



Toward a diagnostic CART model for Ischemic heart disease and idiopathic dilated cardiomyopathy based on heart rate total variability

Agostino Accardo¹ · Luca Restivo² · Miloš Ajčević¹ · Aleksandar Miladinović^{1,3} · Katerina Iscra¹ · Giulia Silveri¹ · Marco Merlo² · Gianfranco Sinagra²

Received: 30 November 2021 / Accepted: 15 June 2022 / Published online: 9 July 2022
© The Author(s) 2022

Abstract

Diagnosis of etiology in early-stage ischemic heart disease (IHD) and dilated cardiomyopathy (DCM) patients may be challenging. We aimed at investigating, by means of classification and regression tree (CART) modeling, the predictive power of heart rate variability (HRV) features together with clinical parameters to support the diagnosis in the early stage of IHD and DCM. The study included 263 IHD and 181 DCM patients, as well as 689 healthy subjects. A 24 h Holter monitoring was used and linear and non-linear HRV parameters were extracted considering both normal and ectopic beats (heart rate total variability signal). We used a CART algorithm to produce classification models based on HRV together with relevant clinical (age, sex, and left ventricular ejection fraction, LVEF) features. Among HRV parameters, MeanRR, SDNN, pNN50, LF, LF/HF, LFn, FD, Beta exp were selected by the CART algorithm and included in the produced models. The model based on pNN50, FD, sex, age, and LVEF features presented the highest accuracy (73.3%). The proposed approach based on HRV parameters, age, sex, and LVEF features highlighted the possibility to produce clinically interpretable models capable to differentiate IHD, DCM, and healthy subjects with accuracy which is clinically relevant in first steps of the IHD and DCM diagnostic process.

Keywords Computer-aided diagnosis · Heart rate variability · Ischemic heart disease · Dilated cardiomyopathy · Interpretable machine learning

1 Introduction

Ischemic heart disease (IHD), in its chronic stable form, is a subtle pathology due to its silent behavior before developing in unstable angina, myocardial infarction or, possibly, sudden cardiac death. This condition typically occurs when there is an imbalance between myocardial oxygen supply and demand, typically due to atherosclerotic heart disease.

The diagnosis in the early stage of the IHD is necessary to improve clinical outcomes, which can often be challenging. Clinical diagnosis relies on the patient's symptoms, especially chest pain, on the pathological ECG and on echocardiography, while only invasive coronary angiography, including the use of possibly toxic contrast means, can provide a definite diagnosis. At the same time, dilated cardiomyopathy (DCM) is a non-ischemic and non-valvular heart muscle disease frequently characterized by significant left ventricular (LV) or biventricular systolic dysfunction at the time of the diagnosis despite asymptomatic or scarcely symptomatic patients [1] reflecting a long period of asymptomatic silent disease progression [2]. Diagnosis of DCM, particularly in the early stages of the disease, can often be difficult and rely on advanced echocardiography (speckle tracking analysis), cardiac magnetic resonance imaging, including a comprehensive tissue characterization analysis, and genetic testing that often are not available or difficult to deliver to patients.

✉ Agostino Accardo
accardo@units.it

¹ Department Engineering and Architecture, University of Trieste, Trieste, Italy

² Cardiovascular Department, Azienda Sanitaria Universitaria Giuliano Isontina (ASUGI) and University of Trieste, Trieste, Italy

³ Institute for Maternal and Child Health – IRCCS Burlo Garofolo, Trieste, Italy

Therefore, novel biomarkers, preferably based on non-invasive techniques, are needed.

Heart disease-related pathophysiologic changes and subsequent alteration of heart rate variability (HRV) can provide important prognostic information [3]. In addition, HRV-based biomarkers have a potentially important role in risk stratification for individuals with suspected heart disease [4]. Nevertheless, the diagnostic role of HRV differentiation between IHD and DCM is still in the early research stage [3]. Indeed, the features extracted from ECG alone may not be able to discriminate these pathological conditions, but they might be complementary to other clinical and instrumental parameters.

Despite the growing use of machine learning-based prediction models in medicine [5–9], clinicians still struggle to rely on these models in clinical practice [10]. Machine learning methods were also applied to produce heart disease detection and prediction models [11] based on clinical history and ECG features [12], magnetocardiography [13], photoplethysmography signal parameters [14], and HRV and blood pressure variability features [15].

One of the problems of the complex machine learning methods (e.g., random forest [16], neural networks [17]) is that the published results are mostly focused on a classification/regression model performance metric, but rarely on practical usability for prediction in medicine [18, 19]. Here is a need of ensuring that machine learning models used in healthcare are interpretable [10, 20]. The classification and regression tree (CART) [21] is an approach which produce interpretable models not only providing output information about a certain disease but also help to intrinsically evaluate the plausibility of the model by examining the selected thresholds and branches in comparison to the existing knowledge [22]. The classification tree modeling, even though in practice it might represent slightly lower accuracy in comparison to the black-box models [10], provides better interpretability and practical usability in clinical application. In addition, their simple visualization allows clinicians to follow a set of rules and thresholds for selected clinical and instrumental features.

The aim of this study was to investigate, by means of CART modeling, the predictive power of HRV features together with non-invasive clinical parameters to support the diagnosis in the early stage of ischemic heart disease and dilated cardiomyopathy.

2 Methods

2.1 Study population

In this study, we analyzed clinical data and processed ECG signals of 1133 subjects who were consecutively enrolled

from December 2016 to October 2018 at the Cardiovascular Department of Trieste University Hospital (Trieste, Italy). In particular, the study encompassed 263 patients affected by IHD (207 males, aged 71 ± 10 , and 56 females, age 76 ± 10 years), 181 patients suffering from DCM (111 males, age 59 ± 12 years, and 70 females, age 63 ± 15 years), and 689 healthy controls (321 males, age 62 ± 15 y, and 368 females, age 64 ± 16 years). The assessment of ischemic heart disease (IHD) was carried out from clinical and laboratory findings, and it was systematically confirmed by coronary angiography [23]. The IHD patients did not present acute coronary syndrome in the 3 months before the Holter monitoring. The DCM patients were enrolled after clinical assessment based on coronary artery disease sufficient to explain the dysfunction or in presence of a left ventricular ejection fraction (LVEF) $< 50\%$ without evidence of pressure or volume overload [1]. Coronary angiography was systematically performed in patients older than 35 years, with cardiovascular risk factors and/or without familial history for DCM. Patients with known trigger factors, such as toxic insults from alcohol or drugs, and tachyarrhythmias were also excluded. Both groups of patients were on beta-blocker pharmacological treatment. The exclusion criteria for healthy controls (HC) were the presence of peripheral artery disease, thyroid disorders, history of myocardial revascularization, hypertensive heart disease, pulmonary hypertension, or severe valvulopathy. The study was conducted according to the principles of the Declaration of Helsinki and was approved by the Trieste Hospital Ethic Committee (Project Number: N.O.43/2009, prot.2161). All participants released their written informed consent.

2.2 Heart rate total variability acquisition and processing

All subjects underwent a 24 h Holter ECG recording using the ambulatory electrocardiographic recorder SpiderView (Sorin Group, Italy) with a sampling rate of 200 Hz. The RR intervals were extracted and labeled by using SyneScope analysis software (Sorin Group, Italy). The RR intervals were labeled as normal (N), premature ventricular contractions—ectopic beats (V), artifacts (A), and calibration (C). The RR interval records were cut into 5 min segments without overlap. Each RR 5-min segment was included in the analysis only if the longest ectopic beats subsequence (labeled with V) or the longest artifact subsequence (labeled with A) does not exceed 10 s. The RR marked with a calibration label was ignored. These segments were interpolated with cubic spline and resampled at 2 Hz, producing the heart rate total variability (HRV) signal. Subsequently, in each segment, linear and non-linear HRV parameters were extracted. In particular, the linear parameters MeanRR, SDNN, RMSSD, NN50, and pNN50 evaluating the RR

variability were calculated directly from RR sequence [24], while in the frequency domain, the absolute powers in low (LF = 0.04–0.15 Hz) and high (HF = 0.15–0.40 Hz) frequency bands, related to the vagal and sympathetic nerve control on the heart rhythm, were estimated from the interpolated HRV signal. Moreover, the normalized low and high-frequency powers (LFn, HFn) and their ratio (LF/HF) were calculated from the latter parameters. The non-linear analysis was carried out by calculating Poincaré plot parameters (SD1, SD2, SD1/SD2), reflecting the short- and the long-term variability [25], estimating the fractal dimension (FD) [26] and the power-law beta exponent [27], quantifying the complexity of the system generating the signal. Finally, the median values of all parameters during 24 h were calculated and used as the input feature vector of the classifier.

2.3 Classification method, features selection, and performance measurements

The CART method [21], used for diagnostic modeling because of its easy interpretability in the clinical domain, was employed to produce models capable of differentiating between three groups (IHD, DCM, and HC). The models were at first produced considering HRV features (Table 1) together with subjects’ age and sex. Three different models were considered: in the first set, all the 17 features were taken in consideration (Model_{All_features}), in the second and third model, only the selected features were used as input of the CART. Stepwise regression algorithm, selecting only the most significant explanatory variables, and correlation analysis (excluding those variables presenting a regression coefficient less than 0.8) were used to operate the selection, producing Model_{Stw} and Model_{Corr}, respectively. Stepwise regression, which was applied, is the step-by-step iterative

algorithm for the selection of independent variables by adding or removing potential explanatory variables in succession and testing for statistical significance after each iteration.

Subsequently, another non-invasive parameter, the left ventricular volume ejection fraction (LVEF) obtained by the Simpson biplane method [28] useful to discriminate some heart pathologies, was also included. This parameter requires the execution of a further ecographic examination, which is performed in many cases and was added to understand if it was necessary to make it routine. Thus, further three different CART models were produced including all the features (Model_{All_features+LVEF}) or, as previously, only those selected by stepwise (Model_{Stw+LVEF}) or low correlation (Model_{Corr+LVEF}).

The CART uses an algorithm to construct the decision tree by essentially producing a set of rules represented by decisional nodes, branches, and leaves (i.e., terminal nodes) which are assigned to a class. The algorithm is based on a recursive segmentation (each non-leaf node has only two branches) and the generated decision tree is a simple structure in which each decision step can be divided into yes-or-no questions about each feature. The two steps of the CART are binary recursive partitioning to construct the complex binary tree and then prune it back to find an optimal subtree. In this work, the Gini coefficient, representing a variance estimate based on all comparisons of possible pairs of values in a subgroup, has been used as a loss function. Cross-validation was used as a technique to avoid overfitting and to produce a model that generalizes better to unseen data. The classification on the dataset was estimated using tenfold cross-validation. The process was then repeated 10 times, using each of the subsamples only once as the validation data. Therefore, the

Table 1 The set of linear and non-linear HRV features

| HRV parameter | Definition |
|--------------------------------|--|
| MeanRR (ms) | Mean of RR intervals |
| SDNN (ms) | Standard deviation of RR intervals |
| RMSSD (ms) | Root mean square of the squared differences of successive RR intervals |
| NN50 | Number of differences of successive RR intervals greater than 50 ms |
| pNN50 | Proportion of NN50 divided by the total number of RR intervals |
| LF (ms ²) | Low-frequency power (from 0.04 to 0.15 Hz) |
| HF (ms ²) | High-frequency power (from 0.15 to 0.40 Hz) |
| LF/HF | Low-frequency power/high-frequency power |
| LFn | Low-frequency power/total power |
| HFn | High-frequency power/total power |
| Beta exp (ms ² /Hz) | Beta exponent |
| SD1 (ms) | Short-term variability of the RR sequence—from Poincaré Plot |
| SD2 (ms) | Long-term variability of the RR sequence—from Poincaré Plot |
| SD1/SD2 | Short-term variability/long-term variability of the RR sequence |
| FD | Fractal dimension |

overall cross-validation accuracy was calculated as a mean of all 10 validation folds. To evaluate the trade-off between model interpretability and classification performance of produced decision tree, the obtained classification accuracy was compared with classification accuracy of other selected machine learning approaches: Logistic regression, Naïve Bayes, and support vector machine (SVM). All analyses were performed and implemented in MATLAB using the Statistics and Machine Learning toolbox.

3 Results

The demographic and clinical characteristics of subjects included in the three considered groups, as well as, the mean values of linear and non-linear HRV parameters are reported in Table 2.

When LVEF was not considered, the features selected by using the stepwise regression method were MeanRR, SDNN, LFn, FD, sex, and age (Model_{Stw}), while the outcome of correlation analysis yielded in the selection of MeanRR, SDNN, LF, LF/HF, Beta exp, sex, and age (Model_{Corr}). On the other hand, when LVEF was added to the other parameters, the features selected by using the stepwise regression method were MeanRR, pNN50, LF/HF, FD, sex, age, and LVEF (Model_{Stw+LVEF}) and those identified by the correlation analysis were MeanRR, SDNN, LF, LF/HF, Beta exp, sex, age, and LVEF (Model_{Corr+LVEF}). The selected features were used as input vector to produce six different decision tree models, as described in the methods section.

Table 3 reports the features selected by CART approach (in bold) together with the performance metrics of produced models. In general, models which included LVEF showed a higher accuracy (about 10%) compared to those based only on HRV and demographic parameters. The highest accuracy on the test set was observed for the CART Model_{Stw+LVEF} (Fig. 1) while among the models based only on HRV and demographic parameters without LVEF the best classification performance was observed for the CART Model_{Corr} (Fig. 2). In addition, in Table 4 are reported AUC values for each group and model.

The comparison of classification accuracies obtained by different machine learning approaches is reported in Table 5. The classification accuracies obtained with models produced with Logistic regression, Naïve Bayes, and support vector machine (SVM) methods were similar to those obtained with CART.

Table 2 Mean and standard deviation values of the features sets

| | IHD | DCM | HC |
|--------------------------------|-------------|-------------|-------------|
| Age | 72 ± 11 | 61 ± 13 | 63 ± 15 |
| Sex (M/F) | 207/56 | 111/70 | 321/378 |
| LVEF (%) | 53 ± 13 | 44 ± 12 | 59 ± 6 |
| HRV | | | |
| MeanRR (ms) | 942 ± 145 | 880 ± 130 | 877 ± 138 |
| SDNN (ms) | 87 ± 68 | 85 ± 61 | 71 ± 53 |
| RMSSD (ms) | 71 ± 115 | 61 ± 111 | 38 ± 91 |
| NN50 | 70 ± 90 | 65 ± 85 | 50 ± 73 |
| pNN50 | 0.21 ± 0.27 | 0.19 ± 0.23 | 0.15 ± 0.21 |
| LF (ms ²) | 350 ± 2400 | 521 ± 1500 | 460 ± 2200 |
| HF (ms ²) | 640 ± 8100 | 626 ± 6005 | 276 ± 5800 |
| LF/HF | 0.97 ± 1.05 | 1.22 ± 1.30 | 2.06 ± 1.90 |
| LFn | 0.40 ± 0.20 | 0.44 ± 0.21 | 0.59 ± 0.22 |
| HFn | 0.60 ± 0.20 | 0.58 ± 0.21 | 0.44 ± 0.22 |
| Beta exp (ms ² /Hz) | 0.67 ± 0.55 | 0.75 ± 0.58 | 1.06 ± 0.57 |
| SD1 (ms) | 46.4 ± 42.4 | 45.5 ± 40.9 | 34.8 ± 34.1 |
| SD2 (ms) | 95.5 ± 70.1 | 92.8 ± 57.3 | 83.1 ± 54.2 |
| SD1/SD2 | 0.44 ± 0.15 | 0.44 ± 0.16 | 0.37 ± 0.14 |
| FD | 1.63 ± 0.15 | 1.63 ± 0.16 | 1.53 ± 0.16 |

Table 3 Feature sets used as an input vector to produce the six models (in bold the features selected by CART algorithm) and their classification performance measures

| | Model Features | CA | F1 | Precision | Recall |
|------------------------------------|---|-------|-------|-----------|--------|
| Model _{Stw} | MeanRR, SDNN, LFn, FD, sex, age | 60.2% | 58.6% | 57.7% | 60.1% |
| Model _{Corr} | MeanRR, SDNN, LF, LF/HF, Beta exp, sex, age | 61.4% | 59.1% | 58.1% | 61.4% |
| Model _{All_features} | MeanRR, SDNN, RMSSD, NN50, pNN50, LF, HF, LF/HF, LFn, HFn, Beta exp, SD1, SD2, SD1/SD2, FD, sex, age | 60.3% | 58.2% | 57.2% | 60.3% |
| Model _{Stw+LVEF} | MeanRR, pNN50, LF/HF, FD, sex, age, LVEF | 73.3% | 71.3% | 70.8% | 72.9% |
| Model _{Corr+LVEF} | MeanRR, SDNN, LF, LF/HF, Beta exp, sex, age, LVEF | 72.8% | 71.3% | 70.8% | 72.7% |
| Model _{All_features+LVEF} | MeanRR, SDNN, RMSSD, NN50, pNN50, LF, HF, LF/HF, LFn, HFn, Beta exp, SD1, SD2, SD1/SD2, FD, sex, age, LVEF | 72.6% | 70.8% | 70.3% | 72.6% |

Fig. 1 Decision tree model based on pNN50, FD, sex, age, and LVEF features. HC, healthy control; DCM, dilated cardiomyopathy; IHD, ischemic heart disease

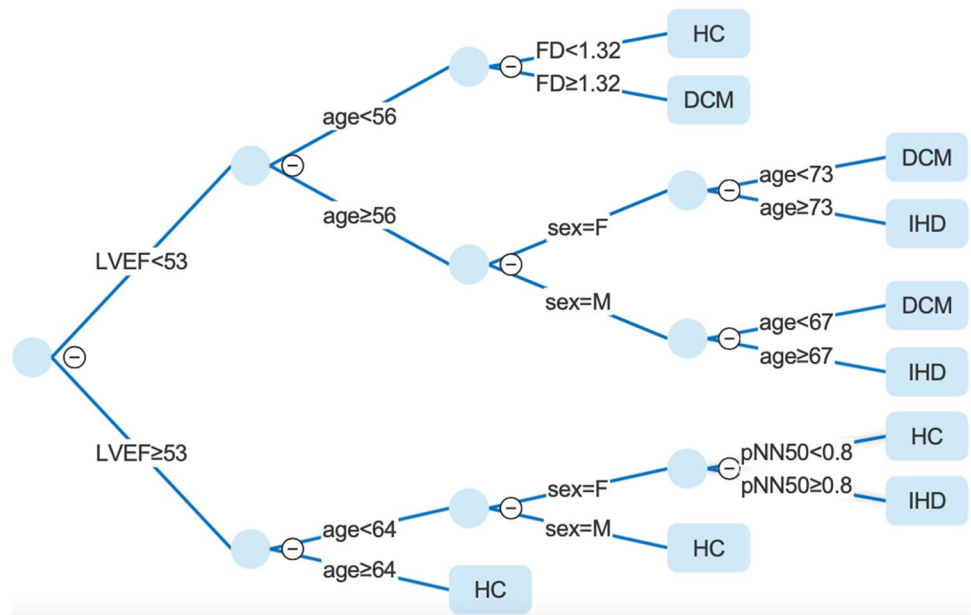


Fig. 2 Decision tree model based on MeanRR, SDNN, LF, LF/HF, Beta exp, sex, age features. HC, healthy control; DCM, dilated cardiomyopathy; IHD, ischemic heart disease

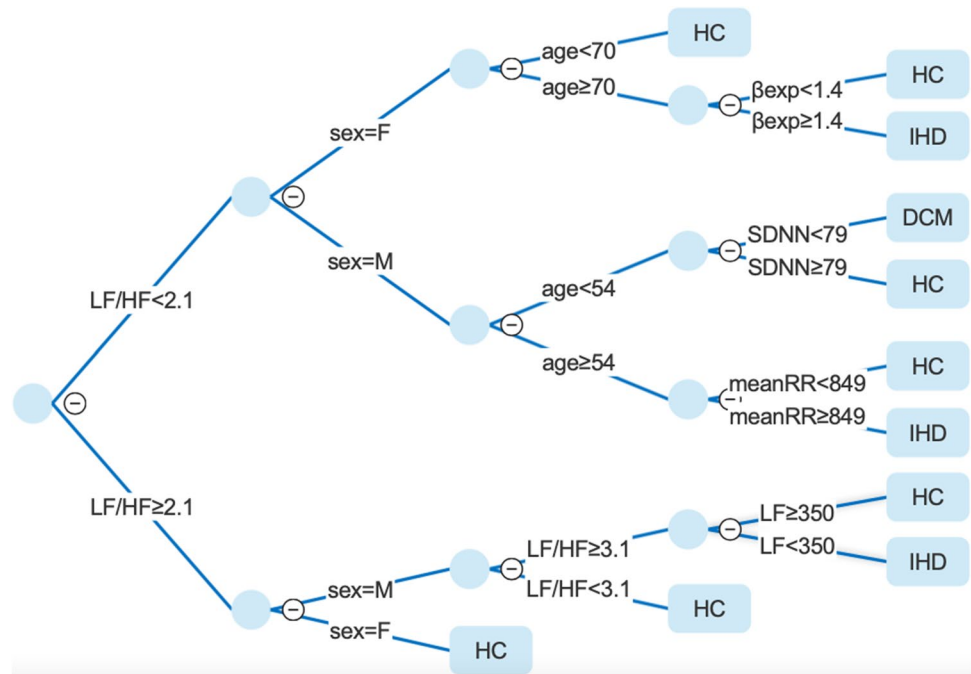


Table 4 AUC values for each group and model

| | AUC | | |
|------------------------------------|-------|-------|-------|
| | HC | DCM | IHD |
| Model _{Stw} | 69.5% | 62.7% | 75.3% |
| Model _{Corr} | 70.8% | 65.0% | 73.8% |
| Model _{All_features} | 69.9% | 63.0% | 72.5% |
| Model _{Stw+LVEF} | 84.1% | 83.0% | 73.6% |
| Model _{Corr+LVEF} | 83.2% | 83.5% | 74.9% |
| Model _{All_features+LVEF} | 83.6% | 84.3% | 74.9% |

Table 5 Classification accuracies obtained by different machine learning algorithms

| | CART | Logistic regression | Naive Bayes | SVM |
|-------------------------------------|-------|---------------------|-------------|-------|
| Model _{Stw} | 60.2% | 64.7% | 63.3% | 54% |
| Model _{Stw + LVEF} | 73.3% | 72.1% | 74.2% | 60.3% |
| Model _{Corr} | 61.4% | 64% | 61.6% | 49.8% |
| Model _{Corr + LVEF} | 72.8% | 72.9% | 71.7% | 63.7% |
| Model _{AllFeatures} | 60.3% | 65.4% | 58% | 50.4% |
| Model _{AllFeatures + LVEF} | 72.6% | 71.9% | 62% | 63.7% |

4 Discussion

Diagnosis of etiology in early-stage IHD and DCM patients may be challenging. IHD patients are generally asymptomatic or exhibit no typical signs and symptoms until the disease manifests as angina, myocardial infarction, or sudden cardiac death. Similarly, DCM represents a particular etiology of heart failure with reduced ejection fraction, frequently carrying a genetic background, which usually affects young patients with few co-morbidities, remaining asymptomatic for a long time. For this reason, there is research interest in the identification of novel, preferably non-invasive, biomarkers.

The main finding of this study is that the models including parameters extracted from the heart rate total variability signal are capable to differentiate three groups with accuracy, which is clinically relevant in first steps of the IHD and DCM diagnostic process. These findings support the hypothesis that HRV analysis emerges as an important, accessible, reproducible, supplementary tool for the IHD and DCM diagnosis.

Nowadays, different mathematical approaches for decision support systems have been proposed for the automatic classification of heartbeats and machine learning techniques have become a useful research diagnostic tool for physicians in the analysis of cardiovascular disease [29–32]. A recent study used artificial neural networks considering age, sex, and HRV features to classify ischemic heart disease patients and healthy subjects with an accuracy of 71.8% [29]. Moreover, considering also the left ventricular ejection fraction as a feature, even higher classification accuracy was obtained. Other authors identified only RR segments of DCM patients using HRV parameters as input of complex classifiers in which the CART was combined with other machine learning techniques [31, 32]. In particular, Thirugnman et al. used different machine learning techniques applied to HRV parameters for identifying DCM and healthy ECG segments with an accuracy of 99.93% considering 22 ECG samples, 16 belonging to DCM and 6 to healthy subjects [32]. Moreover, Mahesh et al. used classification trees with logistic regression on a combination of linear and non-linear HRV parameters to identify ECG segments of DCM subjects with an accuracy of 95.61% (13 cardiomyopathy RR segments). Both studies took the data from a diagnostic ECG database [31]. Only Dua et al. used the classification and regression tree analysis to distinguish IHD patients from healthy subjects [30]. They analyzed the HRV signals of 20 subjects obtaining an accuracy of 81.1% applying principal component analysis to non-linear HRV parameters.

In our study, we aimed to discriminate not so much a specific pathologic or normal ECG segment belonging

to subjects affected or not by cardiac pathology as rather to differentiate subjects belonging to the three groups (IHD, DCM, and HC) mainly by exploiting non-invasive HRV features. Among HRV parameters, MeanRR, SDNN, pNN50, LF, LF/HF, LFn, FD, and Beta exp were identified as the most informative. The produced interpretable decision tree models based only on HRV and demographic features, as well as including also LVEF, showed similar classification accuracies to those produced with logistic regression, Naïve Bayes, and support vector machine (SVM) methods. Even though in general the decision tree modeling might present a slightly lower accuracy in comparison to other commonly applied methods, we used CART algorithm because of its better interpretability and practical usability in clinical application. Models based only on HRV features and demographic parameters presented an accuracy of about 62%. From a clinical perspective, even the results obtained without LVEF (61.4%) is relevant, especially in the early differential diagnosis phase, as it allows to avoid invasive coronary angiography in selected patients. In the CART model without LVEF Model_{Corr}, the most important feature is the LF/HF, which reflects sympatho-vagal balance that can be altered in the patients affected by cardiomyopathies [33]. In fact, it can be observed a denser grouping of the pathologies, and exclusive existence of DCM, in the upper subtree. We also observed that in the bottom subtree the final decision of IHD classification is based again on LF/HF and LF parameters, which confirm their high discriminatory power. The second most important feature observed in the CART Model_{Corr} is sex, which in the upper subtree is strictly related to patient age. It has been already reported that DCM affects men more commonly than women [34], which is also observable in our CART Model_{Corr}. Regarding the age, the identified threshold of 54 years for men, is in line with previous findings, as most of DCM male patients become symptomatic between 20 and 60 years [35]. Finally, the SDNN is an additional parameter that allows fine classification between DCM and HC, as it reflects all the long-term HRV components and it is sensitive to low frequencies heart rate alteration present in DCM. Regarding IHD, in the CART Model_{Corr} was observed that it is more likely to be affected by the disease if the subject is older independently of sex. Furthermore, for beta exponent, a non-linear parameter related to the complexity of the signal generators is more likely to be altered in the patients affected with IHD [36], which can be also observed in produced decision tree.

The inclusion of LVEF parameter beside the HRV ones made it possible to improve the performance by about 10%. In particular, we observed the highest accuracy (73.3%) for the model based on pNN50, FD, sex, age, and LVEF features (Fig. 1). Concerning CART Model_{Stw+LVEF}, the most

important feature is LVEF. Our tree confirms that the cut-off for the diagnosis of DCM is around 50%. Indeed, in the subtree where the LVEF is higher than 53%, there are no branches to DCM identification. Moreover, we also observed the high relation of LVEF with patient age. In particular, the CART Model_{Stw+LVEF} showed that if the subject has high LVEF and it is older than 64 years old, it is more probable that it belongs to HC group. In the cases of LVEF below the 54% cut-off, the sex differences plays important role. Indeed, in our model it can be observed that females and males have different age thresholds to be classified as DCM or IHD (male age < 67, female age < 73).

LVEF showed that the model mainly based on HRV parameters classifies better IHD subjects than DCM and vice versa the model which takes into account also LVEF classifies better DCM subjects than IHD. This fact clearly indicates the contribution of the LVEF parameter, as discriminatory feature to identify the DCM but confounding to discriminate IHD patients. In order to further improve the classification performance, future studies could also take into account the circadian rhythm related physiological parameters variation [37–39].

5 Conclusions

In conclusion, the proposed approach based on HRV parameters, age, sex, and LVEF features highlighted the possibility to produce clinically interpretable models capable to differentiate IHD, DCM, and healthy subjects with accuracy which is clinically relevant in first steps of the IHD and DCM diagnostic process. These results support the hypothesis that HRV analysis emerges as an important, accessible, reproducible, complementary tool for the IHD and DCM diagnosis, potentially avoiding invasive and toxic exams especially in healthy subject cases.

Funding Open access funding provided by Università degli Studi di Trieste within the CRUI-CARE Agreement.

Declarations

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will

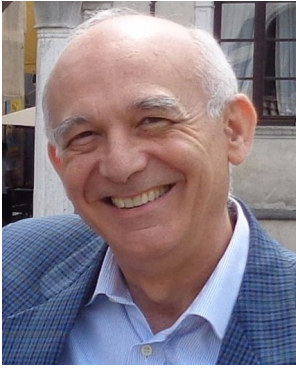
need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Elliott P, Andersson B, Arbustini E, Bilinska Z, Cecchi F, Charron P, Dubourg O, Kühl U, Maisch B, McKenna WJ, Monserrat L, Pankuweit S, Rapezzi C, Seferovic P, Tavazzi L, Keren A (2008) Classification of the cardiomyopathies: a position statement from the European Society Of Cardiology Working Group on Myocardial and Pericardial Diseases. *Eur Heart J* 29:270–276. <https://doi.org/10.1093/eurheartj/ehm342>
- Merlo M, Cannatà A, Gobbo M, Stolfo D, Elliott PM, Sinagra G (2018) Evolving concepts in dilated cardiomyopathy. *Eur J Heart Fail* 20:228–239. <https://doi.org/10.1002/ejhf.1103>
- Kleiger RE, Miller JP, Bigger JT, Moss AJ (1987) Decreased heart rate variability and its association with increased mortality after acute myocardial infarction. *Am J Cardiol* 59:256–262. [https://doi.org/10.1016/0002-9149\(87\)90795-8](https://doi.org/10.1016/0002-9149(87)90795-8)
- Shah AS, Lampert R, Goldberg J, Bremner JD, Li L, Thames MD, Vaccarino V, Shah AJ (2020) Alterations in heart rate variability are associated with abnormal myocardial perfusion. *Int J Cardiol* 305:99–105. <https://doi.org/10.1016/j.ijcard.2020.01.069>
- Darcy AM, Louie AK, Roberts LW (2016) Machine learning and the profession of medicine. *JAMA* 315:551–552. <https://doi.org/10.1001/jama.2015.18421>
- Deo RC (2015) Machine learning in medicine. *Circulation* 132:1920–1930. <https://doi.org/10.1161/CIRCULATIONAHA.115.001593>
- Kononenko I (2001) Machine learning for medical diagnosis: history, state of the art and perspective. *Artif Intell Med* 23:89–109. [https://doi.org/10.1016/s0933-3657\(01\)00077-x](https://doi.org/10.1016/s0933-3657(01)00077-x)
- Miladinović A, Ajčević M, Jarmolowska J, Marusic U, Colussi M, Silveri G, Battaglini PP, Accardo A (2021) Effect of power feature covariance shift on BCI spatial-filtering techniques: a comparative study. *Comput Methods Programs Biomed* 198:105808. <https://doi.org/10.1016/j.cmpb.2020.105808>
- Obermeyer Z, Emanuel EJ (2016) Predicting the future - big data, machine learning, and clinical medicine. *N Engl J Med* 375:1216–1219. <https://doi.org/10.1056/NEJMp1606181>
- Elshawi R, Al-Mallah MH, Sakr S (2019) On the interpretability of machine learning-based model for predicting hypertension. *BMC Med Inform Decis Mak* 19:146. <https://doi.org/10.1186/s12911-019-0874-0>
- Nagavelli U, Samanta D, Chakraborty P (2022) Machine learning technology-based heart disease detection models. *J Healthc Eng* 2022:7351061. <https://doi.org/10.1155/2022/7351061>
- Alsaffar M, Alshammari A, Alshammari G, Aljaloud S, Almurayziq TS, Abdoon FM, Abebaw S (2021) Machine learning for ischemic heart disease diagnosis aided by evolutionary computing. *Appl Bionics Biomech* 2021:6718029. <https://doi.org/10.1155/2021/6718029>
- Tao R, Zhang S, Huang X, Tao M, Ma J, Ma S, Zhang C, Zhang T, Tang F, Jianping Lu, Shen C, Xie X (2019) Magnetocardiography-based ischemic heart disease detection and localization using machine learning methods. *IEEE Trans Biomed Eng* 66:1658–1667. <https://doi.org/10.1109/TBME.2018.2877649>
- Pal P, Ghosh S, Chattopadhyay BP, Kumar Saha K, Mahadevappa M (2020) Screening of ischemic heart disease based on PPG signals using machine learning techniques. *Annu Int Conf IEEE Eng Med Biol Soc* 2020:5980–5983. <https://doi.org/10.1109/EMBC44109.2020.9176447>
- Rodriguez J, Schulz S, Giraldo BF, Voss A (2019) Risk stratification in idiopathic dilated cardiomyopathy patients using

- cardiovascular coupling analysis. *Front Physiol* 10:841. <https://doi.org/10.3389/fphys.2019.00841>
16. Liaw A, Wiener M (2002) Classification and Regression by randomForest. *R News* 2(3):18–22
 17. Specht DF (1990) Probabilistic neural networks. *Neural Netw* 3:109–118. [https://doi.org/10.1016/0893-6080\(90\)90049-Q](https://doi.org/10.1016/0893-6080(90)90049-Q)
 18. Basu Roy S, Teredesai A, Zolfaghari K, Liu R, Hazel D, Newman S, Martinez A (2015) Dynamic hierarchical classification for patient risk-of-readmission. In: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Association for Computing Machinery, New York, NY, USA, pp 1691–1700. <https://dl.acm.org/doi/10.1145/2783258.2788585>
 19. Futoma J, Morris J, Lucas J (2015) A comparison of models for predicting early hospital readmissions. *J Biomed Inform* 56:229–238. <https://doi.org/10.1016/j.jbi.2015.05.016>
 20. Stiglic G, Kocbek P, Fijacko N, Zitnik M, Verbert K, Cilar L (2020) Interpretability of machine learning-based prediction models in healthcare. *WIREs Data Min Knowl Discovery* 10:e1379. <https://doi.org/10.1002/widm.1379>
 21. Breiman L, Friedman JH, Olshen RA, Stone CJ (1984) Classification and regression trees. Routledge, Boca Raton
 22. Ahmad MA, Eckert C, Teredesai A (2018) Interpretable machine learning in healthcare. In: Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics. Association for Computing Machinery, New York, NY, USA, pp 559–560. <https://doi.org/10.1145/3233547.3233667>
 23. Knuuti J, Wijns W, Saraste A, Capodanno D, Barbato E, Funck-Brentano C, Prescott E, Storey RF, Deaton C, Cuisset T, Agewall S, Dickstein K, Edvardsen T, Escaned J, Gersh BJ, Svtil P, Gilard M, Hasdai D, Hatala R, Mahfoud F, Masip J, Muneretto C, Valgimigli M, Achenbach S, Bax JJ, ESC Scientific Document Group (2020) 2019 ESC Guidelines for the diagnosis and management of chronic coronary syndromes. *Eur Heart J* 41:407–477. <https://doi.org/10.1093/eurheartj/ehz425>
 24. Heart rate variability: standards of measurement, physiological interpretation and clinical use Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology (1996) *Circulation* 93:1043–1065. <https://doi.org/10.1161/01.CIR.93.5.1043>
 25. Woo MA, Stevenson WG, Moser DK, Trelease RB, Harper RM (1992) Patterns of beat-to-beat heart rate variability in advanced heart failure. *Am Heart J* 123:704–710. [https://doi.org/10.1016/0002-8703\(92\)90510-3](https://doi.org/10.1016/0002-8703(92)90510-3)
 26. Higuchi T (1988) Approach to an irregular time series on the basis of the fractal theory. *Physica D*. [https://doi.org/10.1016/0167-2789\(88\)90081-4](https://doi.org/10.1016/0167-2789(88)90081-4)
 27. Cusenza M, Accardo A, D’Addio G, Corbi G (2010) Relationship between fractal dimension and power-law exponent of heart rate variability in normal and heart failure subjects. *Computing in Cardiology*, 2010, pp 935–938
 28. Lang RM, Badano LP, Mor-Avi V, Afilalo J, Armstrong A, Ernande L, Flachskampf FA, Foster E, Goldstein SA, Kuznetsova T, Lancellotti P, Muraru D, Picard MH, Rietzschel ER, Rudski L, Spencer KT, Tsang W, Voigt J-U (2015) Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *J Am Soc Echocardiogr* 28:1–39.e14. <https://doi.org/10.1016/j.echo.2014.10.003>
 29. Accardo A, Silveri G, Merlo M, Restivo L, Ajčević M, Sinagra G (2020) Detection of subjects with ischemic heart disease by using machine learning technique based on heart rate total variability parameters. *Physiol Meas* 41:11. <https://doi.org/10.1088/1361-6579/abc321>
 30. Dua S, Du X, Sree SV, Ahamed VIT (2012) Novel classification of coronary artery disease using heart rate variability analysis. *J Mech Med Biol* 12(04):1240017. <https://doi.org/10.1142/S0219519412400179>
 31. Mahesh V, Kandaswamy A, Vimal C, Sathish B (2009) Cardiac disease classification using heart rate signals. *Int J Electron Healthc* 5:211–230. <https://doi.org/10.1504/ijeh.2010.034173>
 32. Thirugnanam M, Santhoshi PM (2020) Hybrid feature extraction and stacking based ensemble classifier model for cardiomyopathy classification. *Int J Adv Sci Technol* 29:1396–1413
 33. Tanabe T, Iwamoto T, Fusegawa Y, Yoshioka K, Shiina Y (1995) Alterations of sympathovagal balance in patients with hypertrophic and dilated cardiomyopathies assessed by spectral analysis of RR interval variability. *Eur Heart J* 16:799–807. <https://doi.org/10.1093/oxfordjournals.eurheartj.a060999>
 34. Fairweather D, Cooper LT, Blauwet LA (2013) Sex and gender differences in myocarditis and dilated cardiomyopathy. *Curr Probl Cardiol* 38:7–46. <https://doi.org/10.1016/j.cpcardi.2012.07.003>
 35. Reichart D, Magnussen C, Zeller T, Blankenberg S (2019) Dilated cardiomyopathy: from epidemiologic to genetic phenotypes. *J Intern Med* 286:362–372. <https://doi.org/10.1111/joim.12944>
 36. Krstacic G, Krstacic A, Martinis M, Vargovic E, Knezevic A, Smalcelj A, Jembrek-Gostovic M, Gamberger D, Smuc T (2002) Non-linear analysis of heart rate variability in patients with coronary heart disease. In: *Computers in Cardiology*, pp 673–675
 37. Accardo A, Merlo M, Silveri G, Del Popolo L, Dalla Libera L, Restivo L, Cinquetti M, Cannatà A, Sinagra G (2021) Influence of ageing on circadian rhythm of heart rate variability in healthy subjects. *J Cardiovasc Med (Hagerstown)* 22:405–413. <https://doi.org/10.2459/JCM.0000000000001048>
 38. Accardo A, Silveri G, Ajčević M, Miladinović A, Pascasio L (2021) Influence of smoking and other cardiovascular risk factors on heart rate circadian rhythm in normotensive and hypertensive subjects. *PLoS ONE* 16:e0257660. <https://doi.org/10.1371/journal.pone.0257660>
 39. Miladinović A, Ajčević M, Silveri G, Liguori L, Pascasio L, Accardo A (2021) Ambulatory blood pressure monitoring versus office blood pressure measurement: are there sex differences? *Procedia Computer Science* 192:2912–2918. <https://doi.org/10.1016/j.procs.2021.09.062>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Agostino Accardo, Associate professor of Biomedical Engineering and head of the Biomedical Engineering group at the University of Trieste, Italy. His main research interest focuses on physiological signal processing by means of linear and non-linear methods.



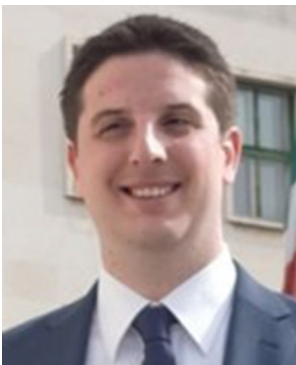
Katerina Iscra, PhD candidate in Biomedical Engineering at the University of Trieste. She graduated in Biomedical Engineering in 2019 in Trieste, Italy. Her main research interest focuses on biomedical signal and image processing applied in neurology and cardiology.



Luca Restivo, graduated from the University of Palermo Medical School in 2017 and currently resident (Cardiovascular Disease) at Trieste University Hospital. He is currently involved in several research projects in heart failure and cardiomyopathies.



Giulia Silveri, Clinical Engineer and Ph.D. in information engineering at University of Trieste. Shi is a IEEE member and her research concerns biomedical signals in particular blood pressure and heart rate variability.



Miloš Ajčević, Senior research associate at the Biomedical Engineering Group, University of Trieste, Italy. His research activities concern biomedical signal and image processing and analysis applied in neurology, ICT solutions for chronic patients, and the development of computer-aided diagnosis.



Marco Merlo, Associate Professor of Cardiology and head of the Cardiomyopathy and Heart Failure Clinic at Trieste University Hospital. His main research interests: genetics, imaging and prognostic stratification of cardiomyopathies.



Aleksandar Miladinović, is a post-doc researcher at the IRCCS Burlo Garofolo, Trieste. His research interest are biomedical signal processing, development of the visual system in children, and development of computer-aided diagnosis.



Gianfranco Sinagra, Full Professor of Cardiology and Director of Cardiovascular Department of Trieste University Hospital. Graduated in Medicine at the Palermo University in 1987. Fellow of the European Society of Cardiology.