Remote Health Monitoring of Heart Failure With Data Mining via CART Method on HRV Features

Leandro Pecchia, Paolo Melillo*, and Marcello Bracale

***Abstract*—Disease management programs, which use no advanced information and computer technology, 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 worsenings could require more complex and expensive care if not recognized. In this letter, we briefly describe the remote health monitoring platform we designed and realized, which supports heart failure (HF) severity assessment offering functions of data mining based on the classification and regression tree method. The system developed achieved accuracy and a precision of 96.39% and 100.00% in detecting HF and of 79.31% and 82.35% in distinguishing severe versus mild HF, respectively. 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*—Classification and regression tree (CART), data mining (DM), heart failure (HF), heart rate variability (HRV), home monitoring (HM).**

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 disease management programs (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 as DMPs, but less efficient because it is 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 incorporate 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, heart failure (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 noninvasive 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]–[17]. Many studies applied DM 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, classification and regression tree (CART). In this letter, 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], while other powerful methods of DM are not easy for humans to understand.

In this letter, 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,

Manuscript received July 13, 2010; revised October 5, 2010; accepted November 4, 2010. Date of publication November 15, 2010; date of current version February 18, 2011. This work was supported in part by the Regione Campania with the research project Remote Health Monitoring (R.H.M.) and in part by European Union with the research project TEMPUS Curricula Reformation and Harmonization in the field of Biomedical Engineering (CRH-BME). *Asterisk indicates corresponding author.*

L. Pecchia and M. Bracale are with the Department of Biomedical, Electronic, and Telecommunication Engineering, University Federico II of Naples, Naples 80125, Italy (e-mail: leandro.pecchia@unina.it; bracale@unina.it).

*P. Melillo is with the Department of Biomedical, Electronic and Telecommunication Engineering, University Federico II of Naples, Naples 80125, Italy (e-mail: paolo.melillo@unina.it).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TBME.2010.2092776

TABLE I
DEVICES, SOFTWARE AND COMMUNICATION LINES INTEGRATED AND TESTED

DEVICES	DESCRIPTION (SIGNALS)	SOFTWARE HARDWARE	LINE
CG-6106, (Cardguard)	trans-telephonic ECGph (ECG)	• IVR • telephone	PSTN/GPRS
Propaq Encore (Welch Allyn)	Multi-channel monitor (ECG, SpO2, BP, T)	• client-sw • palm/PC	GPRS/DSL
Easy Ecg (Ates M.D.)	ECGph Bluetooth (ECG)	• client-sw • palm/PC	GPRS/DSL
BioHarness TM (Zephyr)	Wearable chest monitor (HR; RF; A; T; P)	• client-sw • palm/PC	GPRS/DSL
Mocalab (Aditech Srl)	Wearable arm monitor (HR, ST, A, GSR)	• client-sw • palm/PC	GPRS/DSL

is at this moment undergoing randomized controlled trials involving real patients enrolled *ad hoc*. The aim of this letter was to describe briefly the platform, to present methods employed, and to present the preliminary results of the DM 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 an audio user interface. The WSA was used for raw data processing, signal analysis, and DM. 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 multiparametric monitors, recording ECG, blood pressure, heart rate, SpO2, temperature, galvanic skin response, skin near-body temperature, respiratory frequency, activity, and posture.

According to the scenario, the software/hardware and the communication line to allow data sending also varied (see 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 telemonitoring systems [26]. We focused initially on a metric of four critical technical factors: connectivity, usability, quality of data transmitted, and 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 fourth parameter accounted for interference with other medical equipment.

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 EPR 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, we focus here on ECG records, as this is sufficient to describe the DM functions supported by the platform.

V. DATA MINING

A. Preprocessing

The ECGs were processed following the international guidelines on HRV analyses [8]. After filtering, QRS complexes 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 prescreen 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 tradeoff of misclassification probability and tree complexity, defined as its number of nodes. This reduced the risk of overfitting as further detailed in [23]. The performances of both classifiers were assessed using a cross-validation technique [29].

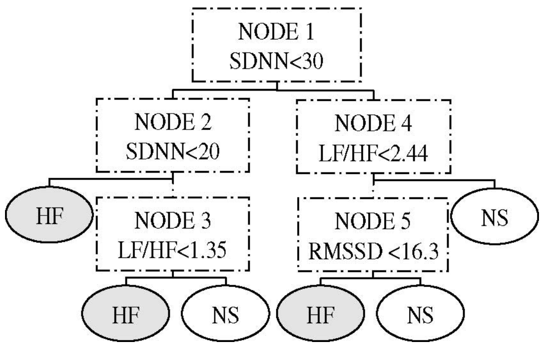

Fig. 1. CART for excerpts classification: patients suffering from heart failure (HF) versus normal subjects (NS).

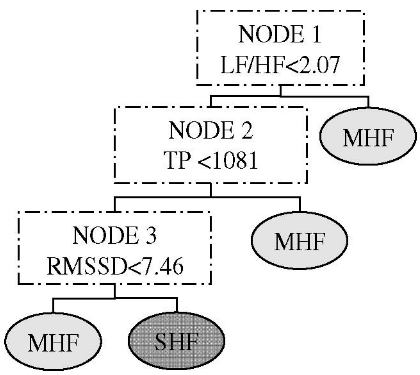

Fig. 2. CART for excerpts classification: severe heart failure (SHF) patients versus mild heart failure (MHF) patients.

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-min HRV were extracted throughout the same day. As shown in Fig. 1, each excerpt was classified as normal or abnormal basing on three standard HRV features: standard deviation, ratio between low frequencies and high frequencies (LF/HF), and square root of the mean of the sum 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. Fig. 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.

VI. Results

As regards the platform evaluation, the tests will continue alongside the clinical trials and further metrics are under consideration. As regards to technical platform evaluation, no significant problems were reported in respect to the four parameters

TABLE II
PERFORMANCE

	T.P.	F.P.	T.N.	F.N.	ACC.	PRE.	SEN.	SPE.	AUC
HF vs NS	26	0	54	3	96.39	100.00	89.66	100.00	94.83
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.

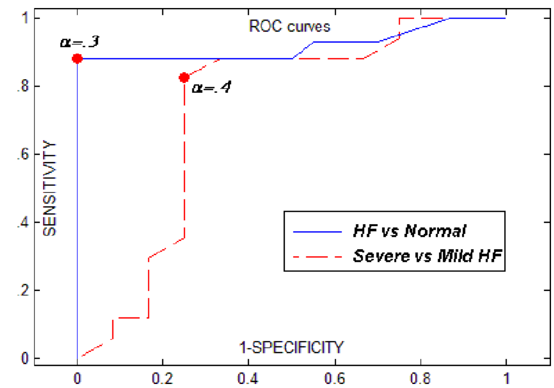

Fig. 3. ROC curves of the two CARTs showed how the sensitivity and the specificity vary according to α. The two red points show the thresholds chosen for the two classifiers ($\alpha = 0.3$ and $\alpha = 0.4$).

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 National Health Service, social services should provide this support. Nonetheless, the most appropriate entry point of the system was the IVR. As regards 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 congestive heart failure 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 (see Table II). Fig. 3 shows the receiver operating characteristic curves for both CARTs.

VII. Discussion

In this letter, we presented a telemedicine platform with advanced functionalities of DM for remote health monitoring of patients suffering from heart failure. The innovative contribution of this study is the integration of the CART method into a telemedicine platform. This contribution is important because DM 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 Fig. 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 (5 min versus 24 h) 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 presents 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 DM on new significant signals, signs, or symptoms. At this regard, we already tested more powerful methods of DM on the same dataset. Nonetheless, our clinician partners are more confident with CART, as this method provides classifiers, which are full understandable.

VIII. Conclusion

The platform improved HM by adding DM 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 letter, 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 everyday clinical practices, this is a prerequisite of paramount importance for DM. Further, data will help to improve classifiers’ performance and trials on patients enrolled on site will provide further insights due to the clinicians’ efforts.

Acknowledgment

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

References

[1] S. Koch, “Home telehealth—Current state and future trends,” *Int. J. Med. Inform.*, vol. 75, no. 8, pp. 565–576, 2006.

[2] C. S. Pattichis, C. N. Schizas, M. S. Pattichis, E. Micheli-Tzanakou, E. C. Kyriakou, and D. I. Fotiadis, “Introduction to the special section on computational intelligence in medical systems,” *IEEE Trans. Inform. Technol. Biomed.*, vol. 13, no. 5, pp. 667–672, Sep. 2009.

[3] S. G. Mougiakakou, I. K. Valavanis, N. A. Mouravliansky, A. Nikita, and K. S. Nikita, “DIAGNOSIS: A telematics-enabled system for medical image archiving, management, and diagnosis assistance,” *IEEE Trans. Instrum. Meas.*, vol. 58, no. 7, pp. 2113–2120, Jul. 2009.

[4] P. A. Bath, “Data mining in health and medical information,” *Annu. Rev. Inform. Sci. Technol.*, vol. 38, pp. 331–369, 2004.

[5] R. Gaikwad and J. Warren, “The role of home-based information and communications technology interventions in chronic disease management: A systematic literature review,” *Health Inform. J.*, vol. 15, no. 2, pp. 122–146, 2009.

[6] A. Martinez, E. Everss, J. L. Rojo-Alvarez, D. P. Figal, and A. Garcia-Alberola, “A systematic review of the literature on home monitoring for patients with heart failure,” *J. Telemed. Telecare*, vol. 12, no. 5, pp. 234–241, 2006.

[7] J. Gonseth, P. Guallar-Castillon, J. R. Banegas, and F. Rodriguez-Artalejo, “The effectiveness of disease management programmes in reducing hospital re-admission in older patients with heart failure: A systematic review and meta-analysis of published reports,” *Eur. Heart J.*, vol. 25, no. 18, pp. 1570–1595, 2004.

[8] M. Jessup, W. T. Abraham, D. E. Casey, A. M. Feldman, G. S. Francis, T. G. Ganiats, M. A. Konstam, D. M. Mancini, P. S. Rahko, M. A. Silver, L. W. Stevenson, C. W. Yancy, S. A. Hunt, M. H. Chin, H. F. W. Comm, and W. C. Members, “2009 focused update: ACCF/AHA guidelines for the diagnosis and management of heart failure in adults a report of the American College of Cardiology Foundation/American Heart Association task force on practice guidelines,” *Circulation*, vol. 119, no. 14, pp. 1977–2016, 2009.

[9] G. Pare, M. Jaana, and C. Sicotte, “Systematic review of home telemonitoring for chronic diseases: The evidence base,” *J. Amer. Med. Inform. Assoc.*, vol. 14, no. 3, pp. 269–277, 2007.

[10] L. Pecchia, U. Bracale, and M. Bracale, “Health technology assessment of home monitoring for the continuity of care of patient suffering from congestive heart failure,” in *Proc. Med. Phys. Biomed. Eng. World. Congr.*, Munich, Germany, 2009, pp. 184–187.

[11] L. Liao, J. G. Jollis, K. J. Anstrom, D. J. Whellan, D. W. Kitzman, G. P. Aurigemma, D. B. Mark, K. A. Schulman, and J. S. Gottdiener, “Costs for heart failure with normal vs reduced ejection fraction,” *Arch. Int. Med.*, vol. 166, no. 1, pp. 112–118, 2006.

[12] D. Lloyd-Jones, R. J. Adams, T. M. Brown, M. Carnethon, S. Dai, G. De Simone, T. B. Ferguson, E. Ford, K. Furie, C. Gillespie, A. Go, K. Greenlund, N. Haase, S. Hailpern, P. M. Ho, V. Howard, B. Kissela, S. Kittner, D. Lackland, L. Lisabeth, A. Marelli, M. M. McDermott, J. Meigs, D. Mozaffarian, M. Mussolino, G. Nichol, V. Roger, W. Rosamond, R. Sacco, P. Sorlie, R. Stafford, T. Thom, S. Wasserthiel-Smoller, N. D. Wong, and J. Wylie-Rosett, "Heart disease and stroke statistics—2010 Update. A report from the American Heart Association," *Circulation*, vol. 121, no. 12, pp. e1–e260, 2009.

[13] M. R. Cowie, A. Mosterd, D. A. Wood, J. W. Deckers, P. A. PooleWilson, G. C. Sutton, and D. E. Grobbee, "The epidemiology of heart failure," *Eur. Heart J.*, vol. 18, no. 2, pp. 208–225, 1997.

[14] M. Malik, J. T. Bigger, A. J. Camm, R. E. Kleiger, A. Malliani, A. J. Moss, and P. J. Schwartz, "Heart rate variability: Standards of measurement, physiological interpretation, and clinical use," *Eur. Heart J.*, vol. 17, no. 3, pp. 354–381, 1996.

[15] J. T. Bigger, J. L. Fleiss, R. C. Steinman, L. M. Rolnitzky, W. J. Schneider, and P. K. Stein, "RR Variability in healthy, middle-aged persons compared with patients with chronic coronary heart-disease or recent acute myocardial-infarction," *Circulation*, vol. 91, no. 7, pp. 1936–1943, 1995.

[16] Musialik-Lydka, B. Sredniawa, and S. Pasyk, "Heart rate variability in heart failure," *Kard. Pol*, vol. 58, no. 1, pp. 10–16, 2003.

[17] G. N. Arbolishvili, V. Y. Mareev, Y. A. Orlova, and Y. N. Belenkov, "Heart rate variability in chronic heart failure and its role in prognosis of the disease," *Kardiologiya*, vol. 46, no. 12, pp. 4–11, 2006.

[18] T. D. J. Smilde, D. J. Van Veldhuisen, and M. P. Van Den Berg, "Prognostic value of heart rate variability and ventricular arrhythmias during 13-year follow-up in patients with mild to moderate heart failure," *Clin. Res. Cardio.*, vol. 98, no. 4, pp. 233–239, 2009.

[19] M. H. Asyali, "Discrimination power of long-term heart rate variability measures," in *Proc. 25th Annu. Int. Conf. IEEE Eng. Med. Bio. Soc.*, Sep. 2003, pp. 200–203.

[20] Y. Isler and M. Kuntalp, "Combining classical HRV indices with wavelet entropy measures improves to performance in diagnosing congestive heart failure," *Comput. Bio. Med.*, vol. 37, no. 10, pp. 1502–1510, 2007.

[21] L. Pecchia, P. Melillo, M. Sansone, and M. Bracale, "Heart rate variability in healthy people compared with patients with congestive heart failure," in *Proc. 9th Int. Conf. Inform. Technol. Appl. Biomed.*, Larnaca, Cyprus, 2009, pp. 1–4.

[22] L. Pecchia, P. Melillo, M. Sansone, and M. Bracale, "Discrimination power of short-term heart rate variability measures for CHF assessment," *IEEE Trans. Inform. Technol. Biomed.*, to be published. DOI: 10.1109/TITB.2010.2091647.

[23] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*. Belmont, CA: Wadsworth, 1984.

[24] K. J. Cios and G. W. Moore, "Uniqueness of medical data mining," *Artif. Intell. Med.*, vol. 26, no. 1, pp. 1–24, 2002.

[25] P. Melillo, L. Pecchia, and M. Bracale, "Interactive voice response system for home telemonitoring of heart failure patients," in *Proc. Med. Phy. Biomed. Eng. World Congr.*, Munich, Germany, 2009, pp. 153–156.

[26] P. Pawar, B.-J. Beijnum van, M. Sinderen van, A. Aggarwal, and F. De Clercq, "Performance analysis of nomadic mobile services on multi-homed handheld devices," in *Proc. Int. Symp. Perform. Evalu. Comput. Telecommun. Syst.*, San Diego, CA, 2007, pp. 387–396.

[27] L. Pecchia, F. Schiraldi, S. Verde, E. Mirante, P. A. Bath, and M. Bracale, "Evaluation of short-term effectiveness of the disease management program "Di.Pro.Di" in continuity of care of patients suffering from congestive heart failure," *J. Amer. Geriat. Soc.*, vol. 58, no. 8, pp. 1603–1604, 2010.

[28] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[29] M. W. Browne, "Cross-validation methods," *J. Math. Psychol.*, vol. 44, no. 1, pp. 108–132, 2000.