

RESEARCH

Open Access



A machine learning-based model for predicting paroxysmal and persistent atrial fibrillation based on EHR

Yuqi Zhang^{1,3†}, Sijin Li^{2,4†}, Peibiao Mai^{5,6†}, Yanqi Yang^{4,7}, Niansang Luo⁶, Chao Tong^{1,3*}, Kuan Zeng^{4*} and Kun Zhang^{6,8*}

Abstract

Background There is no effective way to accurately predict paroxysmal and persistent atrial fibrillation (AF) subtypes unless electrocardiogram (ECG) observation is obtained. We aim to develop a predictive model using a machine learning algorithm for identification of paroxysmal and persistent AF, and investigate the influencing factors.

Methods We collected demographic data, medication use, serological indicators, and baseline cardiac ultrasound data of all included subjects, totaling 50 variables. The diagnosis of AF subtypes is confirmed by ECG observation for at least more than 7 days. Variable selection was performed by spearman correlation analysis, recursive feature elimination, and least absolute shrinkage and selection operator regression. We built a prediction model for AF using three machine learning methods. Finally, the significance of each variable was analyzed by Shapley additive explanations method.

Results After screening, we found the optimal variable set consisting of 10 variables. The model we built achieved good predictive performance (AUC = 0.870, 95%CI 0.858 to 0.882), and had specificity of 0.851 (95%CI 0.844 to 0.858) and sensitivity of 0.716 (95%CI 0.676 to 0.755). Good predictive performance was stably achieved in different age subgroups and different gender subgroups. LA and NT-proBNP were the two most important variables for predicting paroxysmal and persistent AF in all models, except for the female subgroup aged less than 60 years.

Conclusions Our model makes it possible to predict paroxysmal and persistent AF based on baseline data at admission. Early and individualized intervention strategies based on our model may help to improve clinical outcomes in AF patients.

Keywords Atrial fibrillation, Paroxysmal atrial fibrillation, Persistent atrial fibrillation, Machine learning, Prediction model

[†]Yuqi Zhang, Sijin Li and Peibiao Mai contribute equally to the development of this research study.

*Correspondence:

Chao Tong
tongchao@buaa.edu.cn
Kuan Zeng
zengkuan3@mail.sysu.edu.cn
Kun Zhang
zhangk65@mail.sysu.edu.cn

Full list of author information is available at the end of the article



Background

Atrial fibrillation (AF) is the most common persistent cardiac arrhythmia, and its incidence and prevalence are increasing worldwide [1, 2]. More than 50 million people worldwide were affected by AF in 2020, and the prevalence is expected to continue to rise in the future [3, 4]. Therefore, AF, as a major public health issue, has become a huge burden on social and economic development. And this puts forward new requirements for early screening and intervention of AF.

The recent US 2023 Atrial Fibrillation Diagnosis and Management Guidelines divide AF into four stages, including At Risk for AF, Pre-AF, AF, and Permanent AF [4]. The first stage (At Risk for AF) and the second stage (Pre-AF) were first proposed in the guidelines. While enhanced monitoring is recommended in the second stage of AF (Pre-AF), and treatment symptoms are needed for the third stage (AF) and the fourth stage of AF (Permanent AF) according to the latest version of the diagnosis and treatment guidelines for AF [5]. Since current clinical studies have not obtained a positive correlation between opportunistic screening for AF and increased detection rates of AF [6, 7], screening for the first and second stages is not necessary. The fourth stage of AF in the guidelines refers to permanent AF, and no subtype prediction is required. Therefore, the prediction of the third stage of AF is more meaningful. The third stage of AF is further divided into Paroxysmal AF (AF that is intermittent and terminates within ≤ 7 d of onset), Persistent AF (AF that is continuous and sustains for > 7 d and requires intervention), Long-standing persistent AF (AF that is continuous for > 12 months in duration), and Successful AF ablation (Freedom from AF after percutaneous or surgical intervention to eliminate AF) [4]. Among them, paroxysmal and persistent AF patients account for the majority of outpatient AF types in China, accounting for 38.9% and 39.2% respectively [8]. The research conducted on these two subtypes of AF is of great significance in the actual clinical application scenarios in China. The treatment strategies for the four subtypes of stage 3 AF are also different, especially in paroxysmal AF and persistent AF [4, 5]. It is similar to the European guidelines, where catheter ablation is graded as Class I and Class IIb in the management of paroxysmal AF and persistent AF, respectively [9, 10]. Therefore, the prediction for these two subtypes is more meaningful.

A recent review showed that the theoretical burden of AF in patients with non-paroxysmal AF and spontaneous regression is almost 100% [11]. Occasional regression of AF results in an estimated burden of AF of 70–100% in patients with persistent AF [12]. Nevertheless, the burden of AF in persistent AF is about 10 times higher than that in paroxysmal AF. The risk of stroke in patients with persistent or permanent AF is also higher than

that in patients with paroxysmal AF [13, 14]. The result shows that yearly ischemic stroke rates were 2.1% and 3.0% for paroxysmal and persistent AF, respectively [5]. Early identify the type of AF (paroxysmal or persistent) in patients with a new diagnosis of AF is of great significance for the staged management of AF. In the 2024 EHRA guidelines, it specifically emphasize the need to recognize the common misclassification between paroxysmal AF and persistent AF in clinical work [10]. The 2024 ESC guidelines also emphasize that paroxysmal AF and persistent AF are not easy to distinguish [9]. Therefore, more strategies are needed to predict paroxysmal and persistent AF.

Machine learning (ML) is an important artificial intelligence method that uses complex algorithms to discover potential patterns in some massive data sets [15]. With the increasing acceptance of clinicians, ML has been applied to many clinical medical fields including cardiovascular and cerebrovascular diseases [16, 17]. We included all patients with a first diagnosis of AF but lack of specific type and collected their baseline data. Since the diagnosis of paroxysmal and persistent AF requires at least 7 days of observation, we excluded patients with a hospital stay of less than 7 days. We used ML methods to build a model to predict the diagnosis of AF at discharge (Fig. 1-a). We can not only screen out variables that are highly correlated with the clinical classification of AF but also provide an application basis for predicting paroxysmal or persistent AF.

Methods

Patient enrolment and data collection

We enrolled participants with the inclusion criteria: (1) Patients hospitalized in Sun Yat-sen Memorial Hospital from January 2013 to January 2023. (2) Patients with examination records confirming the presence of AF rhythm in the past or during hospitalization (such as surface electrocardiogram, 24-hour dynamic electrocardiogram, and pacemaker memory record). All the patients had a first diagnosis of AF with an unknown specific type. (3) Patients with a discharge diagnosis of atrial fibrillation and the subtype is paroxysmal or persistent. The discharge diagnosis code starts with I48 (AF) and the subtype is I48.x02 (paroxysmal AF) or I48.x00x007 (persistent AF), according to the diagnosis code of the 10th edition of the International Classification of Diseases (ICD-10). Paroxysmal AF is defined as AF that can terminate spontaneously or after intervention within 7 days. Persistent AF is defined as AF that lasts for more than 7 days and requires medication or electrical cardioversion to terminate the attack. (4) Detailed medical history information.

Our exclusion criteria include (1) Patients with rheumatic heart disease, congenital heart disease, primary

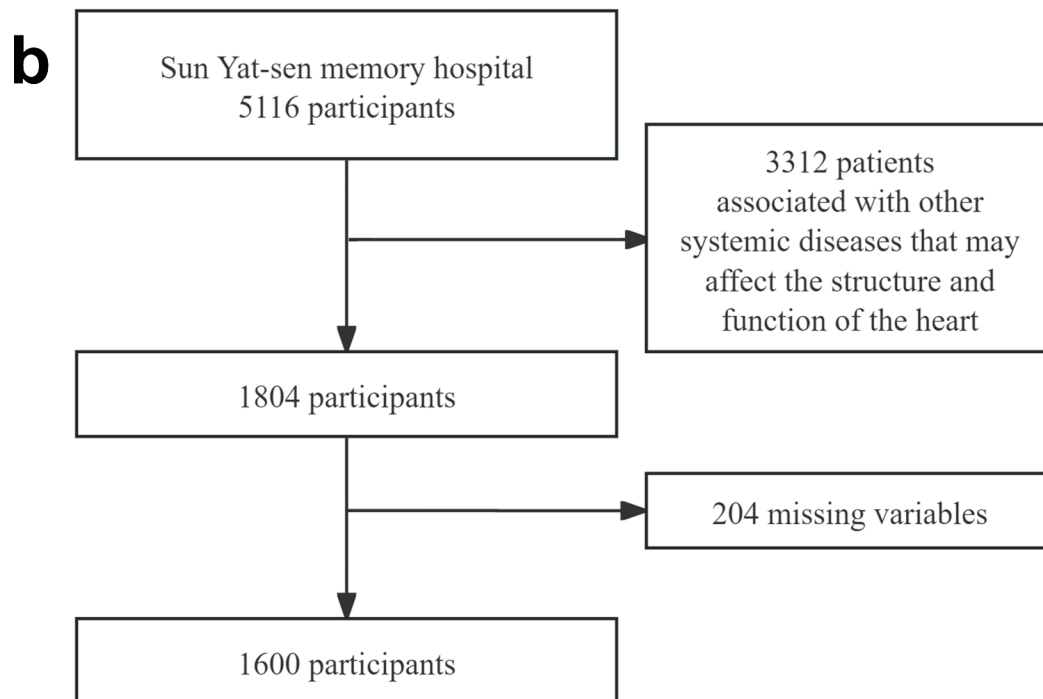
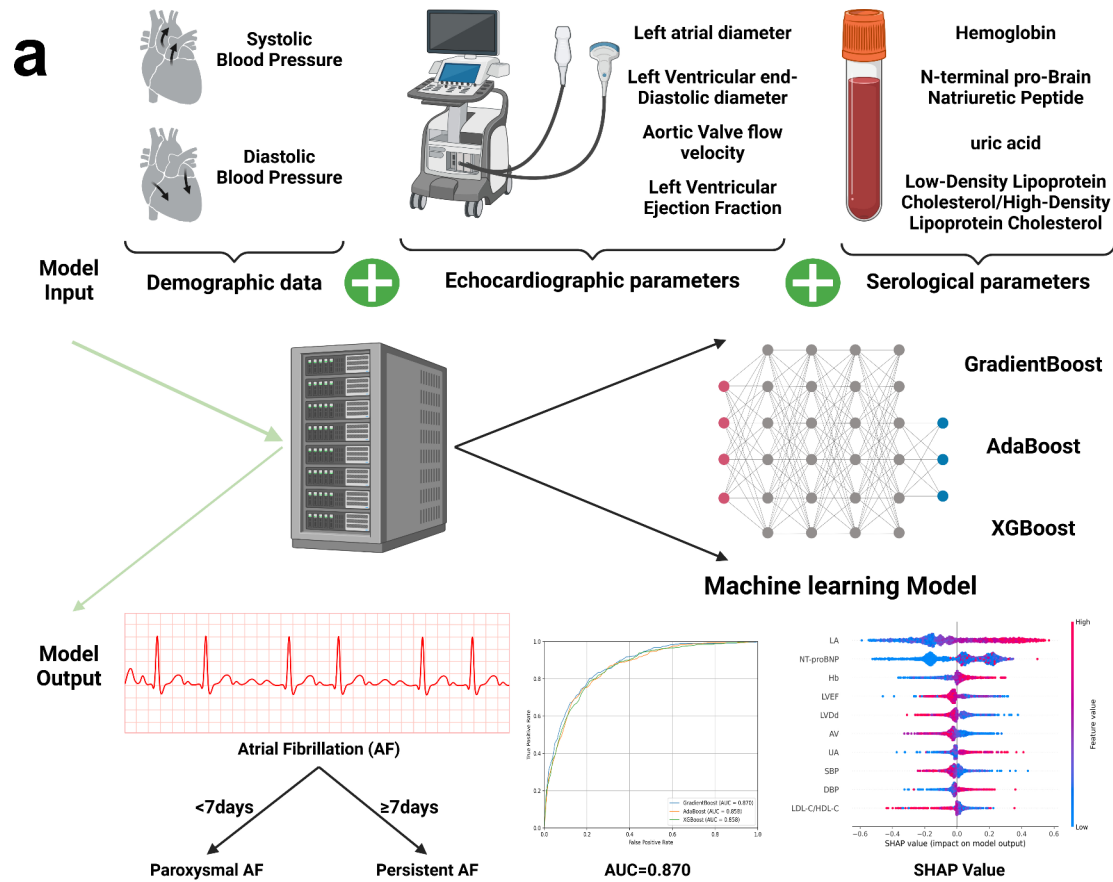


Fig. 1 Central illustration and flow chart of the study design. This figure contains the central illustration and the flow chart. Figure 1-a is a summary of the entire research, explaining the input variables and output results in the model, and the clinical significance of our study. Figure 1-b is the flow chart of this study. After enrolled all the patients matched inclusion criteria, we did the screening according to the exclusion criteria and finally included 1,600 participants

valvular heart disease, cardiomyopathy, pericardial disease, cor pulmonale, malignant tumors, or recent major surgical history; (2) Patients with other systemic diseases that may affect cardiac structure and function, such as acute myocardial infarction, hyperthyroidism, hypothyroidism, amyloidosis, pheochromocytoma, systemic lupus erythematosus, or severe infection; (3) Patients with severe liver dysfunction: alanine aminotransferase > 3 times the upper limit of normal and (alanine aminotransferase/aspartate aminotransferase) ratio > 1; (5) Patients with severe renal dysfunction: estimated glomerular filtration rate < 30 ml/min·1.73m². (6) Patients with severe lack of clinical data, which refers to patients who lack more than half of the variable results or lack records of cardiac ultrasound examination results.

The methods complied with the ethical principles of the Declaration of Helsinki. This study was reviewed and approved by the Ethics Committees of the Sun Yat-sen Memorial Hospital of Sun Yat-sen University (SYSKY-2024-004-01).

Selection of variables

We performed correlation analysis on all recorded variables, and all the variables are detailed in Supplementary Table 1. By using Spearman correlation analysis, we preliminarily screened variables with a strong correlation with a diagnosis of AF stage. Among them, $p < 0.001$ was considered a statistically significant difference. Subsequently, we performed the least absolute shrinkage and selection operator (LASSO) regression correlation analysis and GradientBoost Recursive Feature Elimination (RFE) on this basis to complete feature screening. By combining these two methods, we can comprehensively and accurately evaluate the importance of variables, providing strong support for subsequent model construction.

We collected demographic data, medication use, serological indicators, and baseline cardiac ultrasound data of all included subjects, totaling 50 variables. Since the diagnosis of AF subtypes must rely on ECG, we collected the patients' ECG diagnosis results. But we did not use any parameters in the ECG as input variables of the model. After screening, we finally included 10 variables that can be divided into three categories: demographic data, cardiac ultrasound, and serological indicators. All data were collected from electronic health records (EHR). Demographic data included systolic blood pressure (SBP) and diastolic blood pressure (DBP). Echocardiographic parameters included left atrial diameter (LA), left ventricular end-diastolic diameter (LVDD), aortic valve flow velocity (AV), and left ventricular ejection fraction (LVEF). These indicators were analyzed by routine transthoracic echocardiography (TTE) performed by a certified cardiologist at baseline and collected from the EHR.

Serological parameters included hemoglobin (Hb), N-terminal pro-brain natriuretic peptide (NT-proBNP), uric acid (UA), and the ratio of low-density lipoprotein cholesterol to high-density lipoprotein cholesterol (LDL-C/HDL-C). All serological indicators were obtained from the peripheral blood sample collected for the first time at baseline.

Machine learning algorithms

In this investigation, alongside traditional statistical analyses, we conducted experiments utilizing three widely employed gradient boosting machine algorithms: AdaBoost, GradientBoost, and XGBoost. Each algorithm showcased distinctive strengths, contributing valuable diversity to our study. AdaBoost, short for Adaptive Boosting, is a pioneering algorithm in the boosting family that combines multiple weak learners, typically decision trees, to create a strong classifier. The key principle of AdaBoost is to iteratively adjust the weights of misclassified instances, enabling the model to focus on difficult cases in subsequent iterations [18]. GradientBoost, a classical gradient boosting algorithm, excels when dealing with high-dimensional sparse datasets. It iteratively minimizes the loss function by training decision trees, effectively managing various complex nonlinear relationships [19]. XGBoost, an efficient algorithm rooted in gradient boosting trees, is lauded for its outstanding performance and scalability. Leveraging parallel processing enhances training speed and efficiently handles large-scale datasets [20]. Employing a range of machine learning models enabled us to comprehensively assess their performance on our datasets, providing insights into which models excel in addressing specific problems. We conducted experiments using five-fold cross-validation to comprehensively evaluate their performance on the dataset.

SHAP interpretable analysis for machine learning

SHAP (SHapley Additive exPlanations) [21] stands as an interpretable method grounded in game theory, facilitating a nuanced understanding of each feature's contribution to the model's output. By quantifying the impact of individual features, SHAP empowers us to discern which features play pivotal roles in driving variations in the model's output. This interpretative capability aids researchers in identifying crucial features, leading to a more profound understanding of patterns and regularities within the data. Through the application of SHAP analysis, we were able to visualize the contribution of variables to the model's predictive results, thereby highlighting the key features of the model with clarity and precision.

Statistical analysis

The Kolmogorov-Smirnov test was used to assess the normality of the distribution of continuous variables; normal variables were presented as mean \pm standard deviation (SD), while nonnormal variables were presented as median and interquartile range (IQR). Categorical variables were presented as frequency and percentage. The receiver operating characteristic (ROC) curve and the area under the ROC curve (AUC) were used to measure the discriminative ability of the ML model for predicting paroxysmal and persistent AF. Further analysis was performed to examine the predictive ability of the ML model. The analysis included sensitivity (SEN), specificity (SPE), accuracy (ACC), precision (PRE), recall, and F1 score. Displays restricted cubic spline (RCS) curves with 4 knots to test nonlinear relationships between independent variables and outcomes [22, 23]. SPSS Statistics Version 26.0 and Python 3.7.6 software were used for statistical analysis and graphics, and $p < 0.05$ was considered statistically significant.

Results

Baseline characteristics of participants

According to the strict inclusion and exclusion criteria, 1600 patients at Sun Yat-sen Memorial Hospital of Sun Yat-sen University were enrolled as research subjects. Among them, there were 1,020 cases of paroxysmal AF and 580 cases of persistent AF. The baseline data of the recruits included in this study are shown in Table 1. By comparing the baseline data between the paroxysmal AF and persistent AF groups, it can be found that the differences exist in gender, DBP, Hb, UA, NT-proBNP, LA, LVDd, AV, and LVEF ($p < 0.05$). The male proportion, DBP, Hb, UA, NT-proBNP, and LA in the persistent AF group were higher than those in the paroxysmal AF

group. There was no statistically significant difference in SBP and LDL-C/HDL-C between the two groups.

Variable selection results

We further used Spearman correlation analysis ($p < 0.001$) to screen out variables with a strong correlation with the type of AF diagnosis. For these variables, we used GradientBoost-RFE and LASSO for dimensionality reduction. In the LASSO dimensionality reduction method, when the number of included variables was 32, the highest AUC value could be obtained (Fig. 2-a). Similarly, after dimensionality reduction using the RFE method, we retained 12 variables to obtain the best performance (Fig. 2-b). Then we took the intersection of these two parts of variables and determined 10 variables that were finally included in the model (Fig. 2-c). The correlation between the final included variables and the subtypes of AF was plotted using a heat map (Fig. 2-d).

Results of the AF prediction model

After incorporating 10 variables, we used three methods (GradientBoost, AdaBoost, and XGBoost) to build models. The outputs of the models were the patient's atrial fibrillation subtype (specifically paroxysmal or persistent AF) after a hospitalization of more than 7 days, which was compared with the diagnosis recorded in the EHR at the time of final discharge. By comparing the evaluation indicators, we found that the model established in this study had good performance, whether in AUC, PRE, SEN, or SPE (as shown in Table 2). Among the three machine learning methods used in modeling, GradientBoost performed well in most indicators. Specifically, in the AUC indicator, GradientBoost achieved the best value [0.870, 95% confidence interval (CI): 0.858 to 0.882]. AdaBoost and XGBoost were 0.858 (95% CI: 0.836 to 0.877) and

Table 1 Comparison of general information among paroxysmal and persistent AF

Variables	Total (n = 1600)	Paroxysmal AF (n = 1020)	Persistent AF (n = 580)	p-value
Gender (n, %)				0.021*
Male (%)	944 (59.0%)	580 (56.9%)	364 (62.8%)	
Female (%)	656 (41.00)	440 (43.14)	216 (37.24)	
SBP (mmHg)	127 (116, 142)	128 (116, 143)	126 (115, 140)	0.362
DBP (mmHg)	78 (71, 85)	77 (70, 84)	80 (72, 88)	<0.001*
LDL-C/HDL-C	2.52 (1.94, 3.20)	2.52 (1.95, 3.25)	2.51 (1.93, 3.12)	0.431
Hb (g/L)	135 (123, 147)	133 (120, 144)	139 (126, 150)	<0.001*
UA (μ mol/L)	391 (321, 468)	374 (306, 444)	431 (352, 510)	<0.001*
NT-proBNP (pg/ml)	471 (129, 1192)	245 (75, 788)	899 (472, 1582)	<0.001*
LA (mm)	38 (35, 43)	36 (33, 39)	43 (39, 46)	0.031*
LVDd (mm)	48 (45, 51)	48 (45, 51)	48 (45, 52)	0.013*
AV (cm/s)	129 (110, 152)	133 (115, 156)	121 (103, 143)	<0.001*
LVEF (%)	66 (61, 70)	67 (63, 71)	63 (59, 68)	<0.001*

Values are presented as n (%) as appropriate or the median [interquartile range (IQR)]. AF, Atrial fibrillation; AV, Aortic opening flow velocity; DBP, Diastolic blood pressure; SBP, Systolic blood pressure; Hb, Hemoglobin; HDL-C, High density lipoprotein cholesterol; LA, Left atrial diameter; LDL-C, Low density lipoprotein cholesterol; LVDd, Left ventricular end-diastolic diameter; LVEF, Left ventricular ejection fraction; NT-proBNP, N-terminal brain natriuretic peptide precursor; UA, Uric acid

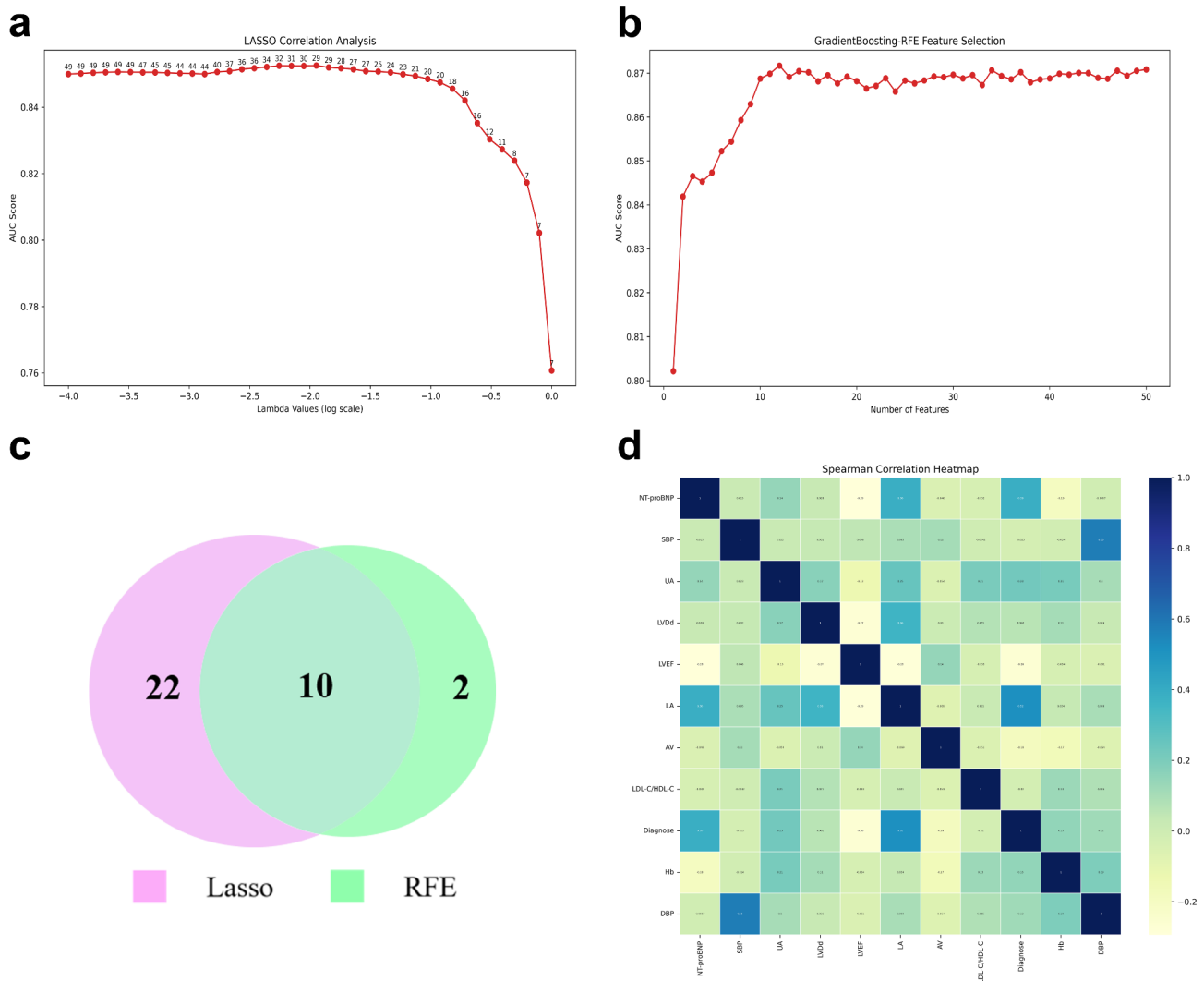


Fig. 2 Selection process of variables. Figure 2 shows the process of variable screening. Figure 2-a and -b correspond to the results of variable screening using LASSO and GradientBoost RFE methods, respectively. The AUC of the model output changes with the change of the model input variables. From Fig. 2-c, it can be found that the number of variables when LASSO and GradientBoost RFE achieve the best AUC is 32 and 12 respectively, and merging them can get a common 10 variables. Figure 2-d illustrates the spearman correlation between the 10 variables

Table 2 Results of model output indicators in all participants

Model	ACC (95%CI)	Precision	Recall	AUC (95%CI)	F1 Score (95%CI)	SEN (95%CI)	SPE (95%CI)
GradientBoost	0.801 (0.775–0.821)	0.730	0.716	0.870 (0.858–0.882)	0.722 (0.686–0.754)	0.716 (0.676–0.755)	0.849 (0.816–0.872)
AdaBoost	0.796 (0.785–0.817)	0.727	0.700	0.858 (0.836–0.877)	0.713 (0.694–0.747)	0.700 (0.673–0.747)	0.851 (0.844–0.858)
XGBoost	0.782 (0.764–0.802)	0.701	0.697	0.858 (0.849–0.872)	0.699 (0.665–0.732)	0.697 (0.648–0.747)	0.831 (0.815–0.843)

ACC, Accuracy; AUC, Area under curve; CI, Confidence interval; SEN, Sensitivity; SPE, Specificity

0.859 (95% CI: 0.849 to 0.872), respectively. The ROC of the AF subtypes prediction model constructed using all participants is shown in Fig. 3-a. The best parameter for evaluating our algorithm in daily clinical practice is precision (positive predictive value). For this indicator, the results of GradientBoost, AdaBoost, and XGBoost were 0.730, 0.727, and 0.701, respectively. In terms of prediction accuracy, GradientBoost outperformed AdaBoost and XGBoost. The results were 0.801(95%CI: 0.775 to

0.821), 0.796(95%CI: 0.785 to 0.817), and 0.782 (95%CI: 0.764 to 0.802), respectively. We have a high SPE with the best AdaBoost algorithm is 0.851 (95% CI: 0.844–0.858). As for the SEN, the best value of 0.716 (95%CI: 0.676 to 0.755) is found in the GradientBoost method.

Interpretation of binary classification model

Since it is difficult for clinicians to accept a prediction model that cannot be directly explained and interpreted,

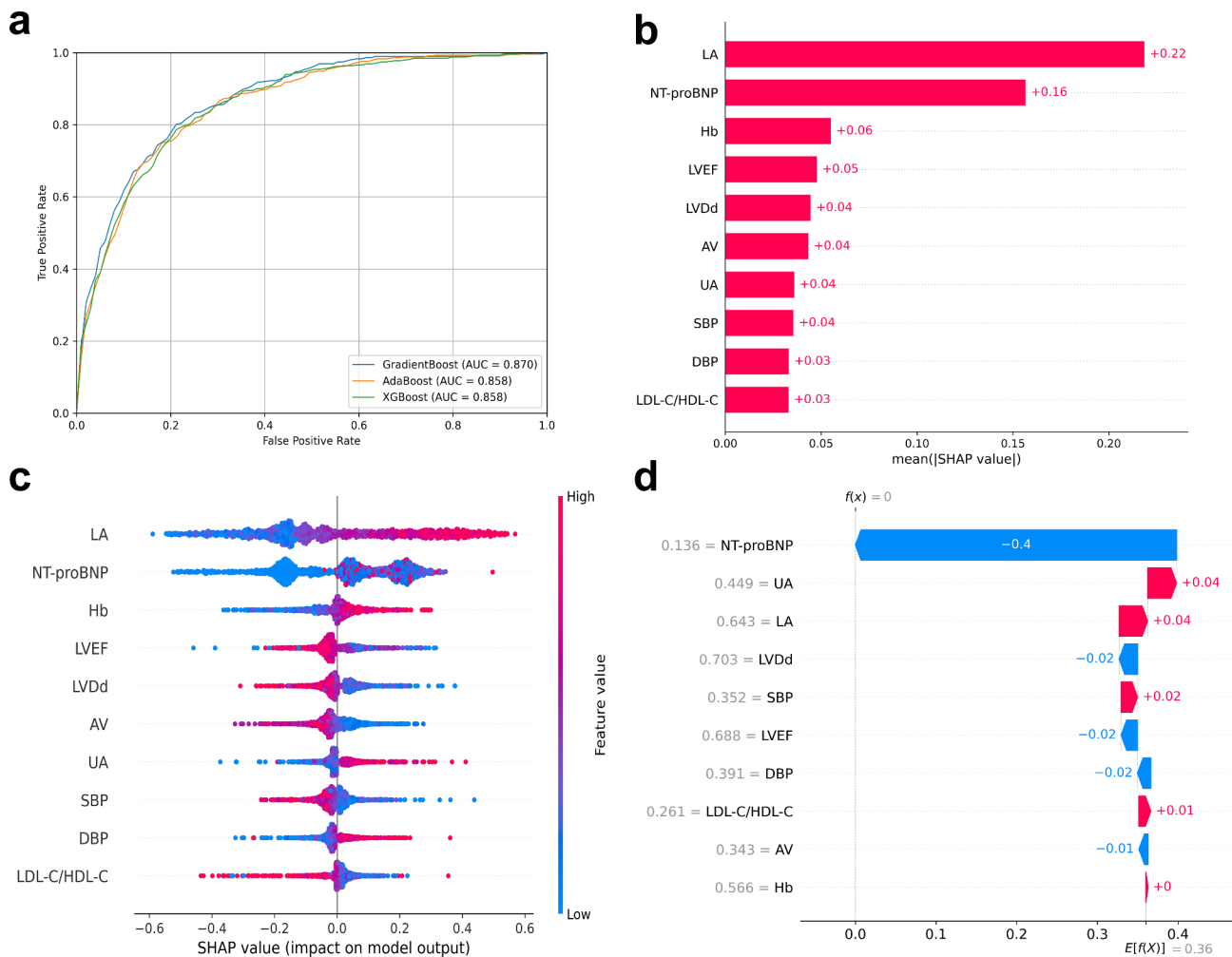


Fig. 3 Model prediction performance by AUC and interpretation by SHAP method. Figure 3 shows the predictive performance of the model and its interpretation using the SHAP method. Figure 3-a is the ROC obtained using three machine learning algorithms and their corresponding AUC. Figure 3-b shows the variables sorted by the absolute value of the mean SHAP value. A high value means a high impact on the model output. Figure 3-c visualizes the different SHAP values of each variable and their impact on the results. The gradient from red to blue indicates the ranking of the variable values from high to low. The horizontal axis indicates the impact on the result. Figure 3-d further shows the application of the SHAP algorithm in a single sample. Different variables with different SHAP values jointly affect the prediction results of the sample

the SHAP method is used to interpret the output of the final model by calculating the contribution of each variable to the prediction. As shown in Fig. 3-b, the SHAP mean evaluates the contribution of the factors to the model and is displayed in descending order. The four variables that have the greatest impact on different diagnostic types of AF are LA, NT-proBNP, Hb, and LVEF. Figure 3-c more intuitively observes the correspondence between different variables and the diagnostic type of AF. It can be observed that the top four variables affect the diagnosis of AF in a certain pattern. For example, LA shows a gradient line from red to blue. In particular, near the SHAP value of 0, there is a clear color boundary, indicating that there is an exploitable pattern between the LA value and the diagnostic type of AF. When the LA value is low, the model tends to predict paroxysmal AF, while

a larger LA value is associated with the diagnosis of persistent AF. As shown in Fig. 3-c, a larger LA, higher NT-proBNP, higher Hb, and lower LVEF are the key factors that lead the model to predict that the patient’s AF type is persistent AF. Figure 3-d specifically shows the impact of each variable obtained by the SHAP method on the model prediction results of a certain sample.

Although the potential pattern between the variables and the diagnosis of AF can be observed in Fig. 3-c, the lack of interaction between different variables is obvious, which is crucial for multi-factor prediction. Therefore, we further explored the interaction between the variables with the top four SHAP mean values (LA, NT-proBNP, Hb, and LVEF) as shown in Fig. 4. In Fig. 4-a ~ -d, we used the values of LA, NT-proBNP, Hb, and LVEF and their SHAP values as the x-axis and y-axis respectively.

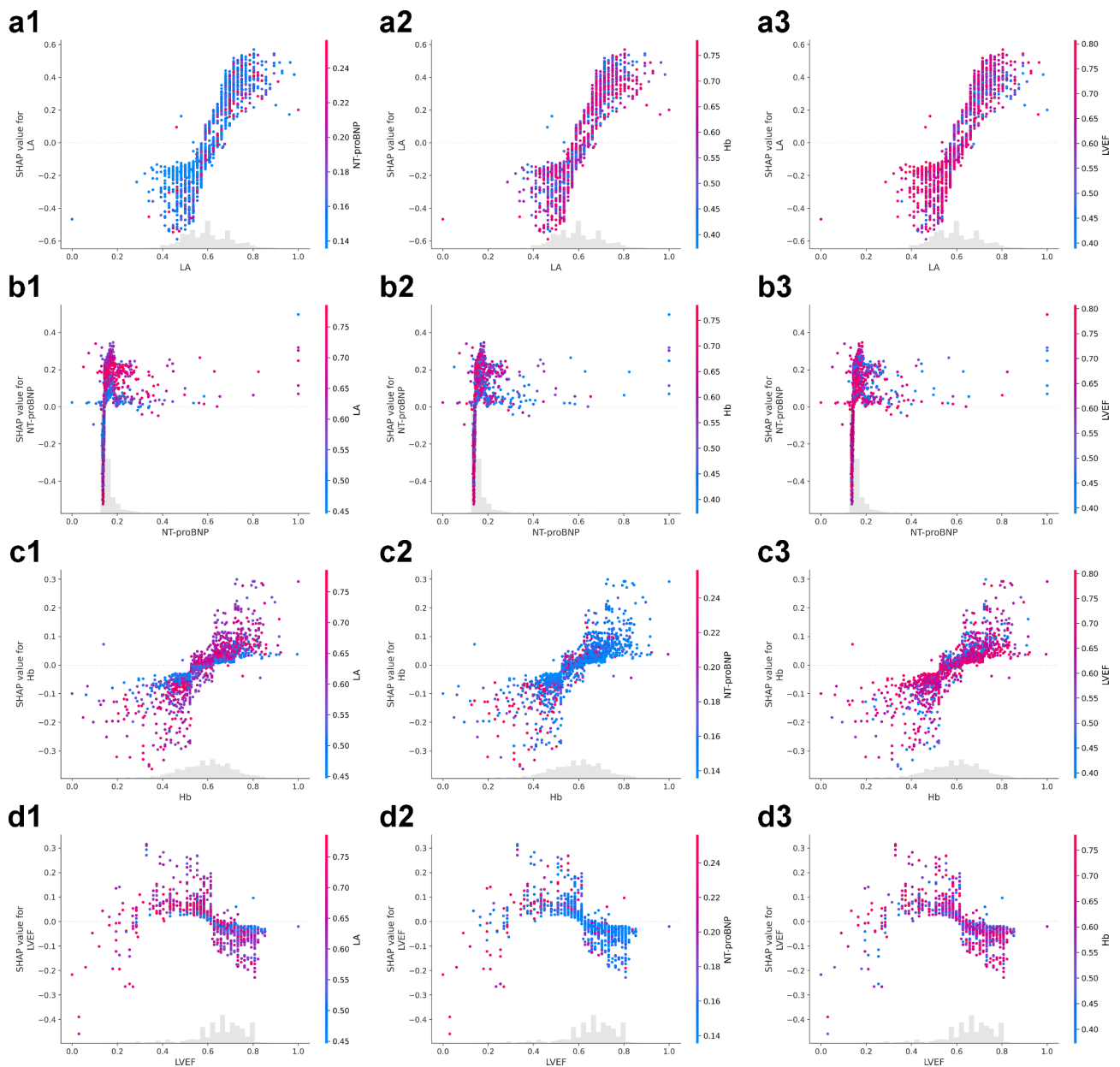


Fig. 4 Interactions between the top four important variables ranked by SHAP method. Figure 4 shows the interactions among the four variables (LA, NT-proBNP, Hb, and LVEF), which are the top four factors affecting the model output obtained by the SHAP method. The horizontal axis of Fig. 4-a1 ~ a3 is the value of LA, and the vertical axis is the SHAP value of LA. The color of the scatter points from red to blue corresponds to the value of the interaction variable, and the interaction variables in Fig. 4-a1 ~ 4-a3 are NT-proBNP, Hb, and LVEF. Similarly, Fig. 4-b1 ~ 4-b3 represents the interaction between NT-proBNP and LA, Hb, and LVEF; Fig. 4-c1 ~ 4-c3 is the interaction between Hb and LA, NT-proBNP, and LVEF; d1 ~ d3 is the interaction between LVEF and LA, NT-proBNP, and Hb

Then remove the variables used to draw the horizontal and vertical coordinates from LA, NT-proBNP, Hb, and LVEF. The remaining three variables are plotted as series 1 ~ 3 in Fig. 4 in the order of SHAP means. For example, in Fig. 4-a, LA is used to draw the horizontal and vertical coordinates, and the remaining NT-proBNP, Hb, and LVEF are used to draw a1, a2, and a3, respectively. Similarly, NT-proBNP is used to draw the horizontal and vertical coordinates, and the remaining LA, Hb, and LVEF

correspond to Fig. 4-b1, b2, and b3. The closer the scatter point is to the zero value of the vertical axis, it means that the sample corresponding to the point has a lower SHAP value in the variable represented by the vertical axis. The color of the scatter points in the figure from blue to red represents the value of the variable from small to large. By comparison, we found the corresponding relationship between the variables. When the patient's LA is a certain value, it corresponds to higher Hb and LVEF and

lower NT-proBNP. The interaction between the variables shown in Fig. 4-c1 is quite significant. Near the zero value of the vertical axis, the number of blue points far exceeds the number of red points, indicating that a smaller LA can reduce the impact of Hb on the results to a certain extent. Figure 4-c2 shows that when patients have the same level of Hb, patients with low NT-proBNP are more likely to be diagnosed with persistent AF.

Further analysis of variables with significant cutoff values

In the SHAP analysis, we also obtained the explanation diagram of the model. The diagrams shown in Fig. 5-b1 ~ b4 visualize 500 samples, with the X-axis representing the size of the variable and the Y-axis representing the size of their impact on the results. The red area indicates a greater tendency to be diagnosed with persistent AF, while the blue area indicates a greater likelihood of being diagnosed with paroxysmal AF. When there is a clear dividing point between the red and blue areas of a variable, this indicates that the variable may have a certain cutoff point, and the values before and after this value point to completely different diagnoses. Therefore, we screened the heat map obtained by SHAP, and Fig. 5-a1 ~ a4 shows all variables with obvious cutoff values (LA, Hb, LVEF, LVDd). For these variables, we further performed restricted cubic spline analysis. Based on further analysis of the RCS results as shown in Fig. 5-a1 ~ a4, we found that these variables do have cutoff points, which make the variable values before and after this value have different correlations with the diagnostic classification. Specifically, when LA > 38 mm, Hb > 135 g/L, LVEF < 65%, and LVDd > 38 mm, patients are more likely to be finally diagnosed with persistent AF. The value of each cutoff point is not completely certain and there will be some fluctuations. At the same time, the trends of these variables as exposure factors are not the same.

Performance and explanation of binary classification models in subgroups

The CHA2DS2-VASc score is currently the most widely used stroke risk assessment tool for patients with AF [24], and age is an important part of it. The international community generally adopts 65 years old as the threshold for the score, but the newly released Chinese guidelines for the diagnosis and treatment of AF adjust this age threshold to 60 years old [25]. Therefore, we used 60 and 65 as the age thresholds for subgroup analysis. We divide the participants of this study into three groups: under 60 years old, 60–65 years old, 65 years old, and above. The model has achieved good prediction performance among different age subgroups, as shown in Fig. 6-a ~ 6-c. The best AUC of people under 60 years old is 0.891 (95% CI: 0.856 to 0.892), with high SPE (0.885, 95% CI: 0.803 to 0.957) and high SEN (0.717, 95%CI: 0.574 to 0.792). The

best AUC of people aged 60 to 65 years old is 0.905 (95% CI: 0.860 to 0.948), with SPE equals 0.875 (95% CI: 0.731 to 0.966) and SEN equals 0.759 (95%CI: 0.625 to 0.904). Among people over 65 years old, the best AUC is 0.837 (95% CI: 0.827 to 0.862), with SPE equals 0.824 (95%CI: 0.789 to 0.855) and SEN equals 0.687 (95% CI: 0.610 to 0.759). We show these results and the specific values of other evaluation indicators in Table 3. The most important and second influencing factors of any age subgroup are LA and NT-proBNP, which is similar to the overall model. More results of the SHAP method are shown in Fig. 6-d ~ 6-f.

Furthermore, we conducted a subgroup analysis that combined gender and age, which was divided into male ≥ 60 years old, male < 60 years old, female ≥ 60 years old, and female < 60 years old. In these subgroups, we made model comparisons by the machine learning method and variable importance analysis by the SHAP method, which are shown in Fig. 7. The baseline data of male and female patients in this study are shown in Supplementary Table 2. In terms of diagnostic performance, the model achieved the highest AUC of 0.893 (95% CI: 0.777 to 0.972), with SPE equals 0.894 (95% CI: 0.789 to 0.967) and SEN equals 0.726 (95% CI: 0.632 to 0.884) in the subgroup of male that age < 60 years old. The lowest AUC was obtained in the subgroup of male ≥ 60 years old, only equals 0.822(95% CI: 0.790 to 0.848), with SPE equals 0.780 (95% CI: 0.764 to 0.799) and SEN equals 0.665 (95% CI: 0.574 to 0.754). More results of model performance evaluation by different machine learning methods can be found in Supplementary Table 3. As for the importance of variables, we found that in all subgroups combining gender and age, the most important and second variables are LA and NT-proBNP, as shown in Fig. 7-b1 ~ b4. Except in the subgroup of female < 60 years old, these two variables (LA and NT-proBNP) became the second and the first respectively.

Discussion

We used ten commonly available clinical indicators to derive a new model for predicting AF subtypes, and it has a high AUC (0.877), SPE (0.716), and SEN (0.851). This is the first time that the subtypes of AF (paroxysmal and persistent AF) could be predicted based on the baseline EHR at admission by machine learning.

AF is a common cause of stroke, heart failure, cardiovascular death, and dementia [26–29]. Recent reviews have proposed that burden-based descriptions of temporal AF patterns associated with outcomes and treatment strategies could improve risk prediction based on the classification of early paroxysmal, persistent, and long-standing persistent AF [11]. In particular, persistent AF is more strongly associated with serious adverse events such as stroke, systemic embolism, hospitalization for

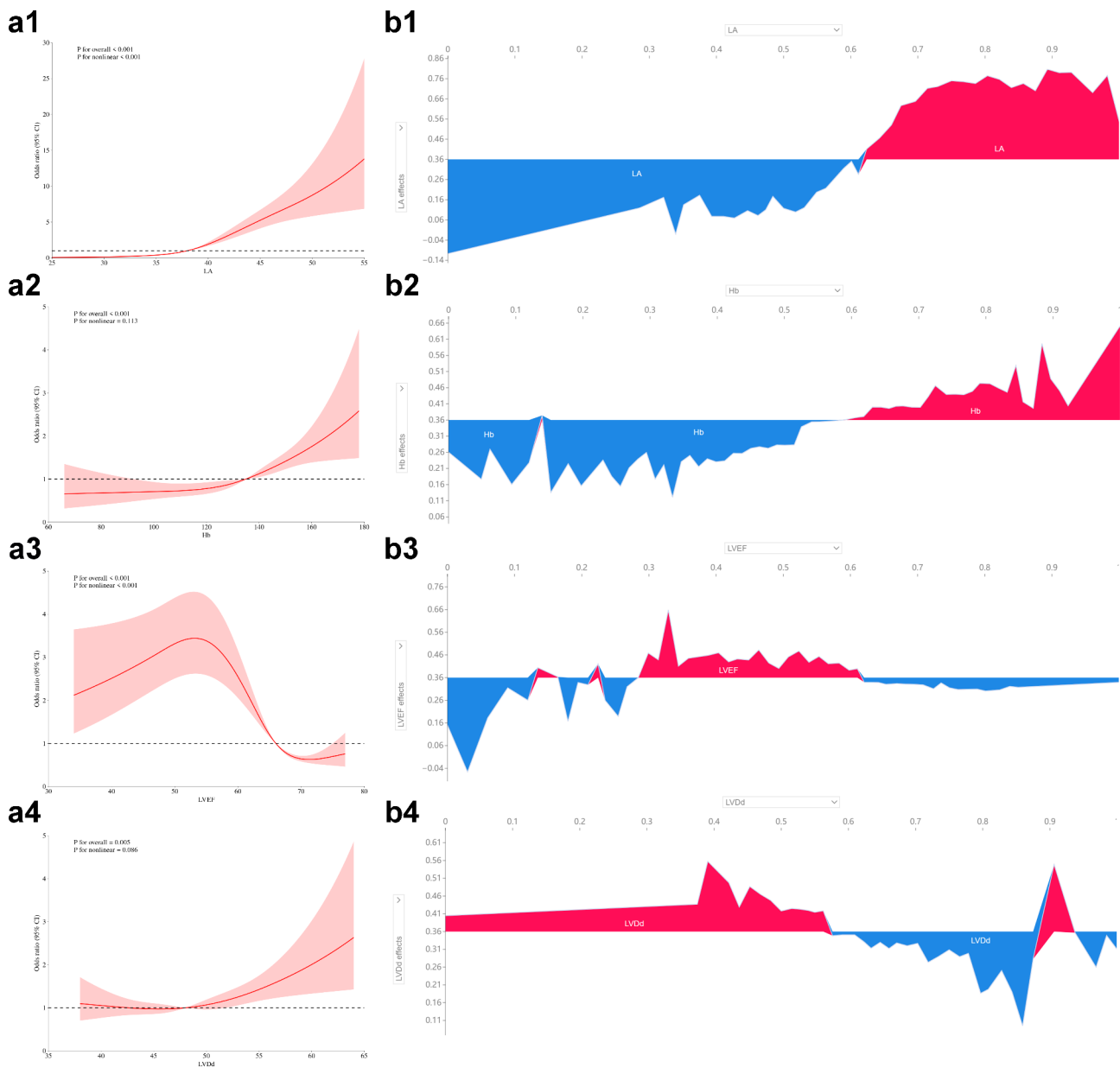


Fig. 5 Heatmap and RCS of variables with sharp cutoffs using the SHAP method. Figure 5 describes the restricted cubic spline analysis of variables with obvious cutoff points and the heatmap of these variables. In order to further observe the correlation between variables and predicted results, we performed RCS. Figure 5-a1 ~ 5-a4 show the nonlinear correlation between variables with obvious cutoff points and predicted results. The horizontal axis of the RCS series of graphs is the variable size, and the vertical axis is the Odds ratio value between paroxysmal and persistent AF. Figure 5-a1 and 5-a2 show that LA and Hb have similar correlations with paroxysmal and persistent AF subtypes, which are protective factors at low values and risk factors at high values. Figure 5-a3 shows that LVEF is a risk factor at low values and a protective factor at high values. For these variables, we further show their heat maps. The charts shown in Fig. 5-b1 ~ b4 visualize 500 samples, with the X-axis representing the size of the variables and the Y-axis representing their impact on the results. The red area indicates a greater tendency to be diagnosed with persistent AF, and the blue area indicates a greater possibility of being diagnosed with paroxysmal AF. When there is a clear dividing point between the red and blue areas of a variable, it means that the variable may have a certain cutoff point, and the values before and after this value point to completely different diagnoses

heart failure, and other cardiovascular morbidity and mortality than paroxysmal AF [30–32]. The risk factors for the early prediction of paroxysmal or persistent AF have become particularly essential.

Our study found 10 variables that are strongly correlated with the prediction of paroxysmal AF and persistent

AF. The results obtained by the SHAP method show that LA is the most important factor associated with the prediction of AF, followed by NT-proBNP. Several studies have found that LA was an effective indicator for predicting the progression of AF [31–33], and its increase is associated with the progression of AF to persistent

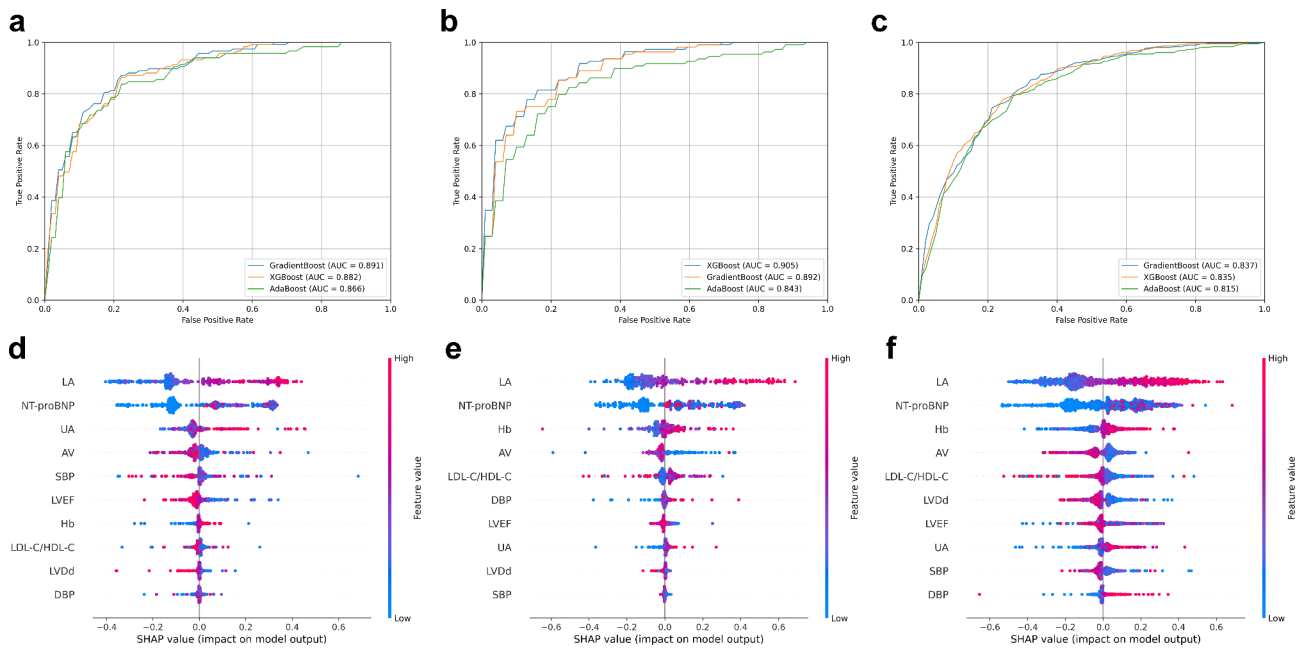


Fig. 6 Model prediction performance by AUC and interpretation by SHAP method in different age subgroup. Figure 6 shows the AUC results for age subgroups and the explanation of the model obtained using the SHAP method. The order of Fig. 6-a~6-c and Fig. 6-d~6-f is people under 60 years old, people aged 60 to 64 years old, and people aged 65 and above. Figure 6-a~6-c show the AUC of the three algorithms for different age subgroups. Figure 6-d~6-f show the order of variable importance, that is, SHAP value, in different age subgroups

Table 3 Results of model output indicators in different age subgroups

Model	ACC (95%CI)	Precision	Recall	AUC (95%CI)	F1 Score (95%CI)	SEN (95%CI)	SPE (95%CI)
Age under 60 years old							
GradientBoost	0.819 (0.783–0.856)	0.731	0.664	0.891 (0.856–0.892)	0.680 (0.544–0.779)	0.664 (0.417–0.825)	0.885 (0.803–0.957)
AdaBoost	0.819 (0.728–0.857)	0.731	0.717	0.866 (0.828–0.911)	0.709 (0.638–0.763)	0.717 (0.574–0.792)	0.862 (0.700–0.961)
XGBoost	0.816 (0.761–0.878)	0.703	0.700	0.882 (0.836–0.918)	0.695 (0.618–0.802)	0.700 (0.535–0.823)	0.866 (0.785–0.924)
Age from 60 to 64 years old							
GradientBoost	0.821 (0.746–0.898)	0.795	0.759	0.892 (0.842–0.929)	0.772 (0.684–0.872)	0.759 (0.625–0.904)	0.863 (0.781–0.934)
AdaBoost	0.776 (0.696–0.813)	0.725	0.722	0.843 (0.724–0.904)	0.721 (0.617–0.774)	0.722 (0.603–0.850)	0.812 (0.750–0.872)
XGBoost	0.825 (0.725–0.905)	0.812	0.749	0.905 (0.860–0.948)	0.776 (0.668–0.884)	0.749 (0.630–0.894)	0.875 (0.731–0.966)
Age 65 years old and above							
GradientBoost	0.765 (0.741–0.793)	0.687	0.687	0.837 (0.827–0.862)	0.686 (0.639–0.732)	0.687 (0.610–0.759)	0.812 (0.782–0.837)
AdaBoost	0.760 (0.721–0.811)	0.689	0.654	0.815 (0.770–0.836)	0.669 (0.605–0.758)	0.654 (0.568–0.790)	0.824 (0.798–0.854)
XGBoost	0.765 (0.746–0.810)	0.695	0.668	0.835 (0.820–0.872)	0.680 (0.639–0.745)	0.668 (0.596–0.742)	0.824 (0.789–0.855)

ACC, Accuracy; AUC, Area under curve; CI, Confidence interval; SEN, Sensitivity; SPE, Specificity

AF [34]. Our research has verified this view in the whole population, different ages, and different age groups of different genders. Recent data suggest that circulating biomolecules, especially elevated NT-proBNP [35–37], can identify patients at risk for AF and stroke because these biomolecules are associated with atrial dysfunction and AF [38]. Our results further validate previous studies suggesting that elevated NT-proBNP is associated with the development of paroxysmal AF [39].

It should be noted that gender and age factors are associated with the progression of AF [39]. The Relationship between increasing age and the progression of AF has been confirmed [34]. To further verify the consistency

of the study conclusions in different age subgroups, we conducted an age subgroup analysis. The cutoff points for grouping were 60 and 65 years old, which was based on the stroke risk score. The CHA2DS2-VASc score is currently the most widely used tool for assessing stroke risk [24], and age is an important part of the CHA2DS2-VASc score that affects this risk of stroke. Although the standard generally adopted in the world is 65 years old as the threshold of an integral, the newly released China guidelines for the diagnosis and management of AF adjust this age threshold to 60 years old [25]. This revision is based on the research evidence in Asia [40–42]. Our results showed that the prediction model had a high diagnostic

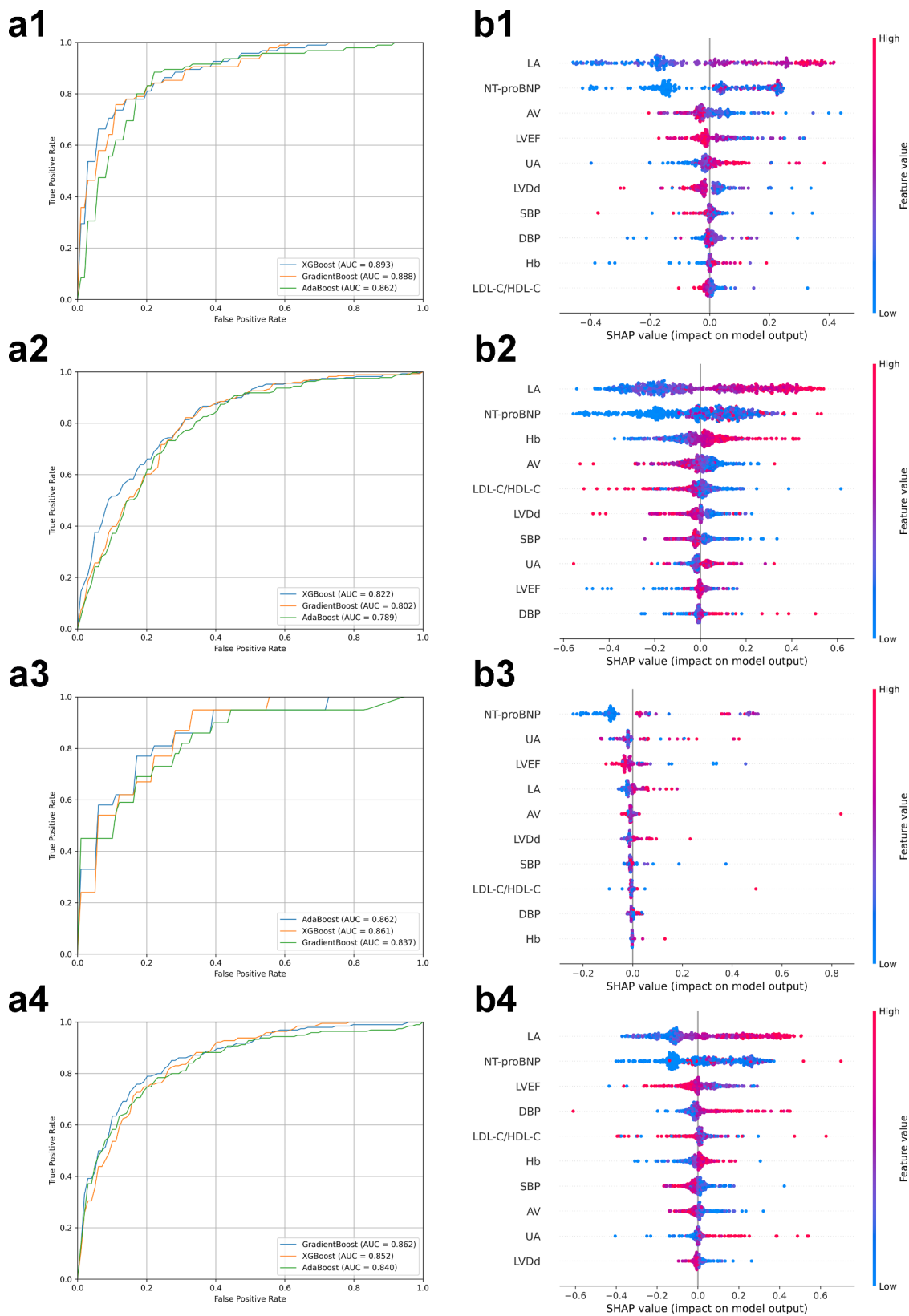


Fig. 7 Model prediction performance by AUC and interpretation by SHAP method in different gender and age group. Figure 7 shows the AUC results for different gender and age subgroups and the interpretation of the model obtained using the SHAP method. Figure 7-a and 7-b series represent the AUC and SHAP value rankings, respectively. 1, 2, 3, and 4 are male younger than 60 years old, male older than or equal to 60 years old, female younger than 60 years old, and female older than or equal to 60 years old, respectively

efficacy in people under 60 years old. This further shows that our study has strong clinical application significance, especially for young AF patients who are newly diagnosed. Among patients with paroxysmal AF, female are less likely to progress [39, 43]. Based on evidence from previous studies, we divided the subjects into four groups according to gender and age. In the groups aged <60 years and ≥ 60 years, male and female subgroups showed different results. Among subjects aged <60 years, male had higher AUC and SEN but weaker SPE than female. In the group aged ≥ 60 years, women had better AUC, SEN, and SPE. In summary, AF progression is a multifactorial disease, and our study also suggests differences between genders. This requires further research.

Our study also proposed some new variables associated with the prediction of paroxysmal AF and persistent AF, including Hb and UA. In our study, we found that the Hb level of patients with persistent AF was higher than that of the paroxysmal AF group, and high hemoglobin values are positively correlated with the prediction of AF. We believe that this conclusion is consistent with the conclusions of previous basic cardiovascular research [44]. Patients with AF have impaired cardiopulmonary function, which is manifested as a decrease in peak oxygen consumption [45]. Studies have shown that oxygen delivery is limited in the state of AF, and limited muscle oxygen uptake further increases tissue cell oxygen uptake [46]. Therefore, patients with AF may increase their Hb levels through compensatory reactions, thereby increasing their oxygen carrying capacity. The Hb level of the population is concentrated in the normal range, so fluctuations within the normal range may be individual differences between samples. Besides, our study is the first to explore the relationship between uric acid levels and AF subtypes (paroxysmal or persistent). High uric acid is positively correlated with persistent AF, and this indicator has a higher predictive value in people under 60 years of age. Although the relationship between uric acid and AF is still unclear, key pathways in the development of AF, such as cardiac electrical remodeling, structural remodeling, immune activation, insulin resistance, endothelial dysfunction, inflammatory response, and oxidative stress imbalance, have been shown to be closely related to UA [47]. In particular, hyperuricemia is independently associated with increased left atrial diameter, which is the physiological basis of AF structural remodeling [48]. Our results also showed that the importance of UA for model prediction was lower in female than in male, which is different from the conclusion of previous studies that the correlation between UA and AF was more significant in female instead of male [49]. Therefore, the underlying mechanism of the differences in the association between UA and AF and AF subtypes in different genders needs further study.

Nevertheless, our study also has some limitations. Due to strict inclusion and exclusion criteria, we only enrolled 1,600 participants from one of the largest hospitals in southern China. The smaller dataset size necessitates validation of the study results with a larger data set. In addition, our participants were all from one hospital with one ethnicity, which ensured the stability of the EHR while requiring research from more centers to generalize the results.

Conclusion

The predictive model developed in this study can be utilized to discern the specific subtypes in patients with newly diagnosed AF. Tailoring individualized treatment strategies based on this predictive model may help to realize early-stage management and treatment, ultimately leading to improved clinical outcomes.

Abbreviations

ACC	Accuracy
AF	Atrial fibrillation
AUC	Area under the curve
AV	Aortic valve flow velocity
BNP	Brain natriuretic peptide
CI	Confidence interval
DNP	Diastolic blood pressure
ECG	Electrocardiogram
ESC	European Society of Cardiology
Hb	Hemoglobin
HDL-C	High-density lipoprotein cholesterol
HER	Electronic health records
ICD	International Classification of Diseases
IQR	Interquartile range
LA	Left atrial diameter
LASSO	Least absolute shrinkage and selection operator
LDL-C	Low-density lipoprotein cholesterol
LVDD	Left ventricular end-diastolic diameter
LVEF	Left ventricular ejection fraction
ML	Machine learning
NT-proBNP	N-terminal pro-brain natriuretic peptide
PRE	Precision
RCS	Restricted cubic spline
RFE	Recursive Feature Elimination
ROC	Receiver operating characteristic
SBP	Systolic blood pressure
SEN	Sensitivity
SPE	Specificity
SD	Standard deviation
SHAP	SHapley Additive exPlanations
TTE	Transthoracic echocardiography
UA	Uric acid

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12911-025-02880-5>.

Supplementary Material 1

Acknowledgements

Not applicable.

Author contributions

Chao Tong, Kuan Zeng and Kun Zhang are the guarantors of the study. All authors (Yuqi Zhang, Sijin Li, Peibiao Mai, Yanqi Yang, Niansang Luo, Chao Tong, Kuan Zeng and Kun Zhang) were involved in the conceptualization and design of the study. Yuqi Zhang and Sijin Li were responsible for the experiment. Data cleaning was done by Sijin Li, Peibiao Mai, Yanqi Yang, and Niansang Luo. Analysis and interpretation were done by Yuqi Zhang and Sijin Li under the supervision and with the support of Niansang Luo, Chao Tong, Kuan Zeng and Kun Zhang. Drafting of the article was done by Yuqi Zhang, Sijin Li and Peibiao Mai. All authors (Yuqi Zhang, Sijin Li, Peibiao Mai, Yanqi Yang, Niansang Luo, Chao Tong, Kuan Zeng and Kun Zhang) revised and contributed to the intellectual content of the article. All authors (Yuqi Zhang, Sijin Li, Peibiao Mai, Yanqi Yang, Niansang Luo, Chao Tong, Kuan Zeng and Kun Zhang) approved the final version of the article, including the authorship list.

Funding

This study is partially supported by National Natural Science Foundation of China (62176016, 72274127), Guizhou Province Science and Technology Project: Research on Q&A Interactive Virtual Digital People for Intelligent Medical Treatment in Information Innovation Environment (supported by Qiankehe [2024] General 058), Capital Health Development Research Project (2022-2-2013), Haidian innovation and translation program from Peking University Third Hospital (HDCXZHKC2023203), Project: Research on the Decision Support System for Urban, Park Carbon Emissions Empowered by Digital Technology - A Special Study on the Monitoring and Identification of Heavy Truck Beidou Carbon Emission Reductions to Chao Tong. Project (YXYXCXRC202401, GCCRCY065, JCYJ20230807110302005) and National Natural Science Foundation of Guangdong Province (2022A1515011041) to Kuan Zeng.

Data availability

Some or all data sets generated during and/or analyzed during the present study are not publicly available but are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

The study reporting adheres to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines and has obtained informed consent from all participants, and only anonymous data was used in the analysis. This study was reviewed and approved by the Ethics Committees of Sun Yat-sen Memorial Hospital of Sun Yat-sen University (SYSKY-2024-004-01). The study adhered to the tenets of the Declaration of Helsinki.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹School of Computer Science & Engineering, Beihang University, Beijing, China

²Department of Cardiology, Joint Laboratory of Guangdong-Hong Kong-Macao Universities for Nutritional Metabolism and Precise Prevention and Control of Major Chronic Diseases, The Eighth Affiliated Hospital, Sun Yat-Sen University, Shenzhen, China

³State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing, China

⁴Department of Cardiovascular Surgery, The Eighth Affiliated Hospital, Sun Yat-Sen University, Shenzhen, China

⁵Department of Cardiology, Fuwai Hospital, Chinese Academy of Medical Sciences (Shenzhen Sun Yat-Sen Cardiovascular Hospital), Shenzhen, China

⁶Department of Cardiology, Sun Yat-Sen Memorial Hospital of Sun Yat-Sen University, Guangzhou, China

⁷Department of Cardiothoracic Surgery, University Hospital, University Linköping, Linköping, Sweden

⁸Department of Cardiology, The Seventh Affiliated Hospital, Sun Yat-Sen University, Shenzhen, China

Received: 6 October 2024 / Accepted: 20 January 2025

Published online: 03 February 2025

References

1. Schnabel RB, Yin X, Gona P, et al. 50 year trends in atrial fibrillation prevalence, incidence, risk factors, and mortality in the Framingham Heart Study: a cohort study. *Lancet*. 2015;386(9989):154–62.
2. Tsao CW, Aday AW, Almarzooq ZI, et al. Heart Disease and Stroke Statistics-2023 update: a Report from the American Heart Association. *Circulation*. 2023;147(8):e93–621.
3. Roth GA, Mensah GA, Johnson CO, et al. Global Burden of Cardiovascular diseases and Risk factors, 1990–2019: Update from the GBD 2019 study. *J Am Coll Cardiol*. 2020;76(25):2982–3021.
4. Joglar JA, Chung MK, Armbruster AL, et al. 2023 ACC/AHA/ACCP/HRS Guideline for the diagnosis and management of Atrial Fibrillation: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice guidelines. *Circulation*. 2024;149(1):e1–156.
5. Writing Committee M, Joglar JA, Chung MK, et al. 2023 ACC/AHA/ACCP/HRS Guideline for the diagnosis and management of Atrial Fibrillation: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice guidelines. *J Am Coll Cardiol*. 2024;83(1):109–279.
6. Uittenbogaart SB, Verbiest-van Gorp N, Lucassen WAM, et al. Opportunistic screening versus usual care for detection of atrial fibrillation in primary care: cluster randomised controlled trial. *BMJ*. 2020;370:m3208.
7. Lubitz SA, Atlas SJ, Ashburner JM, et al. Screening for Atrial Fibrillation in older adults at primary care visits: VITAL-AF Randomized Controlled Trial. *Circulation*. 2022;145(13):946–54.
8. Shi S, Tang Y, Zhao Q, et al. Prevalence and risk of atrial fibrillation in China: a national cross-sectional epidemiological study. *Lancet Reg Health West Pac*. 2022;23:100439.
9. Van Gelder IC, Rienstra M, Bunting KV, et al. 2024 ESC guidelines for the management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS). *Eur Heart J*. 2024;45(36):3314–414.
10. Tzeis S, Gerstenfeld EP, Kalman J et al. 2024 European Heart Rhythm Association/Heart Rhythm Society/Asia Pacific Heart Rhythm Society/Latin American Heart Rhythm Society expert consensus statement on catheter and surgical ablation of atrial fibrillation. *Europace*. 2024;26(4).
11. Becher N, Metzner A, Toennis T, Kirchhof P, Schnabel RB. Atrial fibrillation burden: a new outcome predictor and therapeutic target. *Eur Heart J*. 2024;45(31):2824–38.
12. Charitos EI, Purerfellner H, Glotzer TV, Ziegler PD. Clinical classifications of atrial fibrillation poorly reflect its temporal persistence: insights from 1,195 patients continuously monitored with implantable devices. *J Am Coll Cardiol*. 2014;63(25 Pt A):2840–8.
13. Vanassche T, Lauw MN, Eikelboom JW, et al. Risk of ischaemic stroke according to pattern of atrial fibrillation: analysis of 6563 aspirin-treated patients in ACTIVE-A and AVERROES. *Eur Heart J*. 2015;36(5):281–a7.
14. Pecun RBS, Engler L. Atrial fibrillation patterns are associated with arrhythmia progression and clinical outcomes. *Heart*. 2018;104(19):1608–14.
15. Goecks J, Jalili V, Heiser LM, Gray JW. How machine learning will transform Biomedicine. *Cell*. 2020;181(1):92–101.
16. Zhang Y, Li S, Wu W, et al. Machine-learning-based models to predict cardiovascular risk using oculomics and clinic variables in KNHANES. *BioData Min*. 2024;17(1):12.
17. Zhang Y, Yu M, Tong C, Zhao Y, Han J. CA-UNet Segmentation makes a good ischemic stroke risk prediction. *Interdiscip Sci*. 2024;16(1):58–72.
18. Freund Y, Schapire R, Abe N. A short introduction to boosting. *Journal-Japanese Soc Artif Intell*. 1999;14(771–780):1612.
19. Friedman JH. Greedy function approximation: a gradient boosting machine. *Annals Stat*. 2001;1189–232.
20. Chen T, Guestrin C. Xgboost: A scalable tree boosting system. 2016;2016. p. 785–94.
21. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. *Adv Neural Inf Process Syst*. 2017;30.
22. Gupta S, Glezerman IG, Hirsch JS, et al. Derivation and external validation of a simple risk score for predicting severe acute kidney injury after intravenous cisplatin: cohort study. *BMJ*. 2024;384:e077169.

23. Yang J, Wang T, Li K, Wang Y. Associations between per- and polyfluoroalkyl chemicals and abdominal aortic calcification in middle-aged and older adults. *J Adv Res*. 2024. <https://doi.org/10.1016/j.jare.2024.04.022>
24. van der Endt VHW, Milders J, Penning de Vries BBL, et al. Comprehensive comparison of stroke risk score performance: a systematic review and meta-analysis among 6 267 728 patients with atrial fibrillation. *Europace*. 2022;24(11):1739–53.
25. Ma C, Wu S, Liu S, Han Y. Chinese guidelines for the diagnosis and management of atrial fibrillation. *Pacing Clin Electrophysiol*. 2024;47(6):714–70.
26. Hindricks G, Potpara T, Dagres N, et al. 2020 ESC guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS): the Task Force for the diagnosis and management of atrial fibrillation of the European Society of Cardiology (ESC) developed with the special contribution of the European Heart Rhythm Association (EHRA) of the ESC. *Eur Heart J*. 2021;42(5):373–498.
27. Fabritz L, Crijns H, Guasch E, et al. Dynamic risk assessment to improve quality of care in patients with atrial fibrillation: the 7th AFNET/EHRA Consensus Conference. *Europace*. 2021;23(3):329–44.
28. Rivard L, Friberg L, Conen D, et al. Atrial fibrillation and dementia: a Report from the AF-SCREEN International collaboration. *Circulation*. 2022;145(5):392–409.
29. Kim D, Yang PS, Sung JH, et al. Less dementia after catheter ablation for atrial fibrillation: a nationwide cohort study. *Eur Heart J*. 2020;41(47):4483–93.
30. Potpara TS, Stankovic GR, Beleslin BD, et al. A 12-year follow-up study of patients with newly diagnosed lone atrial fibrillation: implications of arrhythmia progression on prognosis: the Belgrade Atrial Fibrillation study. *Chest*. 2012;141(2):339–47.
31. Ogawa H, An Y, Ikeda S, et al. Progression from paroxysmal to sustained Atrial Fibrillation is Associated with increased adverse events. *Stroke*. 2018;49(10):2301–8.
32. De With RR, Marcos EG, Dudink E, et al. Atrial fibrillation progression risk factors and associated cardiovascular outcome in well-phenotyped patients: data from the AF-RISK study. *Europace*. 2020;22(3):352–60.
33. Dudink E, Erkuner O, Berg J, et al. The influence of progression of atrial fibrillation on quality of life: a report from the Euro Heart Survey. *Europace*. 2018;20(6):929–34.
34. Padfield GJ, Steinberg C, Swampillai J, et al. Progression of paroxysmal to persistent atrial fibrillation: 10-year follow-up in the Canadian Registry of Atrial Fibrillation. *Heart Rhythm*. 2017;14(6):801–7.
35. Brady PF, Chua W, Nehaj F, et al. Interactions between atrial fibrillation and Natriuretic Peptide in Predicting Heart failure hospitalization or Cardiovascular Death. *J Am Heart Assoc*. 2022;11(4):e022833.
36. Chua W, Khashaba A, Canagarajah H et al. Disturbed atrial metabolism, shear stress, and cardiac load contribute to atrial fibrillation after ablation: AXAFA biomolecule study. *Europace*. 2024;26(2).
37. Inohara T, Kim S, Pieper K, et al. B-type natriuretic peptide, disease progression and clinical outcomes in atrial fibrillation. *Heart*. 2019;105(5):370–7.
38. Chua W, Law JP, Cardoso VR, et al. Quantification of fibroblast growth factor 23 and N-terminal pro-B-type natriuretic peptide to identify patients with atrial fibrillation using a high-throughput platform: a validation study. *PLoS Med*. 2021;18(2):e1003405.
39. Nguyen BO, Weberndorfer V, Crijns HJ, et al. Prevalence and determinants of atrial fibrillation progression in paroxysmal atrial fibrillation. *Heart*. 2022;109(3):186–94.
40. Kim TH, Yang PS, Yu HT, et al. Age threshold for ischemic stroke risk in Atrial Fibrillation. *Stroke*. 2018;49(8):1872–9.
41. Li YG, Lee SR, Choi EK, Lip GY. Stroke Prevention in Atrial Fibrillation: focus on Asian patients. *Korean Circ J*. 2018;48(8):665–84.
42. Choi SY, Kim MH, Lee KM, et al. Age-Dependent Anticoagulant Therapy for Atrial Fibrillation patients with Intermediate Risk of ischemic stroke: a Nationwide Population-based study. *Thromb Haemost*. 2021;121(9):1151–60.
43. Mulder BA, Khalilian Ekrami N, Van De Lande ME et al. Women have less progression of paroxysmal atrial fibrillation: data from the RACE V study. *Open Heart*. 2023;10(2).
44. Lim WH, Choi EK, Han KD, Lee SR, Cha MJ, Oh S. Impact of hemoglobin levels and their dynamic changes on the risk of Atrial Fibrillation: a Nationwide Population-based study. *Sci Rep*. 2020;10(1):6762.
45. Lam CS, Rienstra M, Tay WT, et al. Atrial fibrillation in heart failure with preserved Ejection Fraction: Association with Exercise Capacity, Left Ventricular Filling pressures, Natriuretic Peptides, and left atrial volume. *JACC Heart Fail*. 2017;5(2):92–8.
46. Kaye DM, Silvestry FE, Gustafsson F, et al. Impact of atrial fibrillation on rest and exercise haemodynamics in heart failure with mid-range and preserved ejection fraction. *Eur J Heart Fail*. 2017;19(12):1690–7.
47. Tamariz L, Hernandez F, Bush A, Palacio A, Hare JM. Association between serum uric acid and atrial fibrillation: a systematic review and meta-analysis. *Heart Rhythm*. 2014;11(7):1102–8.
48. Giannopoulos G, Angelidis C, Delftereos S. Gout and arrhythmias: in search for causation beyond association. *Trends Cardiovasc Med*. 2019;29(1):41–7.
49. Lin WD, Deng H, Guo P, et al. High prevalence of hyperuricaemia and its impact on non-valvular atrial fibrillation: the cross-sectional Guangzhou (China) Heart Study. *BMJ Open*. 2019;9(5):e028007.

Publisher's note

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