data tier and the “Web Service Area” (WSA) as pure business tier level. The CA aimed to present and to collect data using devices, which differed according to the users and the scenarios. The SA aimed to manage, store and retrieve data and included the Electronic Patient Record (EPR) and an Interactive Voice Response (IVR) [25], which acted as audio user interface. The WSA was used for raw data processing, signal analysis and data mining. IVR allowed users less skilled with web technologies to insert daily ECG records and physiological parameters (pressure, weight and temperature). The IVR, after a login, gave the user all the instructions and recommendations to send data, repeating, when possible, the entered values and asking for further confirmation.

The devices tested and integrated in the system, varied according to the scenario, going from user-friendly ones for self-recording of signals/signs to professional multi-parametric monitors, recording ECG, blood pressure (BP), hart rate (HR), SpO2, temperature (T), Galvanic Skin Response (GSR), skin near-body temperature (ST), respiratory frequency (RF), Activity (A) and Posture (P).

According to the scenario, the software/hardware and the communication line to allow data sending also varied (Table I). All the devices respected standards and requirements recommended in HRV guidelines on the acquisition and sampling of ECG [8].

III. TECHNICAL PLATFORM EVALUATION

The RHM platform was tested in three different scenarios: home care, medical ambulatory and hospital. Several methods have been proposed to analyze performance of remote patient tele-monitoring systems [26]. We focused initially on a metric of four critical technical factors: connectivity, usability, quality of data transmitted, interference with other devices. Connectivity was the capability of the telemedicine system to transmit data between client and server units without any disruptions. Usability accounted for both the ease of transmitting biosignals and data entry by users. Quality of data transmission referred to data integrity, and the forth parameter accounted for interference with other medical equipment.

<table>
<thead>
<tr>
<th>Devices, Software and Communication Lines Integrated and Tested</th>
<th>Description (Signals)</th>
<th>Software Hardware</th>
<th>Line</th>
</tr>
</thead>
<tbody>
<tr>
<td>CG-6106, (Cardguard) (ECG)</td>
<td>trans-telephonic ECGph</td>
<td>• IVR</td>
<td>PSTN/GPRS</td>
</tr>
<tr>
<td>Propaq Encore (Welch Allyn) Multi-channel monitor (ECG, SpO2, BP, T)</td>
<td>• client-sw</td>
<td>GPRS/DSL</td>
<td></td>
</tr>
<tr>
<td>Easy Ecg (Ates M.D.) (ECG)</td>
<td>ECGph Bluetooth</td>
<td>• client-sw</td>
<td>GPRS/DSL</td>
</tr>
<tr>
<td>BioHarness TM (Zephyr) Wearable chest monitor (HR, RF; A; T; P)</td>
<td>• client-sw</td>
<td>GPRS/DSL</td>
<td></td>
</tr>
<tr>
<td>Mocalab (Aditech Srl) Wearable arm monitor (ECG, SpO2, BP, T)</td>
<td>• client-sw</td>
<td>GPRS/DSL</td>
<td></td>
</tr>
</tbody>
</table>

IV. CLINICAL PROTOCOL AND EVALUATION

In previous studies, we defined and tested [27] a clinical protocol for management of elderly patient suffering from HF. This platform was developed according to the knowledge acquired, and its Electronic Patient Record was designed considering the frequency of control visits, the signs, the symptoms and the signals recommended in the guidelines on HF [8]. We also considered the dataset recommended in the guidelines on hypertension, which is often a concomitant pathology. Although the system collected several data, in the rest of this paper we focus on ECG records, as this is sufficient to describe the data-mining functions supported by the platform.

V. DATA-MINING

A. Preprocessing

The ECGs were processed following the international guidelines on HRV analyses [8]. After filtering, QRSs are detected using a standard algorithm [28]. Although this algorithm could be improved in future, we are first interested
in comparing our results with those obtained using other available tools, during clinical trial.

B. HRV Features extraction

We performed standard short-term HRV analysis, according to International Guidelines [8]. We developed the web services using the algorithms and the code of PhysioNet’s HRV Toolkit [28], since it is rigorously validated and because the tool will be used as a valuable benchmark during the clinical trial. This toolkit enables calculation of basic time- and frequency-domain HRV features widely used in the literature.

C. Patient classification

The platform supported a strategy of automatic classifications consisting of two steps: “HF detection” and “HF severity assessment”. The former, discussed in detail elsewhere [22], was used in the platform to pre-screen patients before they underwent the latter. The whole classification aimed to early detection of any worsening, assuming that during worsening, patients will gradually show characteristic of a more severe HF. Both classifiers were based on the CART methods. We pruned the trees according to a trade-off of misclassification probability and tree complexity, defined as its number of nodes. This reduced the risk of over-fitting as further detailed in Breiman [23]. The performances of both classifiers were assessed using a cross-validation technique [29]. Among all the trees achieving a satisfactory accuracy, we selected the one which minimized the divergence between training and testing performance [23].

1) HF detection

For the detection of HF, more excerpts of 5 minutes HRV were extracted throughout the same day. As shown in Figure 1, each excerpt was classified as normal or abnormal basing on three standard HRV features: Standard Deviation (SDNN), ratio between Low Frequencies and High Frequencies (LF/HF) and square Root of the Mean of the Squares of Differences between adjacent NN intervals (RMSSD). Finally, the subjects were considered as suffering from HF if more than \( \alpha = 30\% \) of the excerpts were classified as abnormal.

2) HF severity assessment

We labeled patients as “mild”, if classified by a cardiologist as NYHA I or II, or “severe” if classified NYHA III. We developed and trained several classification trees using all the possible combinations of short-term features. Figure 2 shows the best tree obtained, which used only three features: LF/HF; total power (TP) and RMSSD between adjacent normal beats. The patients were considered as suffering from severe HF if more than \( \alpha = 40\% \) of the excerpts were classified as severe.

As regards the platform evaluation, the tests will continue alongside the clinical trials and further metrics are under consideration. As regard to technical platform evaluation, no significant problems were reported in respect to the four parameters investigated. In terms of connectivity, we found a low rate of failure (<5% at home, <2% in hospital). About the quality of data transmitted, no significant amounts of data were lost. We experienced no interference with other devices. With regard to usability, no significant problems were reported, especially with the IVR, but it should be highlighted that younger relatives in many case supported the elderly. With regard to the effort of patients’ relatives, it has to be remarked that European Union identifies five different models among European countries and classifies Italy as “Mediterranean model”, in which the daily effort of family is very important and the assistance for the elderly is family-based [32]. For this reason, we assumed that family could have an important role in assisting patients and in using the most complex ICT components. In other NHS, social services should provide this support. Nonetheless, the most appropriate entry point of the system was the IVR. As regard to patients’ classification, we obtained the two classifiers described by performing a retrospective analysis on two public databases, including 83 subjects of which 54 were normal and 29 suffering from HF, among which 12 were mild and 17 severe. The data for the normal subjects were retrieved from the Normal Sinus Rhythm RR Interval Database [29]. The data for the CHF group were retrieved from the Congestive Heart Failure RR Interval Database [29]. Subjects were considered positive to the test if classified as “HF” in the first classifier and as “Severe” in the second one (Table II). The Figure 3 shows the ROC curves for both CARTs.

TABLE II

<table>
<thead>
<tr>
<th></th>
<th>T.P.</th>
<th>F.P.</th>
<th>T.N.</th>
<th>F.N.</th>
<th>ACC.</th>
<th>PRE.</th>
<th>SEN.</th>
<th>SPE.</th>
<th>AUC.</th>
</tr>
</thead>
<tbody>
<tr>
<td>HF vs NS</td>
<td>26</td>
<td>0</td>
<td>54</td>
<td>3</td>
<td>96.39</td>
<td>100.00</td>
<td>89.66</td>
<td>100.00</td>
<td>94.83</td>
</tr>
</tbody>
</table>
| SHF vs MHF | 14   | 3    | 9    | 3    | 79.31| 82.35 | 82.35| 75.00 | 78.68%

T.P.(N.): true positive (negative); F.N.(P): false positive(negative); ACC.: accuracy; PRE.: precision; SEN.: sensitivity; SPE.: specificity; AUC.: Area Under the Curve

Figure 3. ROC curves of the two CARTs showed how the sensitivity and the specificity vary according to α. The two red points show the thresholds chosen for the two classifiers (α=3 and α=4).

VII. DISCUSSION

In this paper, we presented a telemedicine platform with advanced functionalities of data mining for remote health monitoring of patients suffering from heart failure. The innovative contribution of this study is the integration of CART method into a telemedicine platform. This contribution is important because data mining represented the benefit of telemedicine compared to other DMPs.

Technical evaluation provided encouraging results, but tests on a greater number of elderly patients are still needed and further parameters should be considered.

The results of the two classifiers were satisfactory. The set of rules, reported in figure 1, is clinically consistent, even if the classifier did not use any a priori clinical knowledge. In fact, the leafs containing abnormal excerpts are on “left”, which reflects a depressed value of all the involved features. This is consistent with the results showed by Bigger [15], Musialik-Lydka [16] and Arbolishvili [17], who stated that standard HRV measures were significantly lower in HF patients than in normal subjects. Similarly, patients suffering from Severe HF, showed an even more depressed LF/HF and TP than those affected by mild HF. Comparisons with other papers had some limitations: difference in the lengths of ECG records (5min versus 24h) and in HRV features. In fact, on the same databases, the performance of our classifiers was higher than or comparable with the one of Asyali’s classifier [19], which were based on HRV long term measures. Moreover, we used all the records, even those rejected by Asyali. The performance of our classifier was lower than those of Isler’s classifier [20], which used HRV short-term measures, including wavelet entropy measures. This may be because of the discrimination power of wavelet entropy measures, which we did not use because they were not standard short-term measures and presumably too complex for most clinicians. In this regard, unlike other studies, we provided a set of rules, which are fully understandable by cardiologists. As regards HF severity assessment, we did not find a similar study. In any case, during the clinical trials, direct physician visits will provide further insights into the results of automatic classification. This study present some limits. Overall is the standardization of the ECG/HRV measurements that could greatly affect the measure (e.g. subject’s and electrodes’ position, the time of day). Moreover, the population of patients used in this preliminary experiment is pretty small. The first step to improve this research will be the enrolling of new patients and the use of data-mining on new significant signals, signs or symptoms. At this regard, we already tested more powerful methods of data-mining on the same data set. Nonetheless, our clinician partners are more confident with CART, as this method provides classifiers, which are full understandable.

VIII. CONCLUSIONS

The platform improved home monitoring by adding data mining functionalities. This was important in order to improve HM effectiveness and efficiency, especially benchmarking telemedicine to other DMPs, and not only to ambulatory follow-up. In this paper, we present preliminary results of classifiers for HF severity detection, which are innovative in comparison to the others previously published. These results are clinically consistent and confirm that patients suffering from HF present a depressed HRV. Similarly those patients suffering from severe HF present a more depressed HRV compared to those affected by mild HF. Compared to the other studies, we obtained higher precision and specificity values, but lower sensitivity. Moreover, our classifier is fully human-understandable. To enter into every-day clinical practices, this is a prerequisite of paramount importance for data-mining. Further data will help to improve classifiers’ performance and trials on patients enrolled on site will provide further insights due to the clinicians’ efforts.

IX. ACKNOWLEDGMENTS

Authors acknowledge Prof. B. Trimarco, Prof. N. De Luca and M.D. L. Argenziano of University Hospital of Federico II of Naples, Prof. F. Schiraldi and M.D. S. Verde of Hospital S. Paolo of Naples and M.D. M. Bartolo of Hospital San Giovanni Addolorata of Rome for their suggestions and contribution in designing the platform, protocols and processing according to clinical needs.

REFERENCES

Remote Health Monitoring of heart failure with data mining via CART method on HRV features.

IEEE Transaction on Biomedical Engineering, 58(3):800-804


Biomedical Engineering at the University Federico II of Naples, Italy.

He was involved in a number of academic and professional collaborations with both University and private companies. With the Department of Biomedical, Telecommunication and Electrical Engineering (DIBET) of University Federico II of Naples, he was scientific responsible of several national (FIRB INTESA, POR CAMPANIA RHM) and international (INTERREG INTERMED, INTERREG INTRAMEDNET and TEMPUS CRH-BME) research projects. His main fields of scientific and professional interest are: electronic and biological instrumentation; biosignal and data analysis; health care system management; health telematics and telemedicine. At present he is author and co-author of 350 publications.

Prof. Bracale is the General Secretary of the “Associazione Italiana di Ingegneria Medica e Biologica” (AIIMB), from 1986 to 1991 he was Chairman of the Developing Countries Committee and the National Secretaries Committee of International Federation of Medical and Biological Society (IFMBE). In 2002 was member of the IFMBE in the UN Economic and Social Council at the United Nation (Vienna).