ORIGINAL ARTICLE

Machine learning-based risk prediction of malignant arrhythmia in hospitalized patients with heart failure

Qi Wang[1,2] ID, Bin Li[1], Kangyu Chen[2], Fei Yu[2], Hao Su[2], Kai Hu[2], Zhiquan Liu[2], Guohong Wu[2], Ji Yan[2] and Guohai Su[1]*

[1]*Division of Cardiology, Jinan Central Hospital, Cheeloo College of Medicine, Shandong University, No. 105 Jiefang Road, Jinan, 250013, China; and* [2]*Heart Failure Center, The First Affiliated Hospital of USTC, University of Science and Technology of China, Hefei, China*

Abstract

Aims   Predicting the risk of malignant arrhythmias (MA) in hospitalized patients with heart failure (HF) is challenging. Machine learning (ML) can handle a large volume of complex data more effectively than traditional statistical methods. This study explored the feasibility of ML methods for predicting the risk of MA in hospitalized HF patients.

Methods and results   We evaluated the baseline data and MA events of 2794 hospitalized HF patients in the HF cohort in Anhui Province and randomly divided the study population into training and validation sets in a 7:3 ratio. The Lasso-logistic regression, multivariate adaptive regression splines (MARS), classification and regression tree (CART), random forest (RF), and eXtreme gradient boosting (XGBoost) algorithms were used to construct risk prediction models in the training set, and model performance was verified in the validation set. The area under the receiver operating characteristic curve (AUC) and Brier score were employed to evaluate the discrimination and calibration of the model, respectively. Clinical utility of the Lasso-logistic regression model was analysed using decision curve analysis (DCA). The median (Q1, Q3) age of the study population was 70 (61, 77) years, and 39.5% were female. MA events occurred in 117 patients (4.2%) during hospitalization. In the training set ($n = 1964$), the AUC of the XGBoost model was 0.998 [95% confidence interval (CI) 0.997–1.000], which was higher than the other models (all $P < 0.001$). In the validation set ($n = 830$), there was no significant difference in AUC of Lasso-logistic model 1 [AUC: 0.867 (95% CI 0.819–0.915)], Lasso-logistic model 2 [AUC: 0.828 (95% CI 0.764–0.892)], MARS model [AUC: 0.852 (95% CI 0.793–0.910)], RF model [AUC: 0.804 (95% CI 0.726–0.881)], and XGBoost model [AUC: 0.864 (95% CI 0.810–0.918); all $P > 0.05$], which were higher than that of CART model [AUC: 0.743 (95% CI 0.661–0.824); all $P < 0.05$]. Brier scores for all prediction models were less than 0.05. DCA results showed that the Lasso-logistic model had a net clinical benefit. Oral antiarrhythmic drug, left bundle branch block, serum magnesium, D-dimer, and random blood glucose were significant predictors in half or more of the models.

Conclusions   The current study findings suggest that ML models based on the Lasso-logistic regression, MARS, RF, and XGBoost algorithms can effectively predict the risk of MA in hospitalized HF patients. The Lasso-logistic model had better clinical interpretability and ease of use than the other models.

Introduction

According to the latest epidemiological survey, the prevalence of heart failure (HF) among Chinese adults (aged ≥ 35 years) has risen to 1.3%, with an estimated 13.7 million people with HF and a mortality rate of 4.1% among hospitalized patients with HF.[1,2] Malignant arrhythmia, represented by persistent ventricular tachycardia or ventricular fibrillation, is the primary cause of sudden cardiac death (SCD) in hospitalized patients with HF, and its initiation and maintenance are the result of a combination of many factors.[3] It is of important practical value to accurately and

effectively predict the occurrence of malignant arrhythmia and screen for the key predictors to guide clinical decisions and improve the prognosis of patients.

However, SCD risk prediction for patients with HF is currently mainly focused on the out-of-hospital follow-up period.[4,5] There is no effective prediction method for malignant arrhythmia events during hospitalization. In addition, the clinical data of patients with HF are structurally complex and contain high-dimensional, interactive, multisource, and non-linear data, which makes traditional hypothesis-driven statistical analysis difficult to effectively perform.[6] At present, traditional HF risk prediction tools have modest predictive power.[7–9]

In recent years, machine learning (ML) technology based on data-driven decision-making has developed rapidly and demonstrated great potential in the diagnosis, classification, and prediction of HF.[10,11] Whether the effective prediction of malignant arrhythmia can be realized through ML urgently needs to be explored.

In this study, we applied multiple ML algorithms to construct a malignant arrhythmia risk prediction model in hospitalized HF patients. Meanwhile, regularization techniques were combined with traditional logistic regression as contrasts to screen out the optimal model and important predictors.

Methods

Study population and protocol

Data for this study were obtained from the Anhui HF cohort (a prospective, multicentre, and observational clinical registry study). The subjects were patients who were hospitalized for HF in the Department of Cardiology of participating hospitals during the period from December 2016 to October 2018. The inclusion criteria were as follows: (i) age > 18 years old and (ii) the diagnostic criteria for HF were met according to the *Chinese Guidelines for the Diagnosis and Treatment of Heart Failure 2014*[12] and the New York Heart Association functional classification II–IV [left ventricular ejection fraction (LVEF) < = 50% or LVEF > 50% but amino-terminal pro-brain natriuretic peptide > 400 ng/L]. The exclusion criteria were as follows: (i) acute myocardial infarction within the past 3 months and (ii) the patient was unable to communicate effectively or follow up as planned. The detailed study protocol is available in the Chinese Clinical Trials Registry (Registration number: ChiCTR-POC-16010100). The study protocol was in accordance with the tenets of the Declaration of Helsinki and approved by the Hospital Ethics Committee (Approved No. of ethics committee: 2016-163).

This study extracted and analysed the baseline data of the Anhui HF cohort study, and the study endpoint was malignant arrhythmic events during hospitalization. Malignant arrhythmias were defined as sustained ventricular tachycardia or ventricular fibrillation requiring intravenous antiarrhythmic medication or electrical cardioversion or defibrillation intervention.

Data collection and preprocessing

Information about the patients' demographics, medical history, history of cardiac surgery, vital signs on admission, first electrocardiogram, echocardiography, laboratory tests (routine blood tests, blood biochemical tests, blood coagulation tests, thyroid hormones, and some cardiac biomarkers), and medication was collected in detail. Clinical features with missing data in ≥30% of patients were removed, and features with missing data for <30% of the patients were subjected to multiple imputation, and sensitivity analysis was performed on the imputed data (Supporting Information, *Table S1*). As the range of different features varied widely and some of the utilized algorithms required the data to be normalized, *Z*-score normalization was performed after imputation. Ultimately, 103 clinical features were used for model development (*Table S2*).

Model development and evaluation

All the subjects were randomly divided into a training set and a validation set at a ratio of 7:3. Model development and preliminary evaluation were performed in the training set, and model performance was verified with the validation set. The risk prediction models were constructed using Lasso-logistic regression, multivariate adaptive regression splines (MARS), classification and regression tree (CART), random forest, and the eXtreme gradient boosting (XGboost) algorithm. Lasso regularization was applied to reduce the data dimension for 103 clinical features. Using 10-fold cross-validation, $\log \lambda 1$ was obtained for the maximum area under the receiver operating characteristic curve (AUC), and $\log \lambda 2$ was obtained for 1 standard error from the maximum AUC (*Figure S1*). The features screened out by Lasso regularization were included in the multivariate logistic regression by the enter method, and Lasso-logistic models 1 and 2 were constructed with the features with $P < 0.05$. The optimal parameter settings for the other ML models are shown in *Table S3*.

Receiver operating characteristic curves were drawn, and the model discrimination was quantified by the AUC, which was taken as the main index of model selection. Platt scaling method was used to probabilistically calibrate the output of the ML model, the calibration curves were drawn, and Brier scores were calculated to evaluate the model calibration. Brier score was defined as the mean squared difference between the observed and the predicted outcomes (range, 0–1), with smaller values indicating better calibration of the model. In addition, the clinical utility of the Lasso-logistic model was evaluated by decision curve analysis.

Statistical analysis

Continuous variables are expressed as the means ± standard deviation or median (first quartile, third quartile). Categorical variables are expressed as frequencies (percentages). Continuous variables were compared between groups using Student's t-test or the Mann–Whitney U test. Categorical variables were compared between groups using the χ^2 test or Fisher's exact test. Delong's test was used to compare the discrimination of the prediction models. All statistical tests were two-tailed, and $P < 0.05$ was set as the significance level. R statistical software, Version 3.6.0 (The R Foundation for Statistical Computing, Vienna, Austria), was used for all statistical analyses.

Results

Participants' characteristics

A total of 3107 patients from 16 participating hospitals were enrolled in the Anhui HF cohort, and 313 patients were excluded due to duplicate enrolment or missing electrocardiogram or echocardiography indicators. Finally, 2794 patients were included in this study for statistical analysis (median age was 70 (61, 77) years, 39.5% female), including 1964 patients in the training set and 830 patients in the validation set (*Figure 1*). The main baseline characteristics of the patients are presented in *Table 1*. (Refer to *Table S4* for complete information). A total of 117 patients (4.2%) experienced malignant arrhythmia during hospitalization, including 76 patients (65.0%) in the training set.

Important features and model performance

In this study, six risk prediction models, including Lasso-logistic model 1, Lasso-logistic model 2, the MARS model, the CART model, the random forest model, and the XGBoost model, were established for the malignant arrhythmia events of hospitalized patients with HF. The important features and model performance of each prediction model are shown in *Table 2*.

Figure 1 Flow chart of the study protocol. CART, classification and regression tree; ECG, electrocardiogram; Echo, Echocardiography; HF, heart failure; MARS, multivariate adaptive regression splines; RF, random forest; ROC, receiver operating characteristic; XGBoost, eXtreme gradient boosting.
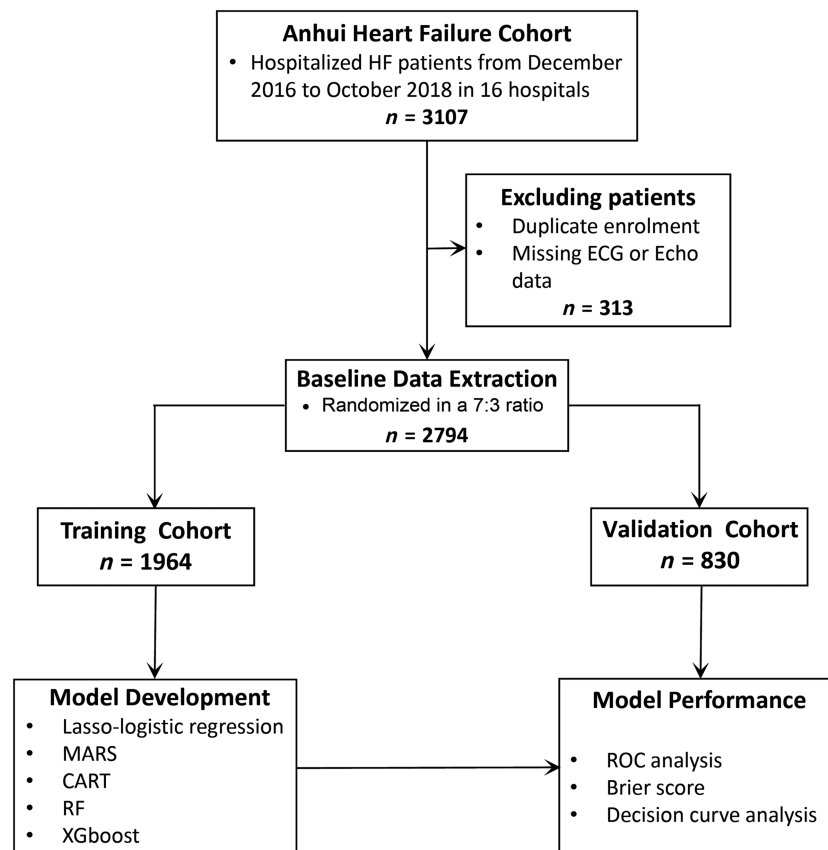
Table 1 Main baseline characteristics of patients in the training and validation sets

Variable	Training set (N = 1964)	Validation set (N = 830)	P value
Age (years)	70 (61,77)	69 (61,77)	0.237
Female	776 (39.5)	327 (39.4)	0.955
BMI (kg/m^2)	23 (21, 26)	24 (21, 27)	0.039
Heart rate (beats/min)	78 (68, 90)	79 (70, 91)	0.370
SBP (mmHg)	128 (114, 143)	129 (115, 144)	0.190
Hospital stay (days)	9 (7, 12)	9 (7, 12)	0.345
NYHA class II	349 (17.8)	159 (19.2)	0.385
Coronary heart disease	898 (45.7)	374 (45.1)	0.748
Hypertension	1008 (51.3)	430 (51.8)	0.815
Atrial flutter or fibrillation	700 (35.6)	296 (35.7)	0.992
Diabetes mellitus	443 (22.6)	176 (21.2)	0.432
COPD	100 (5.1)	39 (4.7)	0.663
QRSd $\geqq$ 120 ms	754(38.4)	302(36.4)	0.318
LBBB morphology	283 (14.4)	114 (13.7)	0.641
LVEF (%)	46 (33, 61)	44 (33, 60)	0.188
NT-proBNP (pg/mL)	2443 (1411, 4657)	2521 (1450, 4930)	0.537
Scr (μmol/L)	85 (69, 109)	84 (69, 108)	0.710
CRT/ICD	92 (4.7)	46 (5.5)	0.339
PCI/CABG	261 (13.3)	116 (14.0)	0.627
ARNI/ACEI/ARB	1287 (65.5)	547 (65.9)	0.849
Beta blocker	1268 (64.6)	523 (63.0)	0.435
Antisterone	1751 (89.2)	746 (89.9)	0.570
Loop diuretic	1829 (93.1)	766 (92.3)	0.432
Cardiotonic drug	1226 (62.4)	530 (63.9)	0.474
Nitrate	1058 (53.9)	457 (55.1)	0.564
Oral antiarrhythmic drug	452 (23.0)	200 (24.1)	0.537
Chinese medicine	144 (7.3)	55 (6.6)	0.508

ACEI, angiotensin converting enzyme inhibitor; ARB, angiotensin receptor antagonist; ARNI, angiotensin receptor neprilysin inhibitor; BMI, body mass index; CABG, coronary artery bypass graft; COPD, chronic obstructive pulmonary diseases; CRT, cardiac resynchronization therapy; ICD, implantable cardioverter defibrillator; LBBB, left bundle branch block; LVEF, left ventricular ejection fraction; NT-proBNP, amino-terminal pro-brain natriuretic peptide; NYHA functional class, New York Heart Association functional class; PCI, percutaneous coronary intervention; QRSd, QRS duration; SBP, systolic blood pressure; Scr, serum creatinine.

Lasso-logistic model

Lasso regularization with logλ1 as the optimal parameter screened out 13 potential predictors. After adjustment by multivariate logistic regression, the following nine factors were independent predictors of malignant arrhythmia in hospitalized patients with HF: left bundle branch block (LBBB) [odds ratio (OR) 14.99 (8.48, 26.50), $P < 0.001$], oral antiarrhythmic drug [OR 5.71(3.28, 9.93), $P < 0.001$], antithrombotic drug [OR 0.28(0.16, 0.50), $P < 0.001$], serum magnesium (Mg) [OR 1.61 (1.37, 1.88), $P < 0.001$], LVEF [OR 0.66 (0.49, 0.88), $P = 0.006$], cardiac metabolic drug [OR 2.57 (1.36, 4.87), $P = 0.004$], D-dimer [OR 1.30 (1.13, 1.50), $P < 0.001$], body mass index [OR 0.56 (0.41, 0.79), $P < 0.001$], and random blood glucose (RBG) [OR 1.40 (1.13, 1.73), $P = 0.002$]. Lasso regularization with logλ2 as the optimal parameter screened out four potential predictors. After adjustment by multivariate logistic regression, LBBB [OR 13.57 (8.04, 22.93), $P < 0.001$], oral antiarrhythmic drug [OR 5.14 (3.07, 8.61), $P < 0.001$], Mg [OR 1.52 (1.30, 1.78), $P < 0.001$], and lactate dehydrogenase [OR 1.28 (1.13, 1.45), $P < 0.001$] were all independent predictors.

Other machine learning models

In other ML models, the contribution of clinical features to outcomes was ranked. The MARS model consisted of 14 features, while the CART model consisted of five features (*Figure 2*). Among the integrated algorithm models, random forest and XGBoost, only the top 15 ranked features were listed (*Figures S2* and *S3*).

Model performance

Figure 3 shows the receiver operating characteristic curves and AUC values for the models. In the training set, the AUC value of the XGBoost model was the highest and was significantly higher than that of the other models (all $P < 0.001$). The MARS model followed, with higher AUC values than Lasso-logistic models 1 and 2, the CART model, and the random forest model ($P = 0.049$, $P < 0.001$, $P < 0.001$, and $P < 0.001$, respectively). Third was Lasso-logistic model 1, which had higher AUC values than Lasso-logistic model 2, the CART model, and the random forest model ($P = 0.037$, $P < 0.001$, and $P < 0.001$, respectively). Fourth was

Table 2 Important features and performance of different prediction models

Model	Important features	Training set AUC (95% CI)	Validation set AUC (95% CI)	Training set Brier score (95% CI)	Validation set Brier score (95% CI)
Lasso-logistic model 1	LBBB, oral antiarrhythmic drug, antithrombotic drug, Mg, LVEF, cardiac metabolic drug, D-dimer, BMI, RBG	0.905 (0.866–0.943)	0.867 (0.819–0.915)	0.027 (0.023–0.035)	0.042 (0.032–0.056)
Lasso-logistic model 2	LBBB, oral antiarrhythmic drug, Mg, LDH	0.881 (0.844–0.918)	0.828 (0.764–0.892)	0.030 (0.026–0.039)	0.041 (0.031–0.054)
MARS	LBBB, Mg, oral antiarrhythmic drug, AST, antithrombotic drug, TRPG, SPAP, D-dimer, ALT, FS, WBC, RBG, globulin, CO₂	0.926 (0.896–0.955)	0.852 (0.793–0.910)	0.025 (0.020–0.032)	0.036 (0.027–0.049)
CART	LBBB, myoglobin, oral antiarrhythmic drug, FT4, LAD	0.773 (0.713–0.832)	0.743 (0.661–0.824)	0.026 (0.020–0.033)	0.042 (0.032–0.057)
Random forest (top15)	Mg, LBBB, neutrophil, RBG, CK, globulin, TP, MCV, CKMB, TSH, LVEDD, WBC, LDH, MCH, CO₂	0.779 (0.720–0.837)	0.804 (0.726–0.881)	0.034 (0.029–0.044)	0.040 (0.030–0.054)
XGBoost (top15)	LBBB, Mg, oral antiarrhythmic drug, CK, TSH, FT4, AST, D-dimer, neutrophil, LAD, MCV, SPAP, age, LVEF, LDL-C	0.998 (0.997–1.000)	0.864 (0.810–0.918)	0.005 (0.003–0.009)	0.037 (0.027–0.051)

ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; CK, creatine kinase; CKMB, creatine kinase MB form; FS, fractional shortening; FT4, free tetraiodothyronine; LAD, left atrial diameter; LBBB, left bundle branch block; LDH, lactate dehydrogenase; LDL-C, low density lipoprotein cholesterol; LVEDD, left ventricular end diastolic diameter; LVEF, left ventricular ejection fraction; MCH, mean corpuscular haemoglobin; MCV, mean corpuscular volume; RBG, random blood glucose; SPAP, systolic pulmonary artery pressure; TP, total protein; TRPG, tricuspid regurgitant pressure gradient; TSH, thyroid stimulating hormone; WBC, leukocyte count.

Figure 2 Classification and regression tree (CART) analysis for malignant arrhythmia. Below 'Yes' and 'No' in the boxes are the proportions of patients in each group with and without malignant arrhythmia. The ratios at the bottom of the boxes represent the number of patients in each group as a percentage of the study population. The colour of the boxes varies with the incidence of malignant arrhythmias. LBBB, left bundle branch block; LAD, left atrial diameter.
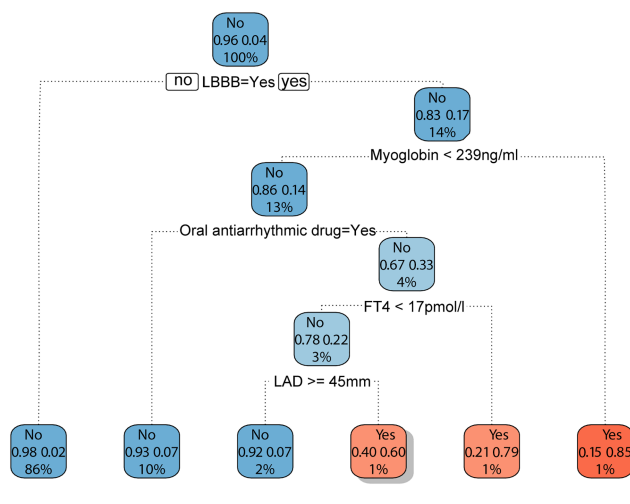
Lasso-logistic model 2, with higher AUC values than the CART model and the random forest model ($P < 0.001$ and $P = 0.004$, respectively). AUC values were similar between the CART model and the random forest model ($P = 0.891$). In the validation set, the AUC value of Lasso-logistic model 1 was the highest, but there were no significant differences

in AUC values among Lasso-logistic models 1 and 2, the MARS model, the random forest model, and the XGBoost model (all $P > 0.05$), and the AUC values of these five prediction models were higher than that of the CART model ($P < 0.001$, $P = 0.002$, $P < 0.001$, $P = 0.049$, and $P = 0.015$, respectively).

Figure 4 shows the calibration curves for the models. Calibration of the models was evaluated by the Brier score method. The Brier scores of all models in both the training set and the validation set were less than 0.05, indicating that the predicted risk was in good agreement with the actual risk.

Figure 5 shows the decision curves of the Lasso-logistic model. Both prediction models exhibited net clinical benefit in the training set and validation set. When the threshold probability was ≤40%, Lasso-logistic model 1 outperformed model 2.

Discussion

In this study, based on the data of the Anhui HF cohort, multiple ML techniques were applied to construct a malignant arrhythmia risk prediction model for the inpatient HF population. Studies have shown that ML models can effectively predict the risk of malignant arrhythmias in hospitalized HF patients. Meanwhile, to overcome the shortcomings of traditional statistical methods in feature selection and data processing, we combined Lasso regularization with logistic regression and found that the joint model had excellent overall performance. In addition, some predictors of

Figure 3 Receiver operating characteristic curves and AUCs of different prediction models. (A) Training set and (B) validation set. AUC, area under the receiver operating characteristic curve; CART, classification and regression tree; MARS, multivariate adaptive regression splines; XGBoost, eXtreme gradient boosting.
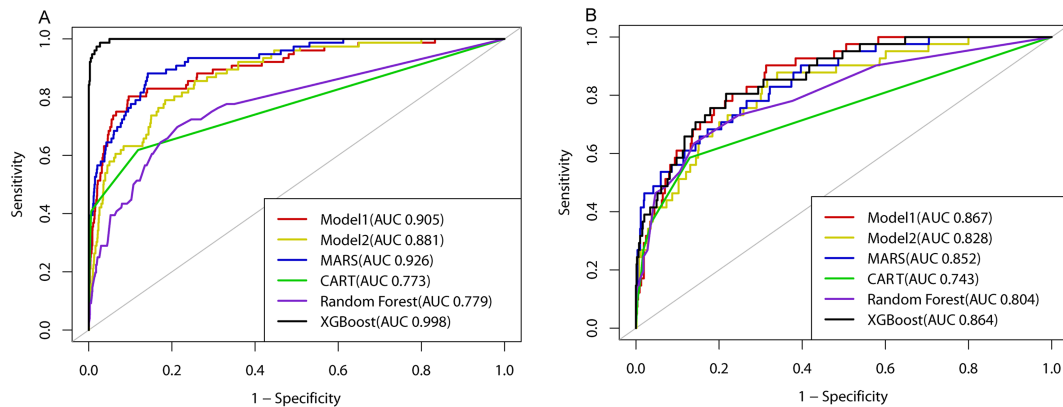
Figure 4 Calibration plots of predicted probabilities and actual proportions for different prediction models. (A) Training set and (B) validation set. CART, classification and regression tree; MARS, multivariate adaptive regression splines; XGBoost, eXtreme gradient boosting.
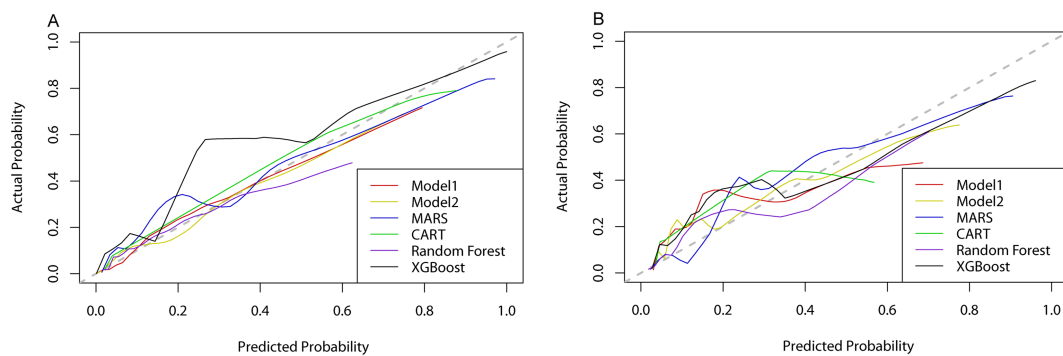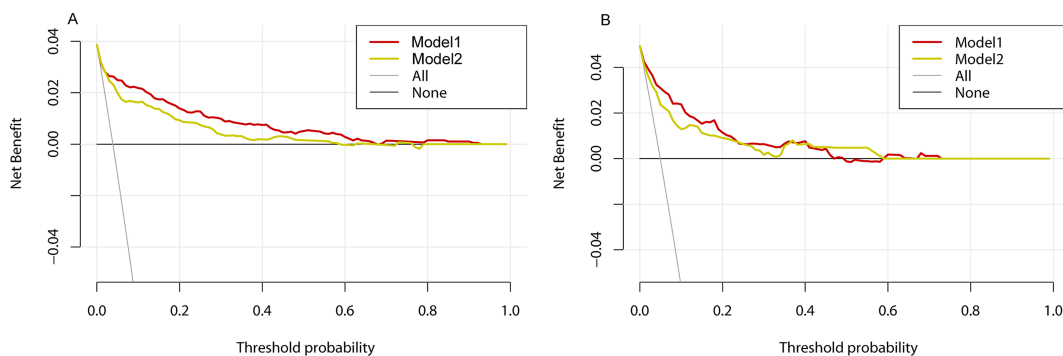
Figure 5 Decision analysis curves of the Lasso-logistic models. (A) Training set and (B) validation set.

malignant arrhythmias, such as oral antiarrhythmic drug, LBBB, Mg, D-dimer, and RBG, were found in this study.

Malignant arrhythmias are the primary cause of SCD in patients with HF, and there is currently a lack of effective prediction methods. In clinical practice, simple risk stratification is mainly based on LVEF. However, the initiation and maintenance of malignant arrhythmias are not only closely related to the complex pathophysiological mechanism of HF

but also influenced by many factors, including changes in autonomic nervous system activity, metabolic disturbances, electrolyte abnormalities, ischaemia, toxins, infectious agents, and cardiac or noncardiac drugs, which makes the predictive sensitivity of LVEF limited.[3] Vaduganathan *et al.* used multiple logistic regression to analyse over 30 baseline characteristics of 4024 HFrEF patients in the EVEREST trial to construct a 1 year SCD risk prediction model.[8] Male gender, black race, diabetes mellitus, and ACEI/ARB use were found to be potential predictors, but the discriminatory power of the model was poor (C-statistic 0.57). For the HF with preserved ejection fraction population, Adabag *et al.* used Cox regression to analyse the baseline characteristics of 4128 patients in the I-PRESERVE trial to construct a 5 year SCD risk prediction model.[5] Age, gender, diabetes mellitus, history of myocardial infarction, LBBB, and amino-terminal pro-brain natriuretic peptide were found to be independent predictors of SCD, and the model had moderate discriminatory power (C-statistic 0.75). Considering the higher risk of deaths from non-sudden causes in HF patients, Shen *et al.* also used the I-PRESERVE trial as the derivation cohort and adopted a competing risk regression model to predict the risk of sudden death.[9] Model performance was found to decline in the external validation of the CHARM-Preserved trial (C-statistic 0.68–0.69) and the TOPCAT trial (C-statistic 0.64–0.73). Because traditional regression methods are difficult to effectively handle high-dimensional interaction information in large data sets, this mechanistically limits the model's ability to predict complex relationships.

Machine learning techniques can overcome these limitations. In dealing with complex data relationships, ML does not need to assume the type of data distribution and the linear or non-linear relationship between features. Modelling by computationally intensive iterative algorithms rather than manually selecting features can help identify potential predictors and improve the predictive accuracy of the model. Currently, several prognostic prediction studies for hypertension, coronary heart disease, and HF have revealed that the XGBoost model and random forest model have higher predictive accuracy than traditional regression models.[13–17] These two ML models are based on the boosting and bagging of integrated algorithms to improve model performance by the concatenation of parallel weak classifiers into strong classifiers. In this study, both the XGBoost model and the random forest model could effectively predict malignant arrhythmia in hospitalized patients with HF, and the discriminatory power of the XGBoost model has potential advantages, while the random forest model is more robust. MARS is a combination algorithm of classical linear regression, spline function, and binary regression that can flexibly construct linear and non-linear models for high-dimensional data and simulate the interaction between features.[18] In this study, MARS was used to explore the complex relationship between HF and malignant arrhythmia, and the overall performance was

found to be close to that of the XGBoost model. CART is an algorithm based on the recursive binary splitting technique, which has been studied for risk stratification of in-hospital mortality in acute HF and SCD risk prediction in chronic HF.[19,20] In this study, the CART model consisting of five predictors had lower discriminatory power than the other models, but its tree structure was concise and comprehensible, which may facilitate rapid bedside evaluation of patients with HF.

The great potential of ML lies in the ability to achieve automated, real-time data updates to continuously 'teach' the model, thereby continuously improving its prediction accuracy. However, real-world applications require a sufficient trade-off between model accuracy, interpretability, and ease of use. In addition to clinical outcomes, both doctors and patients pay more attention to the substantial risk posed by the various predictors, which facilitates physician decision-making and patient education to proactively correct potentially reversible risk factors, thereby improving the prognosis of patients and optimizing the use of medical resources. In this study, we combined Lasso regularization with logistic regression and found that model 2 (consisting of only four predictors) had good accuracy, interpretability, and ease of use and could be used for the admission assessment of patients with HF. When the risk of malignant arrhythmias was ≤40%, model 1, consisting of nine predictors, conferred higher clinical benefit.

Oral antiarrhythmic drug, LBBB, Mg, D-dimer, and RBG were determined to be risk factors in half or more of the models, indicating their great value for the prediction of malignant arrhythmias. In the present study, 23% of patients were prescribed oral antiarrhythmic drugs during hospitalization, including amiodarone in 96% and mexiletine or propafenone in 4%. Antiarrhythmic drugs are well known to have potentially arrhythmogenic effects, and the result of this study is a reminder to standardize antiarrhythmic drug therapy in patients with HF. The result of previous studies have confirmed that LBBB and diabetes are closely associated with SCD,[5,8,21,22] and the risk for SCD increases when glucose tolerance is impaired.[23] This study further revealed that LBBB and RBG were independent predictors of malignant arrhythmia. The relationship between serum magnesium and SCD is still controversial. A retrospective study of CCU patients showed that hypermagnesemia was an independent predictor of increased in-hospital mortality, but serum magnesium levels were not associated with SCD.[24] However, our study revealed a significantly increased risk for malignant arrhythmias with increased serum magnesium levels. Differences in study conclusions may be attributed to different study populations, with 50.7% of patients with acute myocardial infarction in the CCU study. The results of a meta-analysis of seven prospective studies suggested that in HF patients, hypermagnesemia with serum magnesium ≥ 1.05 mmol/L was associated with an increased risk for cardiovascular mortality and all-cause

mortality.[25] In addition, patients with HF have a high risk of thromboembolism.[26] In this study, D-dimer and antithrombotic drugs were risk factors and protective factors for malignant arrhythmia, respectively, suggesting that the risk for thromboembolism might be related to malignant arrhythmia. Of course, this finding needs further verification in subsequent studies.

Limitations

The current study has the following limitations. First, there was an imbalance in the proportion of patients with and without events, which may limit the accuracy of model classification. Therefore, multiple ML classification algorithms were employed in this study to evaluate the discrimination and calibration of the model in the training set and validation set, respectively. Second, excluding features with data missing in ≥30% of the patients [glycosylated haemoglobin, postprandial glucose, C-reactive protein, troponin, and QT interval (atrial fibrillation, bundle branch block, or pacing rhythm affected accurate measurements)] may have affected the prediction accuracy of the model. Third, recent studies have revealed that dynamic changes in clinical characteristics influence SCD risk prediction.[20] Although only the baseline characteristics of the patients were considered in this study, the endpoint of this study was an in-hospital event rather than long-term prognosis; thus, the impact was expected to be limited. Finally, the clinical application of the ML model still needs further validation in an independent external cohort.

Conclusions

In conclusion, our study demonstrates that ML models based on the Lasso-logistic regression, MARS, random forest, and XGBoost algorithms can effectively predict the risk of malignant arrhythmias in hospitalized HF patients. The Lasso-logistic model had better clinical interpretability and ease of use than the other models. Oral antiarrhythmic drugs, LBBB, serum magnesium, D-dimer, and RBG were significant predictors of malignant arrhythmias.

Acknowledgements

Conflict of interest

None declared.

Funding

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Sensitivity analysis in multiple imputation for missing data.
Table S2. Clinical features used for model development.
Table S3. Optimal parameter settings for machine learning algorithms.
Table S4. Baseline characteristics of patients in the training and validation sets.
Figure S1. Lasso regularization for feature selection.
Figure S2. Feature importance plot for the Random Forest model (Top 15).
Figure S3. Feature importance plot for the XGBoost model (Top 15).

References

1. Hao G, Wang X, Chen Z, Zhang L, Zhang Y, Wei B, Zheng C, Kang Y, Jiang L, Zhu Z, Zhang J, Wang Z, Gao R, China Hypertension Survey Investigators. Prevalence of heart failure and left ventricular dysfunction in China: the China Hypertension Survey, 2012-2015. *Eur J Heart Fail* 2019; **21**: 1329–1337.

10. Adler ED, Voors AA, Klein L, Macheret F, Braun OO, Urey MA, Zhu W, Sama I, Tadel M, Campagnari C, Greenberg B, Yagil A. Improving risk prediction in heart failure using machine learning. *Eur J Heart Fail* 2020; **22**: 139–147.

11. Olsen CR, Mentz RJ, Anstrom KJ, Page D, Patel PA. Clinical applications of machine learning in the diagnosis, classification, and prediction of heart failure. *Am Heart J* 2020; **229**: 1–17.

12. Chinese Society of Cardiology of Chinese Medical Association. Editorial Board of Chinese Journal of cardiology. [Chinese guidelines for the diagnosis and treatment of heart failure 2014]. *Zhonghua Xin Xue Guan Bing Za Zhi* 2014; **42**: 98–122.

13. Tokodi M, Schwertner WR, Kovács A, Tősér Z, Staub L, Sárkány A, Lakatos BK, Behon A, Boros AM, Perge P, Kutyifa

V, Széplaki G, Gellér L, Merkely B, Kosztin A. Machine learning-based mortality prediction of patients undergoing cardiac resynchronization therapy: the SEMMELWEIS-CRT score. *Eur Heart J* 2020; **41**: 1747–1756.

14. Wu X, Yuan X, Wang W, Liu K, Qin Y, Sun X, Ma W, Zou Y, Zhang H, Zhou X, Wu H, Jiang X, Cai J, Chang W, Zhou S, Song L. Value of a machine learning approach for predicting clinical outcomes in young patients with hypertension. *Hypertension* 2020; **75**: 1271–1278.

15. Angraal S, Mortazavi BJ, Gupta A, Khera R, Ahmad T, Desai NR, Jacoby DL, Masoudi FA, Spertus JA, Krumholz HM. Machine learning prediction of mortality and hospitalization in heart failure with preserved ejection fraction. *JACC Heart Fail* 2020; **8**: 12–21.

16. Sax DR, Mark DG, Huang J, Sofrygin O, Rana JS, Collins SP, Storrow AB, Liu D, Reed ME. Use of machine learning to develop a risk-stratification tool for emergency department patients with acute heart failure. *Ann Emerg Med* 2021; **77**: 237–248.

17. Zack CJ, Senecal C, Kinar Y, Metzger Y, Bar-Sinai Y, Widmer RJ, Lennon R, Singh M, Bell MR, Lerman A, Gulati R. Leveraging machine learning techniques to forecast patient prognosis after percutaneous coronary intervention. *JACC Cardiovasc Interv* 2019; **12**: 1304–1311.

18. Friedman JH. Multivariate adaptive regression splines. *Ann Stat* 1991; **19**: 1–67.

19. Fonarow GC, Adams KF Jr, Abraham WT, Yancy CW, Boscardin WJ, ADHERE Scientific Advisory Committee, Study Group, and Investigators. Risk stratification for in-hospital mortality in acutely decompensated heart failure: classification and regression tree analysis. *JAMA* 2005; **293**: 572–580.

2. Zhang Y, Zhang J, Butler J, Yang X, Xie P, Guo D, Wei T, Yu J, Wu Z, Gao Y, Han X, Zhang X, Wen S, Anker SD, Filippatos G, Fonarow GC, Gan T, Zhang R. Contemporary epidemiology, management, and outcomes of patients hospitalized for heart failure in China: results from the China Heart Failure (China-HF) registry. *J Card Fail* 2017; **23**: 868–875.

20. Rohde LE, Vaduganathan M, Claggett BL, Polanczyk CA, Dorbala P, Packer M, Desai AS, Zile M, Rouleau J, Swedberg K, Lefkowitz M, Shi V, McMurray JJV,

Solomon SD. Dynamic changes in cardiovascular and systemic parameters prior to sudden cardiac death in heart failure with reduced ejection fraction: a PARADIGM-HF analysis. *Eur J Heart Fail* 2021; **23**: 1346–1356.

21. Kristensen SL, Castagno D, Shen L, Jhund PS, Docherty KF, Rørth R, Abraham WT, Desai AS, Dickstein K, Rouleau JL, Zile MR, Swedberg K, Packer M, Solomon SD, Køber L, McMurray JJV, PARADIGM-HF and AT-MOSPHERE Committees and Investigators. Prevalence and incidence of intra-ventricular conduction delays and outcomes in patients with heart failure and reduced ejection fraction: insights from PARADIGM-HF and ATMO-SPHERE. *Eur J Heart Fail* 2020; **22**: 2370–2379.

22. Ferrini M, Johansson I, Aboyans V. Heart failure and its complications in patients with diabetes: mounting evidence for a growing burden. *Eur J Prev Cardiol* 2019; **26**: 106–113.

23. Eranti A, Kerola T, Aro AL, Tikkanen JT, Rissanen HA, Anttonen O, Junttila MJ, Knekt P, Huikuri HV. Diabetes, glucose tolerance, and the risk of sudden cardiac death. *BMC Cardiovasc Disord* 2016; **16**: 51.

24. Naksuk N, Hu T, Krittanawong C, Thongprayoon C, Sharma S, Park JY, Rosenbaum AN, Gaba P, Killu AM, Sugrue AM, Peeraphatdit T, Herasevich V, Bell MR, Brady PA, Kapa S, Asirvatham SJ. Association of Serum Magnesium on mortality in patients admitted to the intensive cardiac care unit. *Am J Med* 2017; **130**: 229.e5–229.e13.

25. Angkananard T, Anothaisintawee T, Eursiriwan S, Gorelik O, McEvoy M, Attia J, Thakkinstian A. The association of serum magnesium and mortality outcomes in heart failure patients: a systematic review and meta-analysis. *Medicine (Baltimore)* 2016; **95**: e5406.

26. Zorlu A, Yilmaz MB, Yucel H, Bektasoglu G, Refiker Ege M, Tandogan I. Increased d-dimer levels predict cardiovascular mortality in patients with systolic heart failure. *J Thromb Thrombolysis* 2012; **33**: 322–328.

3. Goldberger JJ, Cain ME, Hohnloser SH, Kadish AH, Knight BP, Lauer MS, Maron BJ, Page RL, Passman RS, Siscovick D, Stevenson WG, Zipes DP, American Heart Association; American College of Cardiology Foundation, Heart Rhythm

Society, American Heart Association/American College of Cardiology Foundation/Heart Rhythm Society Scientific Statement on Noninvasive Risk Stratification Techniques for Identifying Patients at Risk for Sudden Cardiac Death. A scientific statement from the American Heart Association Council on Clinical Cardiology Committee on Electrocardiography and Arrhythmias and Council on Epidemiology and Prevention. *J Am Coll Cardiol* 2008; **52**: 1179–1199.

4. Saxon LA, Bristow MR, Boehmer J, Krueger S, Kass DA, De Marco T, Carson P, DiCarlo L, Feldman AM, Galle E, Ecklund F. Predictors of sudden cardiac death and appropriate shock in the comparison of medical therapy, pacing, and defibrillation in heart failure (COMPANION) trial. *Circulation* 2006; **114**: 2766–2772.

5. Adabag S, Rector TS, Anand IS, McMurray JJ, Zile M, Komajda M, McKelvie RS, Massie B, Carson PE. A prediction model for sudden cardiac death in patients with heart failure and preserved ejection fraction. *Eur J Heart Fail* 2014; **16**: 1175–1182.

6. Gottdiener JS, Fohner AE. Risk prediction in heart failure: new methods, old problems. *JACC Heart Fail* 2020; **8**: 22–24.

7. Levy WC, Mozaffarian D, Linker DT, Sutradhar SC, Anker SD, Cropp AB, Anand I, Maggioni A, Burton P, Sullivan MD, Pitt B, Poole-Wilson PA, Mann DL, Packer M. The Seattle heart failure model: prediction of survival in heart failure. *Circulation* 2006; **113**: 1424–1433.

8. Vaduganathan M, Patel RB, Mentz RJ, Subacius H, Chatterjee NA, Greene SJ, Ambrosy AP, Maggioni AP, Udelson JE, Swedberg K, Konstam MA, O'Connor CM, Butler J, Gheorghiade M, Zannad F. Sudden death after hospitalization for heart failure with reduced ejection fraction (from the EVEREST Trial). *Am J Cardiol* 2018; **122**: 255–260.

9. Shen L, Jhund PS, Anand IS, Carson PE, Desai AS, Granger CB, Køber L, Komajda M, McKelvie RS, Pfeffer MA, Solomon SD, Swedberg K, Zile MR, McMurray JJV. Developing and validating models to predict sudden death and pump failure death in patients with heart failure and preserved ejection fraction. *Clin Res Cardiol* 2021; **110**: 1234–1248.