Noname manuscript No.
(will be inserted by the editor)

Multivariate Ensemble Classification for the Prediction of Symptoms in Patients with Brugada Syndrome

Daniel Romero[1] · **Mireia Calvo**[1] · **Virginie Le Rolle**[2] · **Nathalie Behar**[2] · **Phillipe Mabo**[2] · **Alfredo Hernández**[2]

Received: date / Accepted: date

Abstract Identification of asymptomatic patients at higher risk for suffering cardiac events remains controversial and challenging in Brugada syndrome (BS). In this work, we proposed an ECG-based classifier to predict BS-related symptoms, by merging the most predictive electrophysiological features derived from the ventricular depolarization and repolarization periods, along with autonomic-related markers. The initial feature space included local and dynamic ECG markers, assessed during a physical exercise test performed in 110 BS patients (25 symptomatic). Morphological, temporal and spatial properties quantifying the ECG dynamic response to exercise and recovery were considered. Our model was obtained by proposing a two-stage feature selection process, that combined a resampled-based regularization approach with a wrapper model assessment for balancing, simplicity and performance. For the classification step, an ensemble was constructed by several logistic regression base classifiers, whose outputs were fused using a performance-based weighted average. The most relevant predictors corresponded to the repolarization interval, followed by two autonomic markers and two other makers of depolarization

Corresponding author:
Alfredo Hernández
email: alfredo.hernandez@inserm.fr

[1]Institute for Bioengineering of Catalonia (IBEC),
Campus Besòs EEBE-UPC, Ave. E. Maristany 16,
Building C, L5.3 Barcelona, E-08019, Spain

[2]Univ Rennes, CHU Rennes, Inserm, LTSI - UMR 1099, F-35000 Rennes, France

Total number of words: 6926
Abstract words: 196
Number of tables: 3
Number of figure: 5

dynamics. Our classifier allowed for the identification of novel symptom-related markers from autonomic and dynamic ECG responses during exercise testing, suggesting the need for multifactorial risk stratification approaches in order to predict future cardiac events in asymptomatic BS patients.

Keywords Brugada syndrome · depolarization disorders · ensemble classifier · heart-rate recovery · sudden cardiac death

Authors biography

Daniel Romero, PhD, is a postdoctoral researcher at the Institute for Bioengineering of Catalonia, former post-doctoral researcher at University of Rennes 1, with expertise on biomedical signal processing, data analysis and modeling, applied to the study of cardio-respiratory diseases.

Mireia Calvo, PhD, is a postdoctoral researcher at the Institute for Bioengineering of Catalonia, former post-doctoral researcher at University of Rennes 1, with expertise on biomedical signal processing and modeling, applied to the study of cardiac diseases.

Virginie Le Rolle, PhD, is an associate professor at the University of Rennes 1, with expertise on biomedical signal processing and cardiovascualr modeling, applied to the study of cardiac diseases.

Nathalie Behar, MD, PhD, is a cardiology specialist at CHU-Rennes with expertise in cardiac electrophysiology and implantable devices for therapies related to cardiac diseases.

Philippe Mabo, MD, PhD, is a cardiology specialist and full professor at CHU-Rennes, with expertise in cardiac electrophysiology and implantable devices for therapies related to cardiac diseases.

Alfredo Hernández, PhD, is a Research Director at INSERM, with expertise on biomedical data processing and computational modeling applied to the study of cardiorespiratory diseases.

Author contributions

Conceptualization: Daniel Romero, Alfredo Hernández; **Methodology**: Daniel Romero, Alfredo Hernández; **Formal analysis and investigation**: Daniel Romero, Mireia Calvo, Virginie Le Rolle, Nathalie Behar, Philippe Mabo, Alfredo Hernández; **Software and visualization**: Daniel Romero; **Writing - original draft preparation**: Daniel Romero; **Writing - review and editing**: Daniel Romero, Mireia Calvo, Virginie Le Rolle, Nathalie Behar, Philippe Mabo, Alfredo Hernández; **Funding acquisition**: Daniel Romero, Alfredo Hernández, Philippe Mabo; **Resources**: Nathalie Behar, Philippe Mabo, Alfredo Hernández; **Supervision**: Philippe Mabo, Alfredo Hernández.

1 Introduction

Brugada syndrome (BS) is a genetic pathology associated with a high risk for sudden cardiac death (SCD) in patients with structurally normal hearts. It is diagnosed when a coved type-1 electrocardiogram (ECG) morphology (ST-segment elevation $\geq$ 2mm) is observed in at least one right precordial lead (V1 and/or V2), placed in the 2nd, 3rd or 4th intercostal space. This pattern can be observed either spontaneously or after a provocative drug challenge test using intravenous Na^+ channel blockers [22,23].

Implantable cardioverter defibrillator (ICD) therapy is recommended (class I) in BS patients who survived an aborted cardiac arrest and/or have suffered spontaneous sustained ventricular tachycardias (VT). It should also be considered (class IIa) in those patients with a spontaneous type-1 ECG pattern and history of syncope. The decision of implanting an ICD is far more complex in those individuals who remain asymptomatic after BS diagnosis, which represent around 60% of the BS population. Indeed, although these patients show a lower risk of arrhythmic events, as reported in the FINGER study [24], this risk is not insignificant and the consequences of not implanting an ICD in these cases is extreme [28]. Therefore, one of the challenges regarding the treatment of this pathology is the identification of useful quantitative markers to identify asymptomatic patients that may have a higher risk of SCD and that may properly benefit from the implantation of an ICD.

Several clinical and experimental electrophysiological studies have suggested that both depolarization and repolarization disorders may significantly contribute to the underlying pathophysiology of BS [15]. The presence of QRS late potentials (LP), fragmented QRS complexes [17], the prolongation of both the PR and QRS durations [19], as well as the appearance of wider S waves in inferolateral leads and a rightward deviation of the QRS loop axis in its terminal quarter [16], are some depolarization disorders that have been associated with BS, mainly related to a slowing conduction within the right ventricular outflow tract (RVOT). Regarding repolarization, most alterations observed from experimental studies have been related to an imbalance of the ionic currents contributing to phase 1 of the action potential (AP) of cardiac cells.

Furthermore, the role of the autonomic nervous system (ANS) seems to have a significant impact in unmasking symptoms in this pathology, particularly during exercise testing. Tan et al. [29] reported that exercise resulted in an increase of the J-point amplitude in BS patients. Makimoto et al. [13] evaluated the relationship between the post-stress parasympathetic activation and ST-segment changes. They reported that parasympathetic reactivation during early recovery after exercise, assessed through the heart rate recovery (HRR) index [4], tends to be higher in BS patients with prior VF episodes. Furthermore, recent works from our group showed that symptomatic BS patients exhibit greater fluctuations in sinus node response to ANS in 24-h Holter recordings [2] and an increased parasympathetic modulation during incremental exercise and early recovery [3].

Multivariate approaches that involve several electrophysiological markers have demonstrated their usefulness for assessing VF risk in BS patients [7], increasing prognosis accuracy. Since both repolarization and depolarization disorders mentioned above are manifested as a consequence of this genetic disease, studies considering both scenarios simultaneously may provide more meaningful insights than if considered separately. Moreover, the assessment of cardiac electrophysiological markers from both ventricular repolarization and depolarization periods, in addition to other autonomic-related parameters during controlled autonomic maneuvers, such as physical exercise testing, could have a great impact in the prognosis of BS patients.

In this study we aimed at designing a specific machine learning (ML) pipeline to: 1) identify potentially predictive risk markers in BS patients based on their symptomatology that may offer new insight; 2) propose a multivariate classifier to discriminate between symptomatic and asymptomatic BS patients, with a good enough trade-off between performance and generalization. To achieve these goals, depolarization and repolarization ECG-derived features were obtained from 12-lead ECG recordings acquired during a standardized exercise stress test. The HRR marker was also considered as an autonomic-related parameter. The parameters with the greatest predictive potential were identified through a resampled-based regularization approach to obtain sparse and interpretable models, with high generalization capability. Then, the final model was validated after further model assessment of an ensemble classifier proposed herein, using a wrapper-based strategy for balancing simplicity and performance.

2 Materials and methods

2.1 Study population

The study population comprises 110 consecutive patients (mean age = 44.6 ± 13.7 years) suffering from Brugada syndrome, enrolled in a prospective, multicentric study led by the cardiology department of the Rennes University Hospital (CHU de Rennes), in collaboration with other french hospitals. The study protocol was approved by the respective local ethics committees: Comité d'Éthique du CHU de Rennes (ID RCB 2007-A00887-46), Comité d'Éthique du CHU Saint-Pierre, Comité d'Éthique du CHU de Nantes, Comité d'Éthique du CHU de Bordeaux, Comité d'Éthique du CHU de Brest and Comité d'Éthique du Centre Hospitalier de La Rochelle; in accordance with the ICH E6 recommendations, and the ethical principles of the Helsinki Declaration (1996 and 2000). All patients provided their written informed consent to participate in the study.

From the total population, 25 patients presented clinical symptoms including syncope or aborted SCD related to VF (symptomatic group) based on their medical history. The remaining patients (asymptomatic group) had not presented BS-related symptoms when the study was conducted. As indicated on

active guidelines, all patients on the symptomatic group have received an ICD. On the asymptomatic group, 19 of 85 (22%) have received an ICD, based on a positive EPS (Electrophysiological Study) test. Age of participants ranged from 19 to 74 years old (mean: 44.6 ± 13.7) and 74.5% were males. Structural heart disease was discarded by physical examination, patient's clinical history, and resting and exercise electrocardiogram. Moreover, no significant left ventricular hypertrophy was observed on any patient during echocardiographic screening. More details of this clinical study, including information on the genetic background, can be found on [2]. Table 1 summarizes the baseline characteristics of this population.

During a physical exercise test performed on a cycle ergometer (Ergoline 900 Egamed, Piestany, Slovakia), standard 12-lead ECG recordings sampled at 1000 Hz were acquired for all patients. The protocol used for this standardized test is defined as follows:

- Exercise phase (EX): 2 minutes of initial workload by pedaling at 50 W (30 W for women) followed by successive increments of 30 W (20 W for women) every 2 minutes until the patient reached at least the 80% of his/her maximal theoretical heart rate, defined by $HR_{max} = 220 - age$.
- Recovery phase (RE): two successive 3-minute periods, consisting of an active recovery (pedaling at a workload of 50 W), followed by a passive recovery at rest.

2.2 Preprocessing

ECG signals were preprocessed before the automatic extraction of the analyzed features. This preprocessing included automatic QRS complex detection and subsequent visual inspection to guarantee normal beat annotations, cubic spline interpolation for baseline drift attenuation, 4^{th} order Butterworth low-pass filtering at 45 Hz to remove muscular noise and wave delineation using an evolutionary optimization approach [5].

2.3 ECG-derived measures

2.3.1 Depolarization parameters

Classical and unconventional QRS parameters were evaluated at relevant stages of the exercise test for each individual lead. The classical group of parameters included the amplitudes of the R and S waves (R_a and S_a), the duration of the QRS complex (QRS_d) and the maximum vector magnitude (VM) within the spatial QRS loop. The latter was evaluated from the orthogonal leads X, Y and Z, generated by applying the Dower inverse matrix on the standard leads V1-V6, I and II [6].

The unconventional group of parameters included the main slopes (R_U, R_D , S_U) and angles (ϕ_R and ϕ_S) identified on a QRS complex. R_U and R_D

refer to the up- and down-strokes of the R wave, respectively, while S_U is the final up-stroke of the S wave. Details on how to compute these parameters can be found in [26]. The angles ϕ_R and ϕ_S can be calculated from the measured QRS slopes, using (1) and (2), where constants represent the conversion factors needed to standardize angular measures in order to match conventional ECG printouts used in clinical practice [27].

$$\phi_{\mathrm{R}} = \arctan\left(\left|\frac{R_{\mathrm{U}} - R_{\mathrm{D}}}{0.4(6.25 + R_{\mathrm{U}}R_{\mathrm{D}})}\right|\right). \tag{1}$$

$$\phi_{\mathrm{S}} = \arctan\left(\left|\frac{R_{\mathrm{D}} - S_{\mathrm{U}}}{0.4(6.25 + R_{\mathrm{D}}S_{\mathrm{U}})}\right|\right). \tag{2}$$

2.3.2 Repolarization parameters

The spatial dispersion observed in the T wave morphology can be quantified by certain parameters computed from the principal components analysis (PCA) applied on the available ECG leads. The equations below define such parameters:

$$\text{T-wave residuum:} \qquad T_{R_i} = \sum_{l=4}^{L} \delta_{i,l} / \sum_{l=1}^{L} \delta_{i,l} \tag{3}$$

$$\text{T-wave uniformity:} \qquad T_{U_i} = \delta_{i,1} / \sum_{l=1}^{L} \delta_{i,l} \tag{4}$$

$$\text{T-wave complexity:} \qquad T_{C_i} = \sum_{l=2}^{L} \delta_{i,l} / \sum_{j=1}^{L} \delta_{i,l} \tag{5}$$

$$\text{T-wave non-planarity:} \qquad T_{N_i} = \delta_{i,3} / \delta_{i,1} \tag{6}$$

where i is the i-th beat and $\boldsymbol{\delta}_i$ contains the eigenvalues of the interlead repolarization correlation matrix, $\hat{\boldsymbol{R}}_{\boldsymbol{e}_i}$, sorted in descending order. The matrix $\hat{\boldsymbol{R}}_{\boldsymbol{e}_i}$ can be generated as $\hat{\boldsymbol{R}}_{\boldsymbol{e}_i} = \sum_{n=0}^{N-1} \boldsymbol{e}_i(n)\boldsymbol{e}_i^T(n)$, with $\boldsymbol{e}_i(n)$ being the vector for the multilead ECG signal, corresponding to the T-wave interval of the the i-th beat. Note that for each beat, the $N \times L$ matrix $\mathcal{E}_i = [\boldsymbol{e}_{i,1}, \boldsymbol{e}_{i,2}, \ldots, \boldsymbol{e}_{i,L}]$ is obtained, where each column representing the lead l can be seen as an observation of a random process $\boldsymbol{e}$ with zero empirical mean [11].

The parameter T_{R_i} represents the relative energy of the non-dipolar components of the ECG model described in [1], with respect to the total energy (sum of all the eigenvalues). In that model, the sum of the first three components are considered as to represent the energy of the dipolar components in normal conditions. If local heterogeneities are present in the repolarization, larger eigenvalues associated with the non-dipolar components are expected to increase T_{R_i} measures.

T_{U_i} and T_{C_i} are designed to characterize the ST-T loop morphology. T_U values closer to 1 are indicative of narrower ST-T loops that typically lie in the direction of the first eigenvector δ_1. Likewise, T_C quantifies the roundness of the loop, so that values close to 1 indicate an ST-T loop that is mainly contained in a plane. Finally, T_{NP_i} accounts for the non-planarity of the ventricular repolarization.

Spatial dispersion of the repolarization can also be assessed through the analysis of the following ECG intervals: the corrected QT interval (QT_c) using the Bazett's formula, the T wave width, (T_w), the T-wave peak to T-wave end interval (T_{pe}) and the ratio between T_{pe} and QT, (T_{pe}/QT). All these features were calculated in leads V5 and II, usually used for QT related measures in most clinical studies. In order to quantify T-wave symmetry, the following shape-based parameters were also assessed: the ratio of the left- (T_{LA}) and right-side (T_{RA}) areas with respect to T-wave peak, $T_{area} = \frac{T_{LA}}{T_{RA}}$, and the ratio of the time intervals lasted between the T-wave peak and the onset (T_{op}) and end of the T-wave (T_{pe}), respectively, $T_{time} = \frac{T_p - T_o}{T_e - T_p} = \frac{T_{op}}{T_{pe}}$ [12].

2.4 Heart rate recovery

To characterize recovery dynamics, strongly related to vagal response during this period, the heart rate recovery (HRR) curve was assessed by means of the following bi-exponential decay model previously used in [25]:

$$\theta_{\mathrm{HRR}}(t) = A \exp(-\tau_1 t) + B \exp(-\tau_2 t) \tag{7}$$

where τ_1 and τ_2 are the decay rates of two exponential processes, associated with the active and passive periods of recovery in the defined protocol. Parameters A and B account for the amount of change in HR decay caused by the two exponential processes controlled by τ_1 and τ_2.

The HRR curve was fitted after resampling the instantaneous heart rate series (expressed in beats/min) at 4 Hz to have evenly sampled values. Then, the difference between the actual value $\theta_{\mathrm{HRR}}(t)$ and the initial value $\theta_{\mathrm{HRR}}(t_0)$ was computed for $t = \{60, 120, 180\}$ seconds, to obtain the HRR values at minutes 1, 2, and 3 after the peak effort (HRR_1, HRR_2, and HRR_3), respectively. Figure 1 shows an example of the HR evolution for a particular patient and the resulting HRR curve.

2.5 Feature extraction

Mean values ($\bar{\mathcal{Y}}_s^l$) of the ECG markers were computed using beat-to-beat measures from 15-s windows in each lead l, and at three relevant stages (s) including exercise onset (T1), the peak effort (T2) and the end of recovery (T3). For T2 and T3 stages, the 15-s windows were taken just before the maximum heart rate and minute 6^{th} of recovery, respectively. The changes observed between T1 and T2 (exercise period), and between T2 and T3 (recovery period) were

then determined and defined by $\Delta\mathcal{Y}^l_{\mathrm{EX}} = \bar{\mathcal{Y}}^l_{\mathrm{T2}} - \bar{\mathcal{Y}}^l_{\mathrm{T1}}$ and $\Delta\mathcal{Y}^l_{\mathrm{RE}} = \bar{\mathcal{Y}}^l_{\mathrm{T3}} - \bar{\mathcal{Y}}^l_{\mathrm{T2}}$, respectively, where $\mathcal{Y}$ represents the actual marker. Finally, the initial feature database, $\boldsymbol{\mathcal{F}}$, was created including the mean values of all parameters at each stage s, $\bar{\mathcal{Y}}^l_s$, and their dynamics during exercise ($\Delta\mathcal{Y}^l_{\mathrm{EX}}$) and recovery ($\Delta\mathcal{Y}^l_{\mathrm{RE}}$), together with the HRR at the first, second and third minute after the peak effort. Thus, a total of 198 features were initially included in $\boldsymbol{\mathcal{F}}$: 80 concerning cardiac depolarization, 115 from ventricular repolarization and 3 related to the autonomic control.

2.6 Pipeline used for designing the proposed classifier

The proposed BS classifier was designed based on two separate phases (see Fig. 2). Phase 1 is applied to obtain a subset of features with high predictive power for symptoms and is based on the elastic net regularization algorithm using repeated k-fold cross-validation. Phase 2 further exploits this reduced subset of features using a wrapper approach, where the final subset (model) is selected by repeated holdout cross-validation applied on non-regularized models. These two phases are explained in more detail in the following sections.

2.6.1 Phase 1: Selection of high predictive features using elastic net

Since $\boldsymbol{\mathcal{F}}$ is composed of nearly 200 features, some of them probably highly correlated, the first step is to identify the most predictive set of features. In this work, the dimension of the original feature space is greater than the number of patients available. We have thus chosen to apply the elastic net method for feature selection.

Elastic net is a regularized regression method combining LASSO ($\mathcal{L}^1$-norm) and ridge regression ($\mathcal{L}^2$-norm) penalties for the regularization of the model coefficients [31]. This approach solves the limitations of the LASSO method when dealing with highly correlated variables, which tends to select one variable and ignore the others. Elastic net makes it possible to obtain sparse models and thus to better interpret final outcomes. Two parameters are involved in the elastic net regularization: α defines the weight of $\mathcal{L}^1$ versus $\mathcal{L}^2$ optimization, while λ defines the regularization strength used in the optimization process. For $0 < \alpha < 1$ and a non-negative λ, elastic net solves the problem:

$$\min_{\beta_0\beta}\left(\frac{1}{2N}\sum_{i=1}^{N} y_i \cdot (\beta_0 - x_i^T\beta) - \log(1 + e^{\beta_0 - x_i^T\beta}) + \lambda P_\alpha(\beta)\right), \tag{8}$$

where N is the number of observations in the dataset, y_i is the response at observation x_i, a vector containing the input features, and $P_\alpha(\beta)$ represents the penalty term that interpolates between the $\mathcal{L}^1$ and $\mathcal{L}^2$ norms of β coefficients, defined as:

$$P_\alpha(\beta) = \frac{(1-\alpha)}{2}||\beta||_2^2 + \alpha||\beta||_1 = \sum_{j=1}^{p}\left(\frac{(1-\alpha)}{2}\beta_j^2 + \alpha|\beta_j|\right) \tag{9}$$

Parameter α must be defined between 0 and 1. Values of α=1 and α=0 represent the LASSO and ridge regressions, respectively. Other values of α represent elastic net optimization. We have chosen an α value closer to LASSO regression in order to select relevant variables.

Parameter λ can take different values, so that several models with distinct sets of predictive variables can be obtained. An optimal λ value generates a model that seamlessly generalizes new, previously unseen data, while achieving the right balance between simplicity and adjustment of the training data. If λ is too high, the model will be sparse, but with the risk of underfitting the data. Conversely, if λ is too low, the model will be more complex and thus there will be a risk of overfitting the data. Therefore, the best predictive model (with the optimal λ) can be found by selecting the lowest cross-validated error, estimated as the expected deviance of the model applied to new data not seen during training, among all the evaluated λ values.

To find this optimal λ value (hyperparameter optimization), stratified k-fold cross-validation (CV) procedure is often applied [31]. We have implemented a repeated (N=100) K-fold cross-validation approach ($K = 10$) through each λ_i value, with $i = 1, ..., \mathcal{I}$ (number of λ values), as follows:

1. For n from 1 to N
   - Split pseudo-randomly the dataset $\mathcal{F}$ in K folds
   - For k from 1 to K
     - Define L_k (without the k-th fold) and T_k (the k-th fold) as the training and test subsets.
     - For i from 1 to $\mathcal{I}$
       - Build a regularized, CV model $\mathcal{M}^i_{cv}(L_k) = f(L_k, \lambda_i)$.
       - Evaluate $\mathcal{M}^i_{cv}(L_k)$ on T_k and store prediction errors.
   - Calculate the CV error for iteration n in each λ_i, $\epsilon_n(\lambda_i)$.
2. For each λ_i calculate the average error over the N realizations, $\hat{\epsilon}(\lambda_i)$.
3. Define λ_{opt} as the optimal value with the minimal average error.
4. Fit the regularized models $\mathcal{M}^i_{ncv} = f(\mathcal{F}, \lambda_i)$ without CV (ncv) in order to get the variables (with non-zero coefficients) selected for each λ_i, including all models from λ_{opt} to λ_{min} (the simplest model).

The implementation of regularized logistic regression using elastic net, as presented above, requires the application of a feature scaling method, because features' coefficients with large variance are less penalized. As a result, the update speed for the coefficient estimates during the optimization approach depends on their feature scales, and when converged, some coefficients are closer to the optimum than the others. Feature scaling helps with this issue, because coefficients are at the same scale and update roughly with the same speed. Therefore, all variables entering in the regularized features selection process were scaled using the following expression:

$$\hat{f}_q = \frac{f_q - \min(\boldsymbol{f})}{\max(\boldsymbol{f}) - \min(\boldsymbol{f})} \tag{10}$$

where $\boldsymbol{f} = (f_1, ..., f_Q)$ with Q being the total number of patients and $\hat{f}_q$ the normalized data for patient q.

During the above normalization step, no criteria were applied for outliers removal, since their presence in different features and for different patients can lead to a significant reduction of the already small dataset. Nevertheless, after performing the Phase 1, the preselected feature subset was carefully inspected to individually validate features' inclusion in the final evaluation step.

The final objective of this phase is to define optimal feature subsets for further processing. Once λ_{opt} is determined, the features selected on step 4 are thus retained, and defined by the subset $\mathcal{F}_1$. Nevertheless, there may be λ values close to λ_{opt} that produce even simpler models and with very similar CV errors. Variables selected for these other interesting values of λ can also be studied. Therefore, the models corresponding to those λ values between λ_{opt}, denoted as $\mathcal{M}_{ncv}^{Pr}$, and λ_{min} producing the simplest model $\mathcal{M}_{ncv}^{min}$, represent the candidate models to be evaluated in Phase 2. Note that $\mathcal{M}_{ncv}^{min}$ can have one or more features, while Pr represents the model holding all preselected features obtained with λ_{opt}.

2.6.2 Phase 2: Building the classifier using a wrapper-based approach

In order to get a more robust final prediction, the final model used in the proposed BS classifier is determined in a wrapper fashion [9] (see Fig. 3), through an ensemble classification system combining a set of classical, unregularized logistic regression models C, based on the preselected sparse models, $\mathcal{M}_{ncv}^{min}, ..., \mathcal{M}_{ncv}^{Pr}$, $C = \{\mathcal{M}_{Unreg}^{min}, ..., \mathcal{M}_{Unreg}^{Pr}\}$.

Models C were trained using the stacked generalization (SG) approach [10]. SG is a way for combining different models or lower-level sub-models that have been trained for a particular classification task, allowing to derive higher-level models that minimize the generalization error rate, and thus the accuracy of the ensemble. In summary, the entire training set $\mathcal{F}_L$ including only the P features of each model in C, is divided into $j = 1, ..., J$ blocks, and each base classifier c_b^j is first trained on j-1 blocks of the training subset L_b^j, and internally evaluated on the hold out block T_b^j. Then, each trained fold classifier $c_b^j(L_b^j)$ is applied to the test set $\mathcal{F}_T$ (it can be either the training set or an external test set) yielding the output $\mathbf{o}_b^j$. Figure 3 illustrates this stacking procedure using 3 folds, although we set J=10 in our final design.

The final classification decision (right part of Fig. 3) is generated by combining the output of all base classifiers, by means of a fixed rule. In brief, the outputs of each individual base classifier $\boldsymbol{o}_b^j$ (soft outputs or posterior probabilities) were concatenated into the matrix $\boldsymbol{O}_b$, where $\boldsymbol{O}_b = [\boldsymbol{o}_b^1, \boldsymbol{o}_b^2, ..., \boldsymbol{o}_b^J]$, and serves as inputs for a fixed combiner. The proposed combiner is called the trimmed-weighted mean, which discards the worst and best among the low-level base classifiers, and calculates the weighted average with the remaining eight base classifiers [18]. The weight of each classifier, w_j, is defined as its per-

formance obtained on fold j not seen during training. Thus, a well-balanced average posterior probability vector is obtained as:

$$\mu_g(\boldsymbol{x}) = \frac{1}{J}\sum_{j=1}^{J-2} \mathrm{w}_j o_{b,g}^{j}(\boldsymbol{x}) \tag{11}$$

where x represents the actual observation of the input database. This combiner calculates the support for a particular patient group, g= {symptomatic, asymptomatic}, by fusing all individual supports given by the retained base classifiers.

2.6.3 Model performance assessment

To assess the overall performance of the candidate predictive models, C, a repeated stratified holdout cross-validation strategy was used. The whole database was randomly splitted into training ($\mathcal{L}$) and testing ($\mathcal{T}$) subsets using the ratio 70/30. $\mathcal{L}$ was used to train the ensemble classifiers as described in the previous section, resulting in $\mathcal{E}(\mathcal{L})$. Predictive performance was assessed by testing $\mathcal{E}(\mathcal{L})$ on $\mathcal{T}$ through the sensitivity (Se), specificity (Sp), and the area under the ROC curve (AUC) metrics. Since the population size is relatively small, the entire process was repeated $N = 100$ times in order to get an average estimate of the model performance and variance.

As mentioned before, the final model was determined in a wrapper fashion among those candidate subsets shortlisted during the regularization procedure. Therefore, the model producing the best metrics during the performance assessment was then used as our final BS classifier.

3 Results

The initial features dataset $\boldsymbol{\mathcal{F}}$ of 198 features was first reduced by applying the elastic net procedure, thereby discarding the most irrelevant parameters and retaining a few potential candidate models.

3.1 Feature selection

Results obtained from the elastic net-based feature subsets selection and parameter tuning process are illustrated in Figure 4. The model hyperparameters were set as follows: $\alpha = 0.75$ and λ varying between 0.001 and 0.1, logarithmically spaced in 40 steps. The darkest region observed in the heat map (see Fig. 4-a) corresponds to those λ values where CV errors are lower. By averaging such errors, the optimal λ value, λ_{opt}, that produces the smallest mean error, is determined as indicated in Fig. 4-b (vertical dashed line in black). However, an alternative λ can be chosen if the error does not exceed a certain limit referring to that calculated for λ_{opt}. A common practice is to apply the

1 standard error (SE) rule when CV is performed once [31]. Since this rule is no longer valid when using repeated CV because it reduces SE values as a result of averaging, we selected the most predictive model through a wrapper-based approach. In the proposed method, all λ values between λ_{opt} and that one producing the simplest model (just one feature) were tested. The latter is represented in Fig. 4-b by the vertical dashed line in red.

In addition to CV errors, Fig. 4-b shows the number of features selected for each λ. Depending on λ, models with 1 and up to 30 features can be obtained with $\lambda_{min} \geq \lambda \geq \lambda_{opt}$ ($\mathcal{M}_{ncv}^{min}, ..., \mathcal{M}_{ncv}^{Pr}$), representing a significant reduction (85%) from the initial number of variables in the dataset ($P = 198$).

The variables selected in each predictive model can be extracted from the coefficients' pathway generated as a function of λ during the regularization procedure. Table 2 lists the model (feature subsets) obtained for λ_{opt} (P=30), with features ordered by importance. It shows that features related to ventricular depolarization were slightly higher in number (15 features) than those associated with repolarization (13 features). HRR-related indices deserve special attention as well, particularly when evaluated at later intervals of recovery such as at minutes 2 and 3. Of note, the importance of each feature within the model is given by its order of appearance in the list, derived from the coefficient's pathway.

On the other hand, features evaluated at specific phases (T1, T2, T3) of the test protocol were predominant (about two-thirds) as compared to those related to dynamic behaviors either during exercise or recovery. T1 (baseline phase) was the most important among the three studied phases for relevant repolarization features, while T2 and T3 for depolarization. Furthermore, recovery prevailed over exercise in terms of dynamic-related variables retained (9 features for recovery including HRR_2 and HRR_3 and 2 features for exercise). The overall performance achieved with this model, called **Model 1** is included in Table 3, for both the training and testing sets.

3.2 Ensemble classifier performance

Fig. 5-a shows the average performance metrics obtained for all predictive models $C = \{\mathcal{M}_{Unreg}^{min}, ..., \mathcal{M}_{Unreg}^{Pr}\}$, where $\mathcal{M}_{Unreg}^{Pr}$ matches **Model 1**. As observed in Fig. 5-b, the model containing only 7 features, **Model 2**, was the most predictive among all wrapped candidate models, since it presented the smallest Euclidean distance (*dist*) between one unit and the average Se and Sp values.

The features retained for this simpler **Model 2** correspond to the first seven features listed in Table 2, which are delimited by the horizontal dashed line. It is worth noting that these variables also appeared in **Model 1**, but in that case accompanied by other, less relevant features. Specifically, three local features related to repolarization period were preserved and only two from depolarization. HRR_2 and HRR_3 were also retained, confirming their relevance when predicting symptoms in BS. In addition, the first three features

in Table 2 correspond to stage T1, highlighting the importance for the inclusion of basal measures. The average performance metrics of **Model 2** and that of the model based on the first 3 features, **Model 3**, were also included in Table 3 for the training and testing subsets.

For **Model 1**, the results presented in Table 3 show a perfect predictive accuracy when tested with the training subset. However, when the test subset is used, **Model 3** and in particular **Model 2**, achieved similar or even better results than **Model 1** indicating that this simpler model is more generalizable and robust when applied to new, unseen patients.

4 Discussion

In this study, a robust predictive classifier has been implemented to discriminate between Brugada syndrome patients according to their symptomathology. To do this, ECG-derived markers that capture dynamic variations from ventricular repolarization and depolarization during physical exercise were assessed. These markers aimed at quantifying the temporal, spatial and morphological properties of the QRS complex and STT interval, since they might be affected by the electrophysiological disorders associated with this disease. Additionally, autonomic-related markers such as the HRR, estimated immediately after the maximum effort, was also considered.

First, a feature selection approach was implemented to efficiently reduce the initial feature dataset. It was achieved by a regularized optimization approach (elastic net). Then, a wrapper-based approach was used for final model evaluation. The first stage allowed to obtain sparse, predictive models while discarding irrelevant predictors [31]. Then, the unregularized version of these models underwent a second-round test to evaluate their predictive performance using a designed ensemble classifier. The model that presented the best balance between its overall performance and the number of predictors was selected to implement the final classifier.

The above-mentioned approach allowed to obtain a model of only 7 predictive features (**Model 2**) from a total of 198. From those features, the most important were related to the repolarization period, which characterized the morphological shape (T_{time}) and dispersion of the T wave (T_{pe}, T_{pe}/QT), all of them evaluated at baseline stages. Two other features associated with the recovery process of the heart rate, HRR_2 and HRR_3, were found to be relevant when predicting symptoms in BS, highlighting the importance of autonomic control in this pathology, specially the vagal modulation immediately after exercise [3,25] or during sleep. The remaining predictors in the model were related to ventricular depolarization dynamics ($\Delta S_{U_{RE}}^{V2}$, $\Delta S_{a_{RE}}^{V2}$) evaluated during recovery. Thus, the results confirm the fact that Brugada syndrome is a complex multifactorial pathology, that needs to be addressed by considering factors related to the whole ventricular activity and the autonomic control, in an effort to obtain better predictive results [8,21].

The performance metrics achieved can be considered acceptable given the challenge of distinguishing these two groups, whose electrophysiological properties might be quite similar. Of note, these metrics were estimated by repeated cross-validation (hold out) 100 times, so as to obtain a realistic performance estimate when tested on new data, not seen during training. Also noteworthy is the novel configuration used in the design of the classifier, where through the fusion of several base classifiers, trained in different data subsets, a more robust prediction of the classes is obtained. Moreover, it should be noted that these results were obtained exclusively from electrocardiographic markers, without including any clinical parameters, as it is usually done in many clinical studies. Further work is directed towards the integration of such clinical data in a similar multivariate classifier.

In general, the results obtained in this study have shown the usefulness of the stress test to unmask the differences between symptomatic and asymptomatic BS patients. For example, we can highlight from the parameters included in the final model, **Model 2**, that the recovery of the heart rate is more informative when evaluated in later stages of this period, and not in the first minute as it is commonly reported [4]. In fact, in a previous study performed on the same population, we have demonstrated that the best timing for assessing significant differences among these patient groups through the HRR, was found between the minute 2.5 and 3 of recovery [25]. Another finding related to this period was that the two depolarization markers involved in the model, were associated with the second half of the QRS complex, and the dynamics of these markers. This suggests that both ventricular depolarization dynamics and vagal activity reactivation during that period are closely related to the symptomatology in this population. Finally and most importantly, the markers linked to ventricular repolarization appeared to be the most relevant to the model, which were tested separately in **Model 3**. The most remarkable in this case is that these particular markers were assessed at the beginning of the test, which gives them greater relevance because they could be measured at rest, avoiding exercise testing.

Quantifying risk level in asymptomatic Brugada patients is still a major clinical challenge. Better predictive markers are needed to improve prognosis and to optimize the therapy for a given patient. The analysis of ventricular depolarization and repolarization combined with measures related to autonomic control provides valuable information about the patients' condition and prognosis. In a study performed by Postema et al. [20], repolarization abnormalities in Brugada syndrome were reported to be mostly induced by depolarization disorders. This was thought to be a result of the heterogeneity in the action potential duration with a ventricular endo-epi gradient. However, in Meijborg et al. [14], repolarization abnormalities were thought to be related to an increased interventricular and LV-intraventricular dispersion in repolarization time, after dofetilide infusion in an experimental model of dofetilide-induced long QT syndrome type 2 (LQT2). Tokioka et al. [30] reported that the combination of both repolarization and depolarization abnormalities enables potential identification of high- and low-risk Brugada patients. However, all these studies did

not consider autonomic related markers in their analyses, which can provide an added value in the final outcomes. To our knowledge, the present work is the first in providing such quantitative comparison combining depolarization, repolarization and autonomic features for risk stratification in Brugada syndrome.

Limitations

According to results, despite the proposed classifier may seem potentially helpful to differentiate among Brugada syndrome patients, the relatively small sample population may have limited the impact of the results obtained. Moreover, the class imbalance existing between the two patient groups can also affect the prediction of the minority class in favor of the other. Therefore, all markers used in the model should be validated with larger clinical series to confirm its predictive potential. Finally, although the lack of follow-up information does not allow us to propose the use of these features as potential risk markers of future cardiac events, they remain as promising candidates that should be tested in other cohorts having such information.

Conclusions

The results obtained in this study confirm the significance of performing controlled autonomic manoeuvres, such as a physical exercise test, to elucidate the potential differences that may exist between patients with Brugada syndrome. Symptomatic and asymptomatic patients can produce different responses to this type of challenge, which can be detected by means of an exhaustive analysis of the electrophysiological properties and autonomic parameters that are likely to be affected by this pathology. The main conclusion of the study concerns the fact that both repolarization and depolarization characteristics measured at baseline and during exercise, are crucial in understanding the pathogenesis of this disease, and that the analysis of heart rate recovery at longer intervals may emerge as a sensitive risk marker.

Acknowledgments

This work was partly supported by a grant of the French Ministry of Health (Programme Hospitalier de Recherche Clinique - PHRC Regional). D. Romero acknowledges the financial support of the Fondation Lefoulon-Delalande, Institut de France, France. M. Calvo acknowledges the financial support of the social program funded by CaixaBank, Spain.

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Acar, B., Yi, G., Hnatkova, K., Malik, M.: Spatial, temporal and wavefront direction characteristics of 12-lead T-wave morphology. Medical and Biological Engineering and Computing **37**(5), 574–584 (1999). DOI 10.1007/BF02513351
2. Behar, N., Petit, B., Probst, V., Sacher, F., Kervio, G., Mansourati, J., Bru, P., Hernandez, A., Mabo, P.: Heart rate variability and repolarization characteristics in symptomatic and asymptomatic Brugada syndrome. Europace **19**(10), 1730–1736 (2017). DOI 10.1093/europace/euw224
3. Calvo, M., Romero, D., Le Rolle, V., Béhar, N., Gomis, P., Mabo, P., Hernández, A.I.: Multivariate classification of Brugada syndrome patients based on autonomic response to exercise testing. PLoS ONE **13**(5) (2018). DOI 10.1371/journal.pone.0197367
4. Cole, C.R., Blackstone, E.H., Pashkow, F.J., Snader, C.E., Lauer, M.S.: Heart-rate recovery immediately after exercise as a predictor of mortality. New England Journal of Medicine **341**(18), 1351–1357 (1999). DOI 10.1056/NEJM199910283411804
5. Dumont, J., Hernández, A.I., Carrault, G.: Improving ECG beats delineation with an evolutionary optimization process. IEEE Transactions on Biomedical Engineering **57**(3), 607–615 (2010). DOI 10.1109/TBME.2008.2002157
6. Edenbrandt, L., Pahlm, O.: Vectorcardiogram synthesized from a 12-lead ECG: Superiority of the inverse Dower matrix. Journal of Electrocardiology **21**(4), 361–367 (1988). DOI 10.1016/0022-0736(88)90113-6
7. Kawazoe, H., Nakano, Y., Ochi, H., Takagi, M., Hayashi, Y., Uchimura, Y., Tokuyama, T., Watanabe, Y., Matsumura, H., Tomomori, S., Sairaku, A., Suenari, K., Awazu, A., Miwa, Y., Soejima, K., Chayama, K., Kihara, Y.: Risk stratification of ventricular fibrillation in Brugada syndrome using noninvasive scoring methods. Heart Rhythm **13**(10), 1947–1954 (2016). DOI 10.1016/j.hrthm.2016.07.009
8. Kawazoe, H., Nakano, Y., Ochi, H., Takagi, M., Hayashi, Y., Uchimura, Y., Tokuyama, T., Watanabe, Y., Matsumura, H., Tomomori, S., Sairaku, A., Suenari, K., Awazu, A., Miwa, Y., Soejima, K., Chayama, K., Kihara, Y.: Risk stratification of ventricular fibrillation in Brugada syndrome using noninvasive scoring methods. Heart Rhythm **13**(10), 1947–1954 (2016). DOI 10.1016/j.hrthm.2016.07.009
9. Kohavi, R., John, G.H.: Wrappers for feature subset selection. Artificial Intelligence **97**(1-2), 273–324 (1997). DOI 10.1016/s0004-3702(97)00043-x
10. Kuncheva, L.I.: Combining Pattern Classifiers: Methods and Algorithms: Second Edition, vol. 9781118315231 (2014). DOI 10.1002/9781118914564
11. Laguna, P., Martinez Cortes, J.P., Pueyo, E.: Techniques for Ventricular Repolarization Instability Assessment from the ECG (2016). DOI 10.1109/JPROC.2015.2500501
12. Langley, P., Di Bernardo, D., Murray, A.: Quantification of T wave shape changes following exercise. PACE - Pacing and Clinical Electrophysiology **25**(8), 1230–1234 (2002). DOI 10.1046/j.1460-9592.2002.01230.x
13. Makimoto, H., Nakagawa, E., Takaki, H., Yamada, Y., Okamura, H., Noda, T., Satomi, K., Suyama, K., Aihara, N., Kurita, T., Kamakura, S., Shimizu, W.: Augmented ST-segment elevation during recovery from exercise predicts cardiac events in patients with brugada syndrome. Journal of the American College of Cardiology **56**(19), 1576–1584 (2010). DOI 10.1016/j.jacc.2010.06.033
14. Meijborg, V.M., Chauveau, S., Janse, M.J., Anyukhovsky, E.P., Danilo, P.R., Rosen, M.R., Opthof, T., Coronel, R.: Interventricular dispersion in repolarization causes bifid T waves in dogs with dofetilide-induced long QT syndrome. Heart Rhythm **12**(6), 1343–1351 (2015). DOI 10.1016/j.hrthm.2015.02.026
15. Meregalli, P.G., Wilde, A.A., Tan, H.L.: Pathophysiological mechanisms of Brugada syndrome: Depolarization disorder, repolarization disorder, or more? (2005). DOI 10.1016/j.cardiores.2005.03.005
16. Morita, H., Fukushima-Kusano, K., Nagase, S., Miyaji, K., Hiramatsu, S., Banba, K., Nishii, N., Watanabe, A., Kakishita, M., Takenaka-Morita, S., Nakamura, K., Saito, H., Emori, T., Ohe, T.: Sinus node function in patients with Brugada-type ECG. Circulation Journal **68**(5), 473–476 (2004). DOI 10.1253/circj.68.473
17. Morita, H., Kusano, K.F., Miura, D., Nagase, S., Nakamura, K., Morita, S.T., Ohe, T., Zipes, D.P., Wu, J.: Fragmented QRS as a marker of conduction abnormality and a

predictor of prognosis of Brugada syndrome. Circulation **118**(17), 1697–1704 (2008). DOI 10.1161/CIRCULATIONAHA.108.770917
18. Polikar, R.: Ensemble based systems in decision making (2006). DOI 10.1109/MCAS.2006.1688199
19. Postema, P.G., van Dessel, P.F., Kors, J.A., Linnenbank, A.C., van Herpen, G., Ritsema van Eck, H.J., van Geloven, N., de Bakker, J.M., Wilde, A.A., Tan, H.L.: Local Depolarization Abnormalities Are the Dominant Pathophysiologic Mechanism for Type 1 Electrocardiogram in Brugada Syndrome. A Study of Electrocardiograms, Vectorcardiograms, and Body Surface Potential Maps During Ajmaline Provocation. Journal of the American College of Cardiology **55**(8), 789–797 (2010). DOI 10.1016/j.jacc.2009.11.033
20. Postema, P.G., van Dessel, P.F., Kors, J.A., Linnenbank, A.C., van Herpen, G., Ritsema van Eck, H.J., van Geloven, N., de Bakker, J.M., Wilde, A.A., Tan, H.L.: Local Depolarization Abnormalities Are the Dominant Pathophysiologic Mechanism for Type 1 Electrocardiogram in Brugada Syndrome. A Study of Electrocardiograms, Vectorcardiograms, and Body Surface Potential Maps During Ajmaline Provocation. Journal of the American College of Cardiology **55**(8), 789–797 (2010). DOI 10.1016/j.jacc.2009.11.033
21. Postema, P.G., Wilde, A.A.: Risk stratification in Brugada syndrome: Where is the finish line? (2016). DOI 10.1016/j.hrthm.2016.08.008
22. Priori, S.G., Blomström-Lundqvist, C., Mazzanti, A., Bloma, N., Borggrefe, M., Camm, J., Elliott, P.M., Fitzsimons, D., Hatala, R., Hindricks, G., Kirchhof, P., Kjeldsen, K., Kuck, K.H., Hernandez-Madrid, A., Nikolaou, N., Norekvål, T.M., Spaulding, C., Van Veldhuisen, D.J., Kolh, P., Lip, G.Y., Agewall, S., Barón-Esquivias, G., Boriani, G., Budts, W., Bueno, H., Capodanno, D., Carerj, S., Crespo-Leiro, M.G., Czerny, M., Deaton, C., Dobrev, D., Erol, Ç., Galderisi, M., Gorenek, B., Kriebel, T., Lambiase, P., Lancellotti, P., Lane, D.A., Lang, I., Manolis, A.J., Morais, J., Moreno, J., Piepoli, M.F., Rutten, F.H., Sredniawa, B., Zamorano, J.L., Zannad, F.: 2015 ESC Guidelines for the management of patients with ventricular arrhythmias and the prevention of sudden cardiac death: The Task Force for the Management of Patients with Ventricular Arrhythmias and the Prevention of Sudden Cardiac Death of the European Society of Cardiology (ESC). Europace **17**(11), 1601–1687 (2015). DOI 10.1093/europace/euv319
23. Priori, S.G., Wilde, A.A., Horie, M., Cho, Y., Behr, E.R., Berul, C., Blom, N., Brugada, J., Chiang, C.E., Huikuri, H., Kannankeril, P., Krahn, A., Leenhardt, A., Moss, A., Schwartz, P.J., Shimizu, W., Tomaselli, G., Tracy, C., Document Reviewers, Ackerman, M., Belhassen, B., Estes, N.A., Fatkin, D., Kalman, J., Kaufman, E., Kirchhof, P., Schulze-Bahr, E., Wolpert, C., Vohra, J., Refaat, M., Etheridge, S.P., Campbell, R.M., Martin, E.T., Quek, S.C., Heart Rhythm Society, European Heart Rhythm Association, Asia Pacific Heart Rhythm Society: Executive summary: HRS/EHRA/APHRS expert consensus statement on the diagnosis and management of patients with inherited primary arrhythmia syndromes. (2013). DOI 10.1093/europace/eut272
24. Probst, V., Veltmann, C., Eckardt, L., Meregalli, P.G., Gaita, F., Tan, H.L., Babuty, D., Sacher, F., Giustetto, C., Schulze-Bahr, E., Borggrefe, M., Haissaguerre, M., Mabo, P., Le Marec, H., Wolpert, C., Wilde, A.A.: Long-term prognosis of patients diagnosed with brugada syndrome: Results from the finger brugada syndrome registry. Circulation **121**(5), 635–643 (2010). DOI 10.1161/CIRCULATIONAHA.109.887026
25. Romero, D., Béhar, N., Mabo, P., Hernández, A.: A biexponential approach for assessing parasympathetic reactivation after submaximal exercise. In: Computing in Cardiology, vol. 44, pp. 1–4 (2017). DOI 10.22489/CinC.2017.306-390
26. Romero, D., Ringborn, M., Laguna, P., Pahlm, O., Pueyo, E.: Depolarization changes during acute myocardial ischemia by evaluation of QRS slopes: Standard lead and vectorial approach. IEEE Transactions on Biomedical Engineering **58**(1), 110–120 (2011). DOI 10.1109/TBME.2010.2076385
27. Romero, D., Ringborn, M., Laguna, P., Pueyo, E.: Detection and quantification of acute myocardial ischemia by morphologic evaluation of QRS changes by an angle-based method. In: Journal of Electrocardiology, vol. 46, pp. 204–214 (2013). DOI 10.1016/j.jelectrocard.2013.02.014
28. Sacher, F., Probst, V., Maury, P., Babuty, D., Mansourati, J., Komatsu, Y., Marquie, C., Rosa, A., Diallo, A., Cassagneau, R., Loizeau, C., Martins, R., Field, M.E., Derval, N., Miyazaki, S., Denis, A., Nogami, A., Ritter, P., Gourraud, J.B., Ploux, S., Rollin,

A., Zemmoura, A., Lamaison, D., Bordachar, P., Pierre, B., Jais, P., Pasquic, J.L., Hocini, M., Legal, F., Defaye, P., Boveda, S., Iesaka, Y., Mabo, P., Haissaguerre, M.: Outcome After Implantation of a Cardioverter-Defibrillator in Patients With Brugada Syndrome A Multicenter Study-Part 2. CIRCULATION **128**(16), 1739–1747 (2013). DOI {10.1161/CIRCULATIONAHA.113.001941}

29. Tan, H.L., Amin, A.S., De Groot, E.A., Ruijter, J.M., Wilde, A.A.: Exercise-induced ECG changes in brugada syndrome. Circulation: Arrhythmia and Electrophysiology **2**(5), 531–539 (2009). DOI 10.1161/CIRCEP.109.862441
30. Tokioka, K., Kusano, K.F., Morita, H., Miura, D., Nishii, N., Nagase, S., Nakamura, K., Kohno, K., Ito, H., Ohe, T.: Electrocardiographic parameters and fatal arrhythmic events in patients with brugada syndrome: Combination of depolarization and repolarization abnormalities. Journal of the American College of Cardiology **63**(20), 2131–2138 (2014). DOI 10.1016/j.jacc.2014.01.072
31. Zou, H., Hastie, T.: Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society. Series B: Statistical Methodology **67**(2), 301–320 (2005). DOI 10.1111/j.1467-9868.2005.00503.x

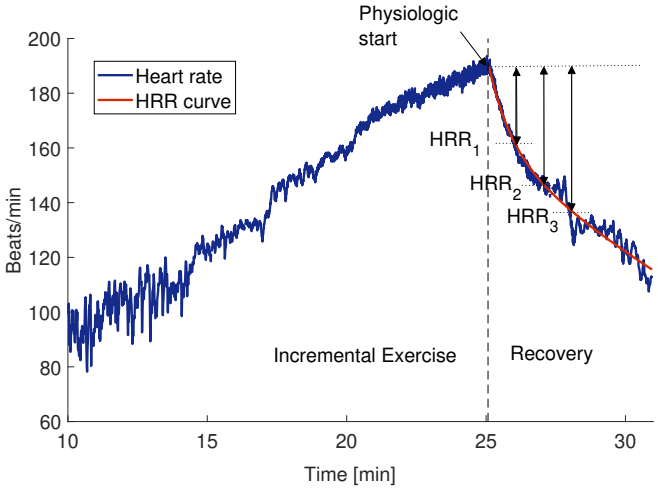

Fig. 1 An example of the HRR curve fitted on HR values during the recovery period of the exercise test. HRR_1, HRR_2, and HRR_3 represent the heart rate recovery measures evaluated at the first, second and third minute of recovery, respectively.

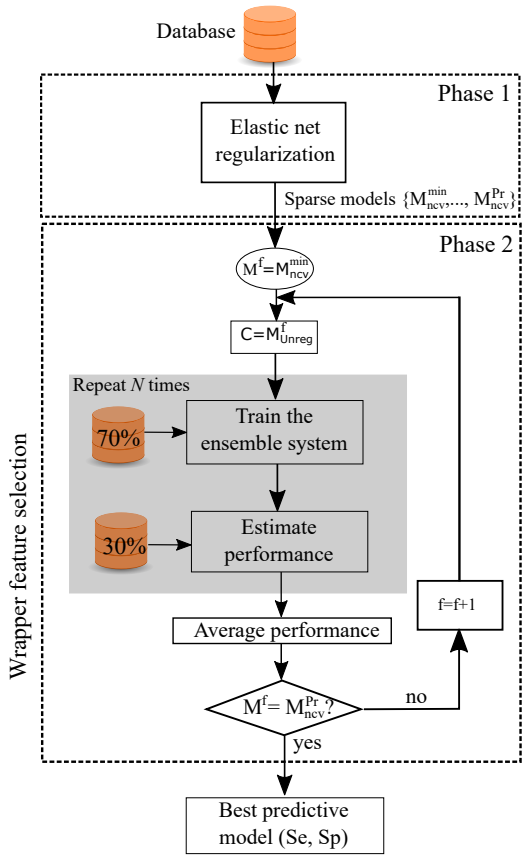

Fig. 2 General work-flow including the regularized feature selection process (Phase 1) and the wrapper step (Phase 2) using the ensemble system. *ncv*: non cross-validated model; *Pr* index of the preselected model obtained for λ_{opt}: Unreg: unregularized model; f: index of the actual wrapped model.

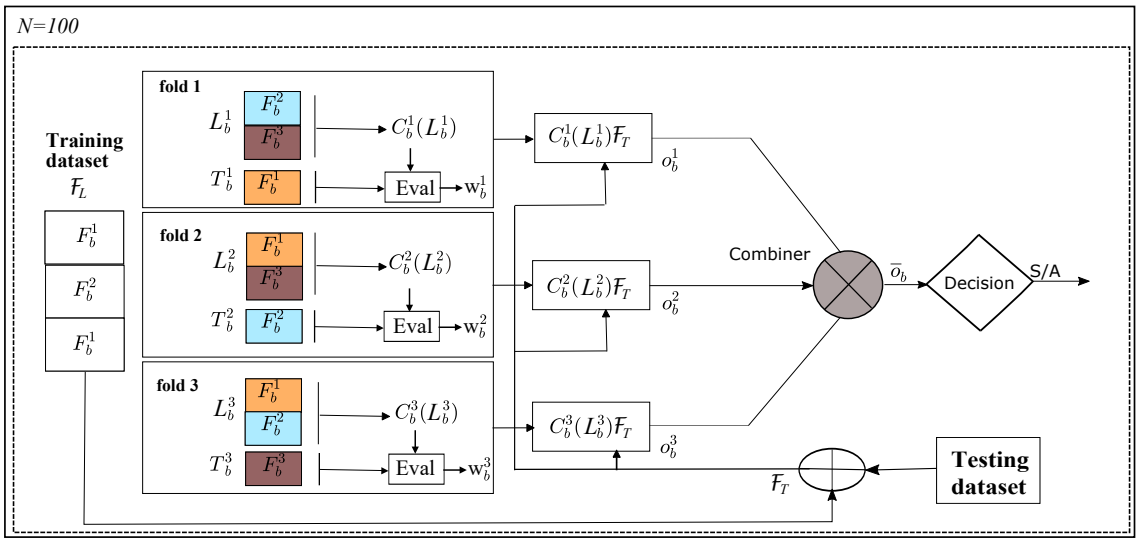

Fig. 3 Ensemble classifier used in the wrapper step. The *stacking* procedure is shown for three folds where three base classifiers are created and their outputs fused through the combiner (arithmetic mean). F_b^k: folds generated when splitting the training set; L_b^k, T_b^k: training and testing subsets of fold k; C_b^k base classifiers; o_b^k soft outputs used as input in the combiner.

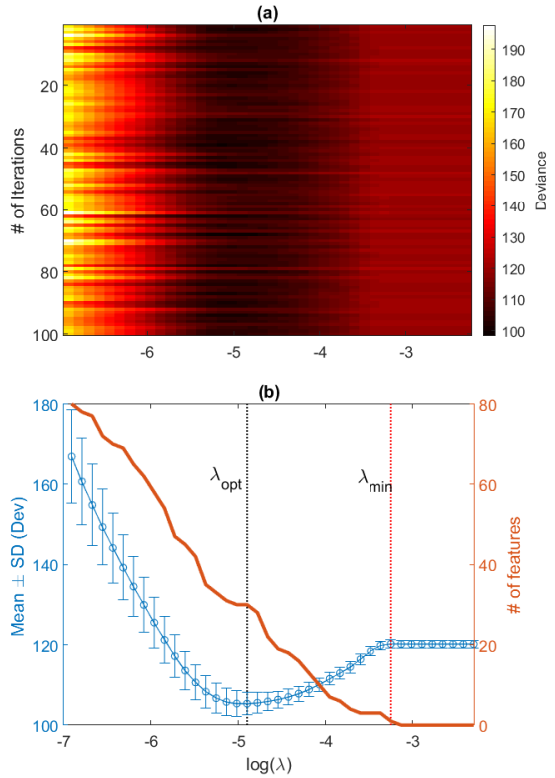

Fig. 4 Results from the elastic net regression of Phase 1, to find the most predictive features to be used in the ensemble classifier. (a) Deviance heat map obtained after 100 realizations (repeated CV: 100 times, 10-fold) using different λ values. (b) Mean $\pm$ SD of the deviance values displayed in a) (left axis), and the number of features selected for each λ value (right axis). Vertical dashed line in black indicates the optimal λ producing the lowest mean deviance.

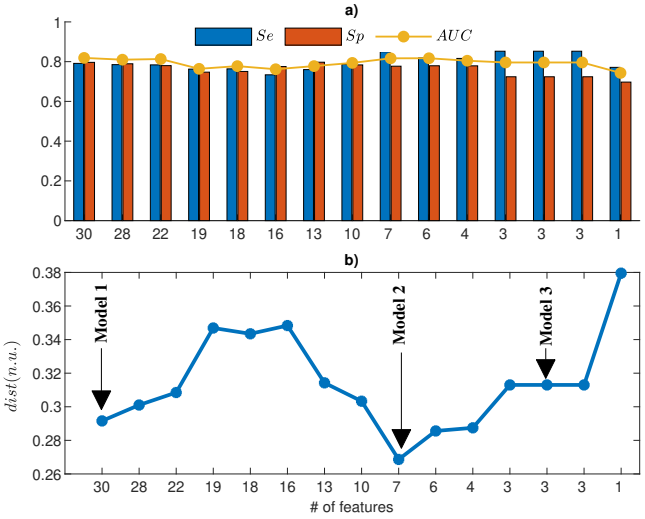

Fig. 5 a) Average performance metrics obtained for the testing in all models $C = \{\mathcal{M}^{min}_{Unreg}, ..., \mathcal{M}^{Pr}_{Unreg}\}$. b) Euclidean distance obtained for each model based on Se an Sp values. **Model 1**, **Model 2** and **Model 3**, corresponds to all features, first 7 features, and first 3 features in Table 2, respectively. Se: Sensitivity; Sp: Specificity; AUC: area under the ROC curve.

Table 1 Baseline clinical characteristics of the study population.

Clinical features	**Symptomatic**	**Asymptomatic**	**p-value**
Sex males	21 (84%)	61 (72%)	0.22
Age (years)	46.8 ±15.5	44.0 ±13.1	0.63
Symptoms			
Cardiac arrest	11 (44%)		
Syncope	14 (56%)		
Spontaneous Type-1	7 (28%)	27 (32%)	0.72
ICD implanted	25 (100%)	19 (22%)	< 0.01
SNC5A mutation (N=79)			
Positive	7 (28%)	22 (26%)	
Negative	13 (52%)	37 (44%)	0.85
Max. HR (beats/min)	151.7 ±17.5	160.9 ±18.5	0.07
Max. workload (Watts)	171.2 ±57.8	170.1 ±57.9	0.66

p-value is obtained using the Mann-Whitney test for continuous features and χ^2 test for categorical features.

Table 2 Ordered subset of the 30 features selected by choosing the optimal λ value, λ_{opt}. The first seven features correspond to the smaller subset associated with the final selected model (**Model 2**). The numbers in the first column indicate the models in which each feature has been involved.

Model	Features	Description
1,2,3	$T_{time}{}^{\mathrm{V5}}_{\mathrm{T1}}$	T-wave intervals ratio in lead V5 at baseline
1,2,3	$T_{pe}/QT^{\mathrm{V5}}_{\mathrm{T1}}$	Ratio T-peak-T-end/QT in lead V5 at baseline
1,2,3	$T_{pe}{}^{\mathrm{V5}}_{\mathrm{T1}}$	T-peak-T-end interval in lead V5 at baseline
1,2	HRR_2	Heart rate recovery evaluated after 2 minutes
1,2	$\Delta S_{U}{}^{\mathrm{V2}}_{\mathrm{RE}}$	S-wave upslope's dynamic in lead V2 during recovery
1,2	HRR_3	Heart rate recovery evaluated after 3 minutes
1,2	$\Delta S_{a}{}^{\mathrm{V2}}_{\mathrm{RE}}$	S-wave amplitude dynamic in lead V2 during recovery
1	$S_{U}{}^{\mathrm{V3}}_{\mathrm{T3}}$	S-wave upslope in lead V3 when recovery ended
1	$T_{time}{}^{\mathrm{V5}}_{\mathrm{T3}}$	T-wave time ratio in lead V5 when recovery ended
1	$\Delta S_{U}{}^{\mathrm{V3}}_{\mathrm{RE}}$	S-wave upslope's dynamic in lead V3 during recovery
1	$T_{C\,\mathrm{T1}}$	T-wave complexity evaluated at baseline
1	$T_{U\,\mathrm{T1}}$	T-wave uniformity evaluated at baseline
1	$\Delta S_{U}{}^{\mathrm{V5}}_{\mathrm{RE}}$	S-wave upslope's dynamic in lead V5 during recovery
1	$S_{U}{}^{\mathrm{V3}}_{\mathrm{T1}}$	S-wave upslope in lead V3 at baseline
1	$S_{a}{}^{\mathrm{V2}}_{\mathrm{T3}}$	S-wave amplitude in lead V2 when recovery ended
1	$R_{a}{}^{\mathrm{V4}}_{\mathrm{T2}}$	R-wave amplitude in lead V4 at maximum effort
1	$\Delta QRS_{d\,\mathrm{RE}}$	QRS duration dynamic during recovery
1	$R_{a}{}^{\mathrm{V4}}_{\mathrm{T3}}$	R-wave amplitude in lead V4 when recovery ended
1	$\Delta S_{a}{}^{\mathrm{V5}}_{\mathrm{EX}}$	S-wave amplitude dynamic in lead V5 during exercise
1	$T_{w}{}^{\mathrm{II}}_{\mathrm{T1}}$	T-wave width in lead II at baseline
1	$\phi_{R}{}^{\mathrm{V6}}_{\mathrm{T2}}$	R-wave angle in lead V6 during maximum effort
1	$\Delta T_{time}{}^{\mathrm{II}}_{\mathrm{EX}}$	T-wave intervals ratio's dynamic, lead II during exercise
1	$T_{pe}{}^{\mathrm{II}}_{\mathrm{T1}}$	T-peak-T-end interval in lead II at baseline
1	$\phi_{R}{}^{\mathrm{V5}}_{\mathrm{T1}}$	R-wave angle in lead V5 during baseline
1	$\phi_{R}{}^{\mathrm{V5}}_{\mathrm{T2}}$	R-wave angle in lead V6 when recovery ended
1	$\Delta T_{time}{}^{\mathrm{II}}_{\mathrm{RE}}$	T-wave intervals ratio's dynamic, lead II during recovery
1	$\Delta QTc{}^{\mathrm{V5}}_{\mathrm{RE}}$	Corrected QT interval dynamic,lead V5 during recovery
1	$T_{time}{}^{\mathrm{II}}_{\mathrm{T2}}$	T-wave intervals ratio in lead II during maximum effort
1	$T_{area}{}^{\mathrm{II}}_{\mathrm{T3}}$	T-wave area's ratio in lead II when recovery ended
1	$\phi_{R}{}^{\mathrm{V6}}_{\mathrm{T3}}$	R-wave angle in lead V5 during maximum effort

Table 3 Predictive model performance of the ensemble classifier when trained with the models **Model 1**, **Model 2** and **Model 3**. The symbols † and ⋆ stand for the biased (training set) and cross-validated performances (testing set), respectively.

Model	Number of features	*Se* (%) Mean ± SD	*Sp* (%) Mean ± SD	*AUC* (%) Mean ± SD
Model 1	30			
Training set		100±0.0	100±0.0	$100\pm0.0^{(\dagger)}$
Testing set		79.1±8.7	79.6±10.3	$81.9\pm6.3^{(\star)}$
Model 2	7			
Training set		87.9±4.8	81.9±4.8	$89.4\pm2.9^{(\dagger)}$
Testing set		85.0±11.1	77.7±7.6	$81.7\pm7.6^{(\star)}$
Model 3	3			
Training set		85.6±6.1	71.7±4.9	$82.6\pm3.0^{(\dagger)}$
Testing set		85.3±10.6	72.4±9.6	$79.6\pm7.7^{(\star)}$