Does Ectopic Beats Bring More Discriminatory Information to Diagnose Ischemic Heart Disease?

Katerina Iscra[1], Aleksandar Miladinović[1,2], Miloš Ajčević[1], Luca Restivo[3], Simone Kresevic[1], Marco Merlo[3], Gianfranco Sinagra[3], Agostino Accardo[1]

[1]Department Engineering and Architecture, University of Trieste, Trieste, Italy
[2]Institute for Maternal and Child Health – IRCCS Burlo Garofolo, Trieste, Italy
[3]Cardiovascular Department, ASUGI and University of Trieste, Trieste, Italy

Abstract

Early non-invasive diagnosis of Ischemic Heart Disease (IHD) can often be challenging. HRV features have a potentially important role in risk stratification for subjects with suspected heart disease. However, there is no consensus on the HRV preprocessing steps, particularly on how to properly treat ectopic beats. We aimed to investigate the performance of the models for classification of early IHD versus healthy subjects (HC) based on HRV features extracted from signals excluding ectopic beats and based on the same features extracted from the signals that contain both ectopic and normal heartbeats. This study encompassed 385 subjects (170 IHD and 215 HC). The models were produced by logistic regression method considering two sets of HRV features obtained by two preprocessing approaches. The results showed that the model with the input features from HRV signals including normal and ectopic beats presented a higher classification accuracy (72.7%) than the model based on features extracted only from normal heart beats (67.8%). In addition, the evaluation of the feature importance by analysis of produced nomograms and observed significant differences between features extracted with two preprocessing approaches, showed also that the exclusion of the ectopic beats modifies the features' discriminatory power between HC and IHD.

1. Introduction

The most prevalent cause of cardiovascular mortality is Ischemic Heart Disease (IHD) also referred to as angina and myocardial infarction. The condition typically occurs when there is an imbalance between myocardial oxygen supply and demand [1]. An early and accurate diagnosis of IHD is necessary to improve outcomes. However, diagnosis of IHD can often be challenging because only invasive, and not largely available exams can provide a definite diagnosis. Indeed, only coronary angiography, an invasive tool requiring the use of possibly toxic contrast means, can definitively diagnose IHD.

There is growing research interest in the development of machine learning models for computer-aided diagnosis of different cardiopathies [2], especially those based on features extracted from non-invasive techniques, as heart rate variability (HRV) analysis.

The changes in tonic vagal activity and sympathetic-parasympathetic disbalance, characteristic of ischemic heart disease [3], can be measured by HRV [4], which reflects the fluctuations in beat-to-beat heart rate (RR interval). HRV is calculated by analyzing RR intervals from sinoatrial node beats, and it can be examined in a variety of methods, including time and frequency domain analyses, as well as non-linear analyses [5]. HRV can be utilized to assess several cardiac diseases [6, 7]. However, there is no consensus on the HRV preprocessing steps, that can potentially bring different results. One of the issues still debated is the inclusion of the ectopic beats in HRV analysis [7-13]. Some studies exclude ectopic beats from HRV analysis considering them biological artifacts [8] or irrelevant due to fact that they are not generated by sinoatrial nodes [9, 10]. Their exclusion, however, creates a challenge for interpolation of the RR intervals and can bias the HRV parameters [7, 11, 12], especially when these are caused by the cardiovascular autonomic tone changes [11]. Indeed, albeit still much debated, such bias and the inclusion of ectopic beats can potentially be relevant for discrimination of IHD [13]. For this reason, the inclusion of ectopic beats should be potentially considered in HRV feature extraction [14].

Therefore, we aimed to investigate the performance of the models for classification of early IHD versus healthy subjects based on HRV features extracted from signals excluding ectopic beats and based on the same features extracted from the signals that contain both ectopic and normal heartbeats.

2. Methods

In this study, we analyzed clinical data and processed ECG signals of 385 subjects. In particular, the study encompassed 170 patients affected by early IHD (125M/45F, aged 71±11 y) and 215 healthy controls (101M/114F, aged 58±20 y). The assessment of IHD was based on clinical and laboratory findings [15]. Only early-stage IHD were included in the study (patients without cardiac insufficiency symptoms or with cardiac insufficiency symptoms classified by New York Heart Association (NYHA) scale as class 1). IHD patients did not present acute coronary syndrome in the 3 months before the Holter monitoring. Patients with known trigger factors, such as toxic insults from alcohol or drugs, and tachyarrhythmias were also excluded. IHD 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 performed according to the Declaration of Helsinki and all patients gave written consent.

All subjects underwent a 24h Holter ECG recording using the ambulatory electrocardiographic recorder SpiderView (Sorin Group, Italy) with a sampling rate of 200Hz. 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 (E), artifacts (A), and calibration (C). The RR interval records were cut into 5 min segments without overlap. For each segment, in the case where only normal beats were considered, the RR intervals labeled with E were excluded. In the preprocessing step, only 5min segments that contain at least 60 beats were labeled as valid. In the case where ectopic beats were also considered, each RR 5 min segment was included in the analysis only if the longest ectopic beats subsequence (labeled with E) or the longest artifact subsequence (labeled with A) does not exceed 10s. The RR marked with a calibration label was ignored in both cases. In each case these segments were interpolated with cubic spline and resampled at 2 Hz, producing two different HRV signals. Subsequently, for each signal, and in each segment, linear and non-linear HRV features were extracted. In particular, the linear parameters MeanRR and SDNN related to the RR variability were calculated directly from the RR sequence [16], whilst in the frequency domain, the absolute and relative powers in low (LF=0.04−0.15Hz; LFn) and high (HF=0.15−0.40Hz; HFn) frequency bands and their ratio (LF/HF), were estimated from the interpolated HRV signals (the one only with normal beats and the one that contains both normal and ectopic beats). The non-linear analysis was carried out by calculating Poincaré plot parameters (SD1, SD2) reflecting short and long-term variability [17] and extracting Fractal Dimension (FD) [18] quantifying the complexity of the system that generates the signal. Finally, the median of all features from valid 5 min segments during 24h were calculated and used as the input features for the classifier.

The Logistic Regression (LogReg) method [19], used for diagnostic modeling because of its easy interpretability in the clinical domain, was employed to produce models capable of differentiating between the two groups (IHD and HC). The models were produced considering HRV features obtained from 1) signals after exclusion of ectopic beats ($LogReg_N$) and 2) signals which included both normal and ectopic beats ($LogReg_{NE}$). In both cases, the total number of 10 aforementioned HRV features was considered. The classification performance of the produced models was estimated using 5-fold cross-validation. For each model we calculated the classification accuracy (CA), AUC, F1, precision, and recall.

Nomograms were used to interpret the obtained logistic regression models. Beside the prediction, the logistic regression nomogram reveals the structure of the model and the relative impacts of the features on the class probability. The lengths of the lines are related to the spans of odds ratios, providing the information on feature importance. Furthermore, nomograms allow the computation of scores for each feature, which may be used to determine not only the classification outcome but also the class belonging probability [20]. The features that individually contribute at least 10 out of 100 points in the nomograms were plotted and considered for further statistical analysis.

The HRV features represented in the nomograms extracted from signals excluding ectopic beats and obtained from signals which included ectopic beats were compared using the paired t-test. A $p<0.05$ was considered statistically significant.

3. Results

Classification performance of LogReg models based on HRV features extracted from signals after excluding ectopic beats and features obtained from signals which included normal and ectopic beats are reported in Table 1. The CA, AUC, F1, precision, and recall were higher in the model based on features that were extracted from HRV that included both normal and ectopic beats ($LogReg_{NE}$) compared to the logistic regression model constructed with the features obtained from HRV excluding ectopic beats ($LogReg_N$).

Table 1. Classification performance of produced $LogReg_N$ and $LogReg_{NE}$ models

Model	AC	AUC	F1	Precision	Recall
$LogReg_N$	0.678	0.714	0.677	0.677	0.678
$LogReg_{NE}$	0.727	0.810	0.725	0.726	0.727

The produced nomograms for $LogReg_{NE}$ and $LogReg_{N}$ are reported in Figure 1 and the features are listed in order of importance allowing to select the subset of most informative features. The most discriminatory features were SD2, SDNN, LF, HF, MeanRR, LF/HF, SD1 and SD2, SDNN, SD1, LF/HF, HF, MeanRR, LF for $LogReg_{NE}$ and $LogReg_{N}$, respectively.

Table 2. Mean±SD and comparison between $LogReg_{N}$ and $LogReg_{NE}$ features in HC subjects

Features	$LogReg_{N}$ Mean±SD	$LogReg_{NE}$ Mean±SD	p-value
MeanRR	874±137	879±138	<0.001
SDNN	66±43	69±48	0.058
LF	1165±2386	1042±2470	<0.001
HF	1501±4727	1681±5282	<0.001
LF/HF	2.31±2.06	2.23±2.12	<0.001
SD1	30±25	33±31	0.094
SD2	82±50	82±49	0.022

Mean±SD and comparison between $LogReg_{N}$ and $LogReg_{NE}$ features in HC and IHD subjects are reported in Table 2 and Table 3, respectively. All features except SDNN and SD1 in HC subjects (Table 2), and MeanRR and SD2 in IHD subjects (Table 3) were significantly different.

Table 3. Mean±SD and comparison between $LogReg_{N}$ and $LogReg_{NE}$ features in IHD subjects

Features	$LogReg_{N}$ Mean±SD	$LogReg_{NE}$ Mean±SD	p-value
MeanRR	963±151	960±149	0.807
SDNN	81±62	88±72	<0.001
LF	1521±2816	1406±2836	<0.001
HF	3193±6294	4170±9693	<0.001
LF/HF	1.22±1.09	1.10±1.13	<0.001
SD1	40±35	46±44	<0.001
SD2	95 ±69	97±73	0.565

4. Discussion

The main finding of this study is that the model with the input features extracted from RR segments with normal and ectopic beats, also called Heart rate total variability, presented higher classification performance (72.7%) in comparison to the model that uses features based only on normal heart beats (67.8%). In addition, the model based only on heart rate total variability features was able to correctly classify between early-stage IHD subjects and healthy subjects with moderately high accuracy.

The nomogram revealed that the most important features were SD2, SDNN, LF, HF, meanRR, LF/HF and SD1, emphasizing that SD2 and SDNN were among the most discriminatory. The SD2 changes suggest activation of both the parasympathetic and sympathetic nervous systems, namely via a fast vagal response (parasympathetic) and the slow sympathetic response [21]. Indeed, the IHD stroke patients show a typical suppression of the SD2 of the Poincaré plot [22] that represents the long-term HRV changes. The SD2, related to autonomic nervous system dysfunction, provides additional information about the IHD [22], which is in line with our study, as the SD2 was identified by nomograms as the most important feature. Similar information about the dysfunction of the autonomic nervous system in IHD patients is also measurable by LF and HF parameters [22], that in our study takes 3rd and 4th place, in order of importance, in $LogReg_{NE}$ nomogram and 4th and 7th place in $LogReg_{N}$ nomogram.

Furthermore, it can be observed on nomograms that the SD2 was invariant in case that ectopic beats were excluded or not. Indeed, SD2 did not differ between the two preprocessing approaches in the IHD group, and it should be preferred as a feature. The similar behavior was observed for meanRR, although with less discriminatory power. On the other hand, the different SDNN impact on total nomogram score in IHD's was in line with detected differences of this parameter between two preprocessing approaches. Nonetheless the SDNN was less invariant to the exclusion of ectopic beats, it remained the 2nd most important feature. The opposite trend was observed in the HC group for SD2 and SDNN, suggesting that effect of the exclusion of ectopic beats have different influence on these features in two considered groups. Moreover, SD1, a feature that reflects the short-time HRV changes, was significantly different in the IHD, and not in the HC group, which probably caused by the increased number of ectopic beats in the IHDs. However, it can be observed on nomograms that SD1 has more discriminatory power in the

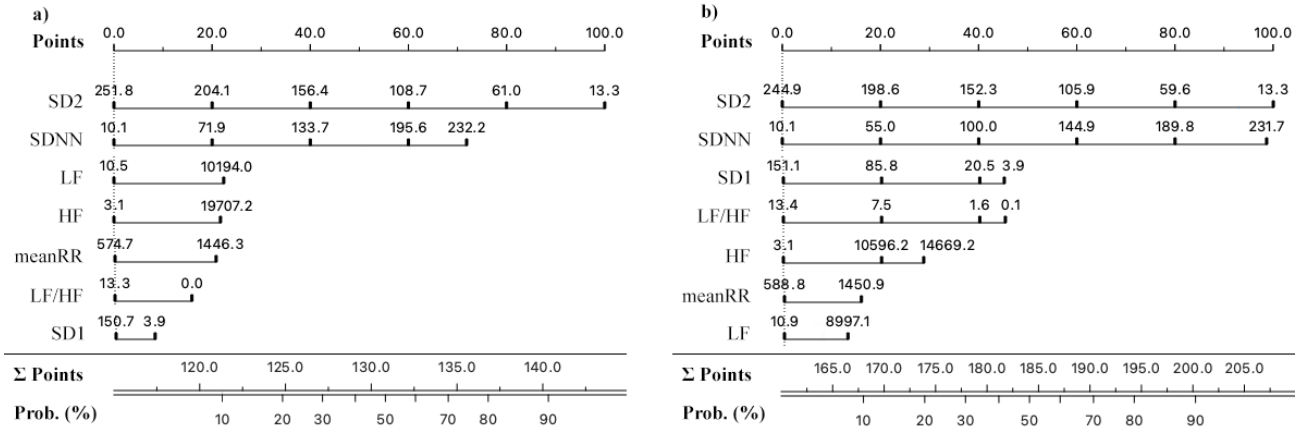

Figure 1. Nomograms for IHD output class for (a) $LogReg_{NE}$ and (b) $LogReg_{N}$ models.

case when ectopic beats are excluded, which can be related to the RR interpolation [11]. LF, HF as well as LF/HF were statistically different between two preprocessing methods in both groups.

In conclusion, our results showed that inclusion of ectopic beats might bring more discriminatory power and help to better identify between early-stage IHD and healthy individuals. The evaluation of the feature importance and the assessed differences between extracted features, showed also that the exclusion of the ectopic beats modifies the features' discriminatory power between HC and IHD. These findings should be confirmed in futures studies on a lager study sample, considering also different methods of ectopic beats exclusion.

Acknowledgments

Work partially supported by the master program in Clinical Engineering of the University of Trieste.

References

[1] G. Lippi, M. Franchini, and G. Cervellin, “Diagnosis and management of ischemic heart disease,” *Semin. Thromb. Hemost.*, vol. 39, no. 2, pp. 202–213, Mar. 2013

[2] M. A. Ahmad, C. Eckert, and A. Teredesai, “Interpretable machine learning in healthcare,” in *Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, New York, NY, USA, Aug. 2018, pp. 559–560

[3] K. Sroka, “On the genesis of myocardial ischemia,” *Z. Kardiol.*, vol. 93, no. 10, pp. 768–783, Oct. 2004

[4] G. Ernst, “Heart-rate variability—more than heart beats?” *Front. Public Health*, vol. 5, p. 240, Sep. 2017

[5] U. Rajendra Acharya, K. Paul Joseph, N. Kannathal, C. M. Lim, and J. S. Suri, “Heart rate variability: a review,” *Med. Biol. Eng. Comput.*, vol. 44, no. 12, pp. 1031–1051, Dec. 2006

[6] A. Accardo, G. Silveri, M. Ajčević, A. Miladinović, and L. Pascazio, “Influence of smoking and other cardiovascular risk factors on heart rate circadian rhythm in normotensive and hypertensive subjects,” *PLOS ONE*, vol. 16, no. 9, p. e0257660, Sep. 2021

[7] L. Zhao, P. Li, J. Li, and C. Liu, “Influence of ectopic beats on heart rate variability analysis,” *Entropy*, vol. 23, no. 6, p. 648, May 2021

[8] M. A. Salo, H. V. Huikuri, and T. Seppänen, “Ectopic beats in heart rate variability analysis: effects of editing on time and frequency domain measures,” *Ann. Noninvasive Electrocardiol. Off. J. Int. Soc. Holter Noninvasive Electrocardiol. Inc*, vol. 6, no. 1, pp. 5–17, Jan. 2001

[9] B. Singh, D. Singh, A. K. Jaryal, and K. K. Deepak, “Ectopic beats in approximate entropy and sample entropy-based HRV assessment,” *Int. J. Syst. Sci.*, vol. 43, no. 5, pp. 884–893, May 2012

[10] P. Albrecht and R. J. Cohen, “Estimation of heart rate power spectrum bands from real-world data: dealing with ectopic beats and noisy data,” in *Proceedings. Computers in Cardiology 1988*, Sep. 1988, pp. 311–314.

[11] N. Lippman, K. M. Stein, and B. B. Lerman, “Comparison of methods for removal of ectopy in measurement of heart rate variability,” *Am. J. Physiol.*, vol. 267, no. 1 Pt 2, pp. H411-418, Jul. 1994

[12] J. Mateo and P. Laguna, “Analysis of heart rate variability in the presence of ectopic beats using the heart timing signal,” *IEEE Trans. Biomed. Eng.*, vol. 50, no. 3, pp. 334–343, Mar. 2003

[13] S. M. Al-Khatib *et al.*, “2017 AHA/ACC/HRS Guideline for management of patients with ventricular arrhythmias and the prevention of sudden cardiac death: executive summary,” *Circulation*, vol. 138, no. 13, pp. e210–e271, Sep. 2018.

[14] A. Accardo, G. Silveri, M. Merlo, L. Restivo, M. Ajčević, and G. Sinagra, “Detection of subjects with ischemic heart disease by using machine learning technique based on heart rate total variability parameters,” *Physiol. Meas.*, Oct.2020

[15] J. Knuuti *et al.*, “2019 ESC Guidelines for the diagnosis and management of chronic coronary syndromes,” *Eur. Heart J.*, vol. 41, no. 3, pp. 407–477, Jan. 2020

[16] “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,” *Circulation*, vol. 93, no. 5, pp. 1043–1065, Mar. 1996.

[17] M. A. Woo, W. G. Stevenson, D. K. Moser, R. B. Trelease, and R. M. Harper, “Patterns of beat-to-beat heart rate variability in advanced heart failure,” *Am. Heart J.*, vol. 123, no. 3, pp. 704–710, Mar. 1992

[18] T. Higuchi, “Approach to an irregular time series on the basis of the fractal theory,” *Phys. Nonlinear Phenom.*, Jun. 1988

[19] A. Urso, A. Fiannaca, M. La Rosa, V. Ravì, and R. Rizzo, “Data mining: prediction methods,” in *Encyclopedia of Bioinformatics and Computational Biology*, S. Ranganathan, M. Gribskov, K. Nakai, and C. Schönbach, Eds. Oxford: Academic Press, 2019, pp. 413–430.

[20] M. Možina, J. Demšar, M. Kattan, and B. Zupan, “Nomograms for visualization of naive bayesian classifier,” in *Knowledge Discovery in Databases: PKDD 2004*, Berlin, Heidelberg, 2004, pp. 337–348.

[21] D. Noronha Osório *et al.*, “Autonomic nervous system response to remote ischemic conditioning: heart rate variability assessment,” *BMC Cardiovasc. Disord.*, vol. 19, p. 211, Sep. 2019

[22] J. T. Korpelainen, K. A. Sotaniemi, A. Mäkikallio, H. V. Huikuri, and V. V. Myllylä, “Dynamic behavior of heart rate in ischemic stroke,” *Stroke*, vol. 30, no. 5, pp. 1008–1013, May 1999

Address for correspondence:
Katerina Iscra
University of Trieste
Via A. Valerio 10, 34127, Trieste, Italy
katerina.iscra@phd.units.it