



Predicting All-Cause Mortality Risk in Atrial Fibrillation Patients: A Novel LASSO-Cox Model Generated From a Prospective Dataset

Yu Chen^{1†}, Shiwan Wu^{1†}, Jianfeng Ye¹, Muli Wu¹, Zhongbo Xiao¹, Xiaobin Ni¹, Bin Wang¹, Chang Chen¹, Yequn Chen^{1,2,3*}, Xuerui Tan^{1,2,3*} and Ruisheng Liu⁴

¹ Department of Cardiology, The First Affiliated Hospital of Shantou University Medical College, Shantou, China, ² Clinical Research Center, The First Affiliated Hospital of Shantou University Medical College, Shantou, China, ³ Institute of Cardiac Engineering, The First Affiliated Hospital of Shantou University Medical College, Shantou, China, ⁴ Department of Molecular Pharmacology and Physiology, Morsani College of Medicine, University of South Florida, Tampa, FL, United States

OPEN ACCESS

Edited by:

Tong Liu,
Tianjin Medical University, China

Reviewed by:

Yutao Guo,
Chinese PLA General Hospital, China
Jiangang Zou,
Nanjing Medical University, China

*Correspondence:

Xuerui Tan
doctortxr@126.com
Yequn Chen
gdcycyq@163.com

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Cardiac Rhythmology,
a section of the journal
Frontiers in Cardiovascular Medicine

Received: 25 June 2021

Accepted: 20 September 2021

Published: 18 October 2021

Citation:

Chen Y, Wu S, Ye J, Wu M, Xiao Z,
Ni X, Wang B, Chen C, Chen Y, Tan X
and Liu R (2021) Predicting All-Cause
Mortality Risk in Atrial Fibrillation
Patients: A Novel LASSO-Cox Model
Generated From a Prospective
Dataset.
Front. Cardiovasc. Med. 8:730453.
doi: 10.3389/fcvm.2021.730453

Background: Although mortality remains high in patients with atrial fibrillation (AF), there have been limited studies exploring machine learning (ML) models on mortality risk prediction in patients with AF.

Objectives: This study sought to develop an ML model that captures important variables in order to predict all-cause mortality in AF patients.

Methods: In this single center prospective study, an ML-based mortality prediction model was developed and validated using a dataset of 2,012 patients who experienced AF from November 2018 to February 2020 at the First Affiliated Hospital of Shantou University Medical College. The dataset was randomly divided into a training set (70%, $n = 1,223$) and a validation set (30%, $n = 552$). A total of 122 features were collected for variable selection. Least absolute shrinkage and selection operator (LASSO) and random forest (RF) algorithms were used for variable selection. Ten ML models were developed using variables selected by LASSO or RF. The best model was selected and compared with conventional risk scores. A nomogram and user-friendly online tool were developed to facilitate the mortality predictions and management recommendations.

Results: Thirteen features were selected by the LASSO regression algorithm. The LASSO-Cox model achieved an area under the curve (AUC) of 0.842 in the training dataset, and 0.854 in the validation dataset. A nomogram based on eight independent features was developed for the prediction of survival at 30, 180, and 365 days following discharge. Both the time dependent receiver operating characteristic (ROC) and decision curve analysis (DCA) showed better performances of the nomogram compared to the CHA₂DS₂-VASc and HAS-BLED models.

Conclusions: The LASSO-Cox mortality predictive model shows potential benefits in death risk evaluation for AF patients over the 365-day period following discharge. This novel ML approach may also provide physicians with personalized management recommendations.

Keywords: atrial fibrillation, machine learning, prediction model, mortality, risk factors

INTRODUCTION

AF is one of the most common chronic cardiovascular health problems globally (1–3). In Europe and the USA, 2–3 % of the population suffers from AF (4), and it is estimated that AF will affect 6–12 million people in the USA by 2,050 and 17.9 million people in Europe by 2,060 (5, 6). The incidence of AF is not high among young people but increases with age, reaching more than 10 % in those >80 years of age (7). The inevitable global aging of the population, combined with a cumulative increase in chronic cardiovascular diseases, will lead to considerable growth in the number of AF patients in the next few decades. AF is associated with a nearly five-fold increased risk of ischemic stroke (8, 9), and provokes significant increases in all-cause mortality along with important financial burden (10, 11). Consequently, higher risk of all-cause mortality associated with AF has become a significant public health issue (1, 11–13).

Several classic risk scores, including CHA₂DS₂-VASc and HAS-BLED scores, predict clinical outcomes, such as for stroke, bleeding and mortality (14–17). Machine learning can learn to identify the underlying pattern and classes from multidimensional data by utilizing computational algorithms (18). Based on novel ML algorithms, more accurate and intelligent models, such as the Global Anticoagulant Registry in the Field (GARFIELD)-AF risk model and the Multilayer Neural Network artificial intelligence model, have been developed (19–21). In contrast to the high awareness regarding clinical outcomes of AF in Europe and the USA, there is limited knowledge for East Asia. In addition, few ML models have used multi-dimensional features to predict future mortality of AF patients.

Advances in supervised ML allow the recognition and translation of multi-dimensional data into valuable models (21, 22). The use of machine learning for predicting clinical outcomes may enable physicians to improve efficiency, reliability, and accuracy of management decisions. In the present study, we used multiple ML approaches that included LASSO feature selection and the Cox proportional hazards regression model to predict all-cause mortality outcome over the 30–365-day period after discharge in patients with AF.

METHODS

Study Cohort

For machine learning model construction, a prospective observational study was undertaken using data from patients who were hospitalized for evaluation and treatment of AF between November 2018 and February 2020 at the First Affiliated Hospital of Shantou University Medical College. Inclusion criteria were a diagnosis of AF and availability of complete data concerning clinical indicators for evaluating AF and follow-up. The diagnosis of AF required recording the heart rhythm by electrocardiogram (ECG). Three diagnostic criteria shown by ECG are: (1) absolutely irregular RR intervals, (2) no discernible, distinct P waves, and (3) an episode lasting at least 30 s. Many individuals with AF have both symptomatic and asymptomatic episodes. The exclusion criteria were pregnant women, age \leq 18, or patients who refused follow-up.

Data Collection

A systemic clinical evaluation for AF was conducted during the hospitalization when patients were enrolled. Overall, 122 variables were initially used for the selection of key features (**Supplementary Table 1**), which included medical histories, physical examinations, laboratory examination results, medications, comorbidities, ultrasonic cardiogram, CHA₂DS₂-VASc score, and HAS-BLED score. Follow-up by outpatient follow-up and/or telephone interview was carried out at 30, 180, and 365 days after discharge. The main outcome of the AF cohort was all-cause death.

This study complied with the principles of the Declaration of Helsinki and was approved by the Ethics Committee of the First Affiliated Hospital of Shantou University Medical College. All participants provided written informed consent to participate in this study. All procedures were performed in conformity with the European Society of Cardiology guidelines (23).

Variable Selection and Model Development

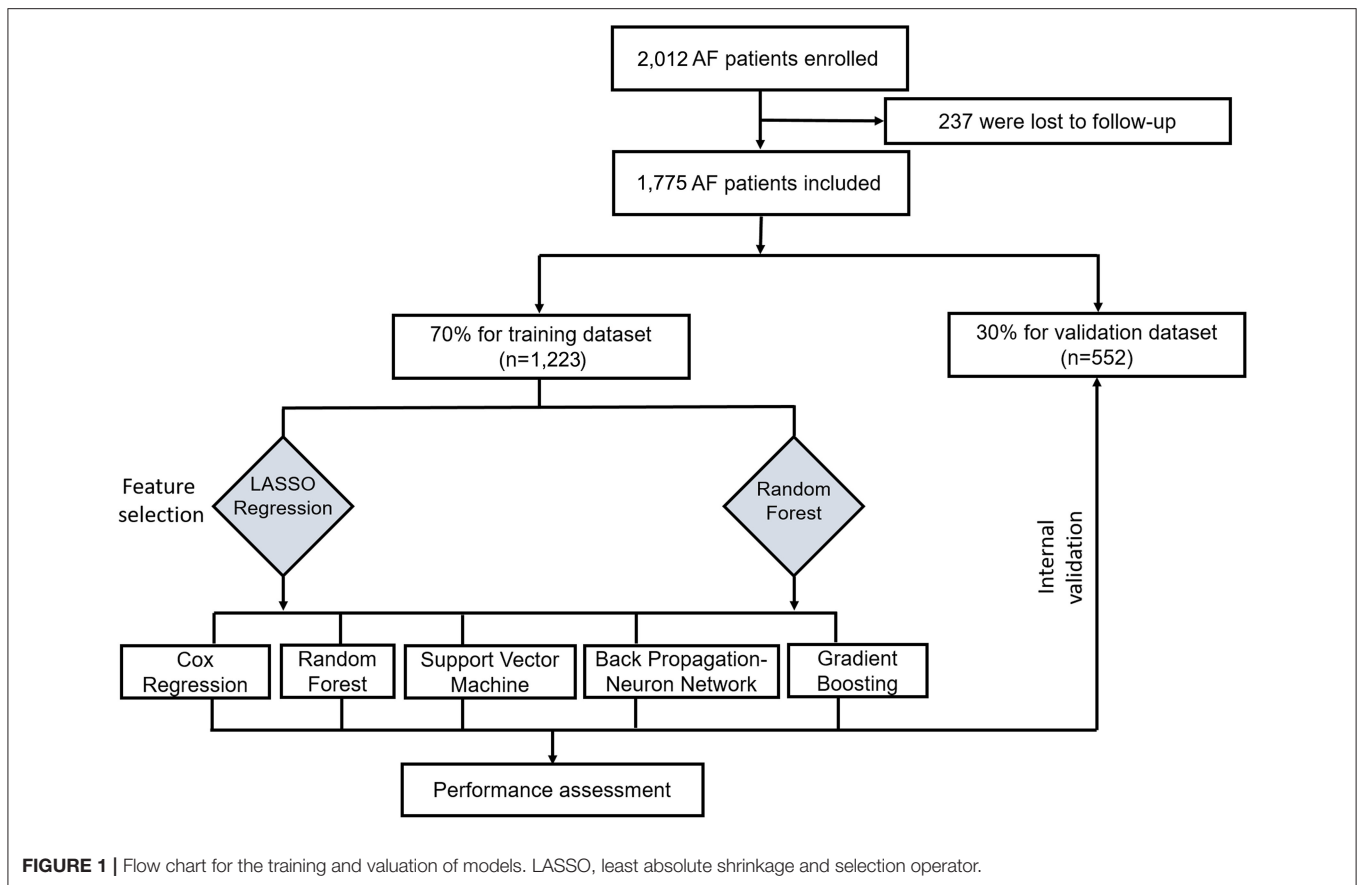
Due to the 122 variables present in the dataset, conducting variable selection was necessary and could lead to improved prediction performance. Both the LASSO algorithm (24) and RF (25) were used to select the features for model training. The top 20 predictor variables were chosen using RF based on relative variable importance (26).

We used five algorithms, including Cox regression, RF, support vector machines (SVM) (27), backpropagation neural networks (BP-NN) (28), and gradient boosting (GB) (29), to train models using the variables that were selected by LASSO and RF. Ultimately, 10 models, including LASSO-Cox, LASSO-RF, LASSO-SVM, LASSO-BP-NN, LASSO-GB, RF-Cox, RF-RF, RF-SVM, RF-BP-NN, and RF-GB, were established.

Statistical Analyses and Model Performance Measures

Statistical analyses were performed using SPSS 23.0 (Inc., Chicago, Illinois, USA), X-tile 3.6.1 (30), and R (version 4.0.2; R Foundation for Statistical Computing, Vienna, Austria) software. Continuous variables are presented as the mean \pm standard derivation. We used multiple imputation to account for missing data on continuous variables if missing data was <30% (31). Missing values were imputed using the “mice” package. Categorical variables are presented as numbers and percentages. Statistical differences of continuous variables were examined by two-tailed *t*-tests or Mann-Whitney *U* tests. Categorical variables were analyzed by the chi-square test or Fisher exact test. Various R packages were used to conduct this study. The glmnet package was used for logistic regression with LASSO regularization (32). Random forest, e1071, neural net, and gbm packages were used for the RF, SVM, BP-NN, and GB models, respectively (29, 33).

The predictive accuracy of the LASSO-Cox model was compared with the performances of CHA₂DS₂-VASc and HAS-BLED scores. The performances of the models were assessed by the AUC derived from receiver operating characteristics curves. A nomogram for predicting the 30-, 180- or 360-day survival was established using the LASSO-Cox regression model, and the cut-off value for mortality risk stratification was calculated.



The nomogram and calibration plots were generated with the rms package. The pROC package was used to plot ROC curves. Kaplan-Meier curves were produced using the survival package. $P < 0.05$ was considered to indicate statistical significance.

RESULTS

Patient Baseline Demographics

This study was conducted according to the flow chart shown in **Figure 1**. Eligible study participants consisted of 1,775 AF patients. A total of 1,223 AF patients were randomly assigned in the training dataset and 552 patients in the validation dataset. Baseline characteristics of the study cohort are shown in **Table 1**. The mean age was 69.22 years (SD = 12.05 years) for the training dataset and 69.02 years (SD = 11.65 years) for the validation dataset. The mean CHA₂DS₂-VASc was 3.37 (SD = 1.18) in the training set and 3.19 (SD = 1.80) in the validation set. There were no significant differences in diabetes, atherosclerosis, prior stroke, heart failure, cerebral hemorrhage, cancer, renal insufficiency, bleeding, current smoker status, statin medication, and urine ketone bodies in the training set compared with the validation set. An all-cause mortality end point event occurred for 194 of the 1,775 patients (10.9%, 111 males and 83 females), 143 in the training set (11.7%) and 51 in the validation set

(9.2%). There was no significant difference in all-cause death rate between the training and validation set.

Feature Selection and Model Performance Comparison

LASSO coefficient profiles of the 122 variables and ten-fold cross-validation for tuning parameter selection in the LASSO model are shown in **Figure 2**. Thirteen variables were selected by the LASSO regression algorithm, including CHA₂DS₂-VASc, stroke, cancer, red cell volume distribution width-coefficient of variation (RDW-CV), statin medication use, lymphocyte ratio, neutrophil-to-lymphocyte ratio, basophilic granulocyte number, urine ketone body (KET), blood glucose (GLU), blood urea nitrogen (BUN), cholinesterase (CHE), and monoamine oxidase (MAO). In addition, the top-20 variables were selected by the RF algorithm (**Supplementary Table 2**). Next, we built 10 models using these two sets of selected features, and their prediction performances were described using AUC, sensitivity, and specificity (**Figure 3**). The key performance of machine learning was evaluated by AUC.

Among the 10 models, LASSO-BP-NN had the highest AUC (0.910, 95% CI: 0.875–0.944) in the training dataset, but a relatively low AUC (0.685, 95% CI: 0.613–0.756) in the validation dataset. The LASSO-Cox model, over the 1-year follow-up,

TABLE 1 | Baseline characteristics of the AF prospective cohort.

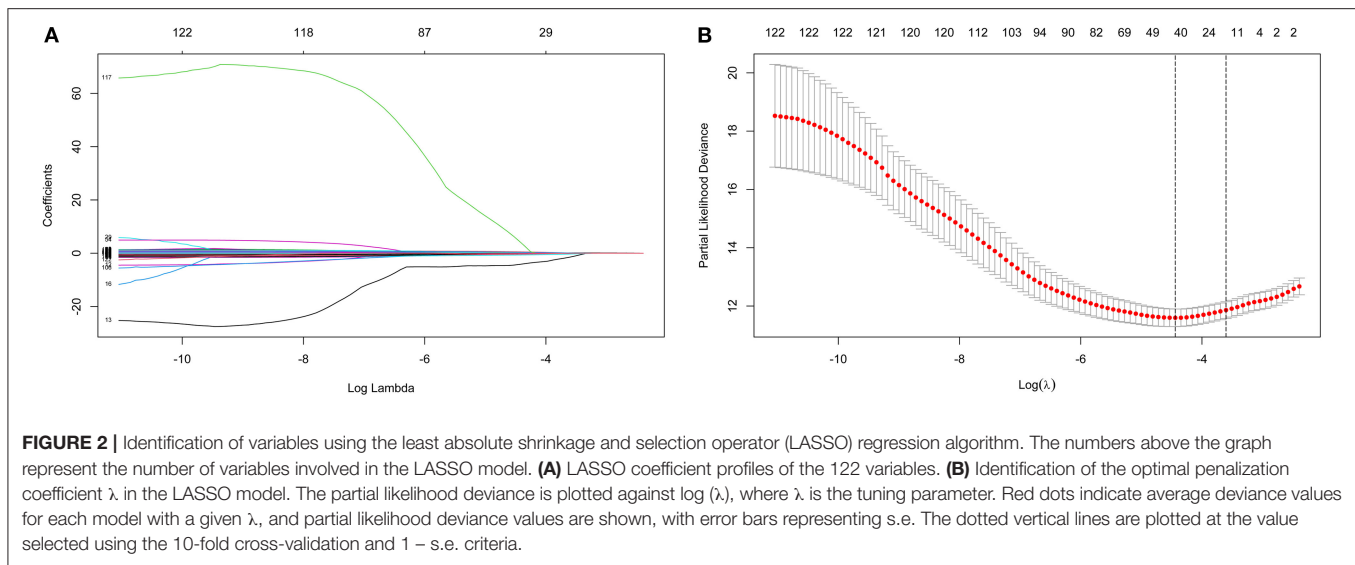
Variable	Training set (n = 1,223)	Validation set (n = 552)	P-value
Age (years)	69.22 ± 12.05	69.02 ± 11.65	0.749
Gender: male	682 (55.8)	334 (60.5)	0.062
Diabetes mellitus	324 (26.5)	163 (29.5)	0.184
Atherosclerosis	446 (36.5)	202 (36.6)	0.959
Hypertension	759 (62.1)	297 (53.8)	<0.001
Prior stroke	319 (26.1)	127 (23.0)	0.167
Cerebral hemorrhage	36 (2.9)	9 (1.6)	0.103
Heart failure	319 (26.1)	119 (21.6)	0.041
Cancer	80 (6.5)	30 (5.4)	0.371
Renal insufficiency	154 (12.6)	67 (12.1)	0.788
Bleeding	47 (3.8)	22 (4.0)	0.886
Current smoker status	347 (28.4)	150 (27.2)	0.603
Alcohol	87 (7.1)	31 (5.6)	0.241
Anticoagulant treatment	696 (56.9)	299 (54.2)	0.281
β-Blocker treatment	546 (44.6)	237 (42.9)	0.502
Digoxin treatment	398 (32.5)	192 (34.8)	0.354
Statin medication	486 (39.7)	230 (41.7)	0.443
WBC (10 ⁹ /L)	8.43 ± 3.71	8.30 ± 3.29	0.47
KET			0.463
0	1,093 (89.4)	497 (90.0)	
1	124 (10.1)	52 (9.4)	
2	6 (0.5)	2 (0.4)	
3	0 (0.0)	1 (0.2)	
Lymphocyte ratio	20.20 ± 10.65	20.59 ± 10.93	0.48
Creatinine (mmol/L)	119.05 ± 78.09	117.84 ± 73.22	0.758
RDW-CV (%)	14.53 ± 2.01	14.36 ± 1.79	0.962
Platelet (10 ⁹ /L)	205.18 ± 69.61	202.30 ± 73.27	0.426
Platelet distribution width (%)	15.62 ± 1.98	15.67 ± 1.93	0.648
GLU (mmol/L)	7.00 ± 3.65	6.76 ± 3.28	0.191
BUN (mmol/L)	7.77 ± 4.75	7.92 ± 5.17	0.542
CHE (U/L)	6.42 ± 1.96	6.41 ± 1.96	0.927
MAO (U/L)	4.89 ± 2.35	4.91 ± 2.68	0.862
Neutrophils/lymphocytes (%)	5.85 ± 6.61	5.96 ± 7.19	0.748
CHA ₂ DS ₂ -VASc score			0.51
0	59 (4.8)	33 (6.0)	
1	139 (11.4)	76 (13.8)	
2	198 (16.2)	99 (17.9)	
3	256 (20.9)	103 (18.7)	
4	250 (20.4)	110 (19.9)	
5	169 (13.8)	68 (12.3)	
6	100 (8.2)	46 (8.3)	
7	42 (3.4)	12 (2.2)	
8	10 (0.8)	5 (0.9)	

Data are represented as mean ± standard deviation or frequency (percentage). WBC, White blood cell; KET, Urine ketone body; RDW-CV, Red cell volume distribution width - coefficient of variation; GLU, Blood glucose; BUN, Blood urea nitrogen; CHE, Cholinesterase; MAO, Monoamine oxidase.

achieved an AUC of 0.842 (95% CI: 0.809–0.875) in the training dataset, and an AUC of 0.854 (95% CI: 0.807–0.901) in the validation dataset. Due to the very good performances in both the training set and validation set, the LASSO-Cox regression was chosen as the best model.

Nomogram Construction

Based on the Cox proportional hazards regression analysis, we identified eight independent risk factors in the training cohort. CHA₂DS₂-VASc (hazard ratio, HR = 1.188, *P* = 0.002), stroke (HR = 1.717, *P* = 0.008), cancer (HR = 2.208, *P* = 0.002),



statin medication use (HR = 0.341, $P < 0.001$), KET (HR = 1.730, $P = 0.006$), BUN (HR = 1.037, $P = 0.003$), CHE (HR = 0.889, $P = 0.032$), and MAO (HR = 1.133, $P < 0.001$) were all significantly associated with mortality in AF patients (**Supplementary Table 3**).

A nomogram based on the eight independent features from the training cohort was developed for the prediction of the 30-, 180-, and 365-day survival (**Figure 4**). The nomogram demonstrated that MAO contributes the most to survival, followed by CHE, KET, BUN, CHA₂DS₂-VASc, stroke, statin use, and cancer. The total score, obtained by adding the scores for each of the eight features, helped in estimating the 30-, 180-, and 365-day survival rate for each individual patient.

Validation and Calibration of the Nomogram

ROC curves were used to evaluate the predictive ability for 30-, 180-, and 365-day survival in both the training and validation sets. Our Cox model demonstrated good discriminative ability in both the training (30-day AUC: 0.848, 180-day AUC: 0.826, 365-day AUC: 0.762) and validation (30-day AUC: 0.834, 180-day AUC: 0.788, 365-day AUC: 0.841) datasets for the 30-, 180-, and 365-day survival rates (**Supplementary Figure 1**). The calibration plots of our nomogram also showed optimal agreement between the actual observations and the predicted outcomes both in the training set and validation set (**Supplementary Figure 2**) for all time points. Thus, the above nomogram-based results displayed good accuracy for predicting the 30-, 180-, and 365-day survival of AF patients.

Comparison of the Nomogram With CHA₂DS₂-VASc and HAS-BLED Models for Predictive Performance

The time-dependent ROCs of the training and validation sets (**Figure 5**) based on the nomogram were higher than those based

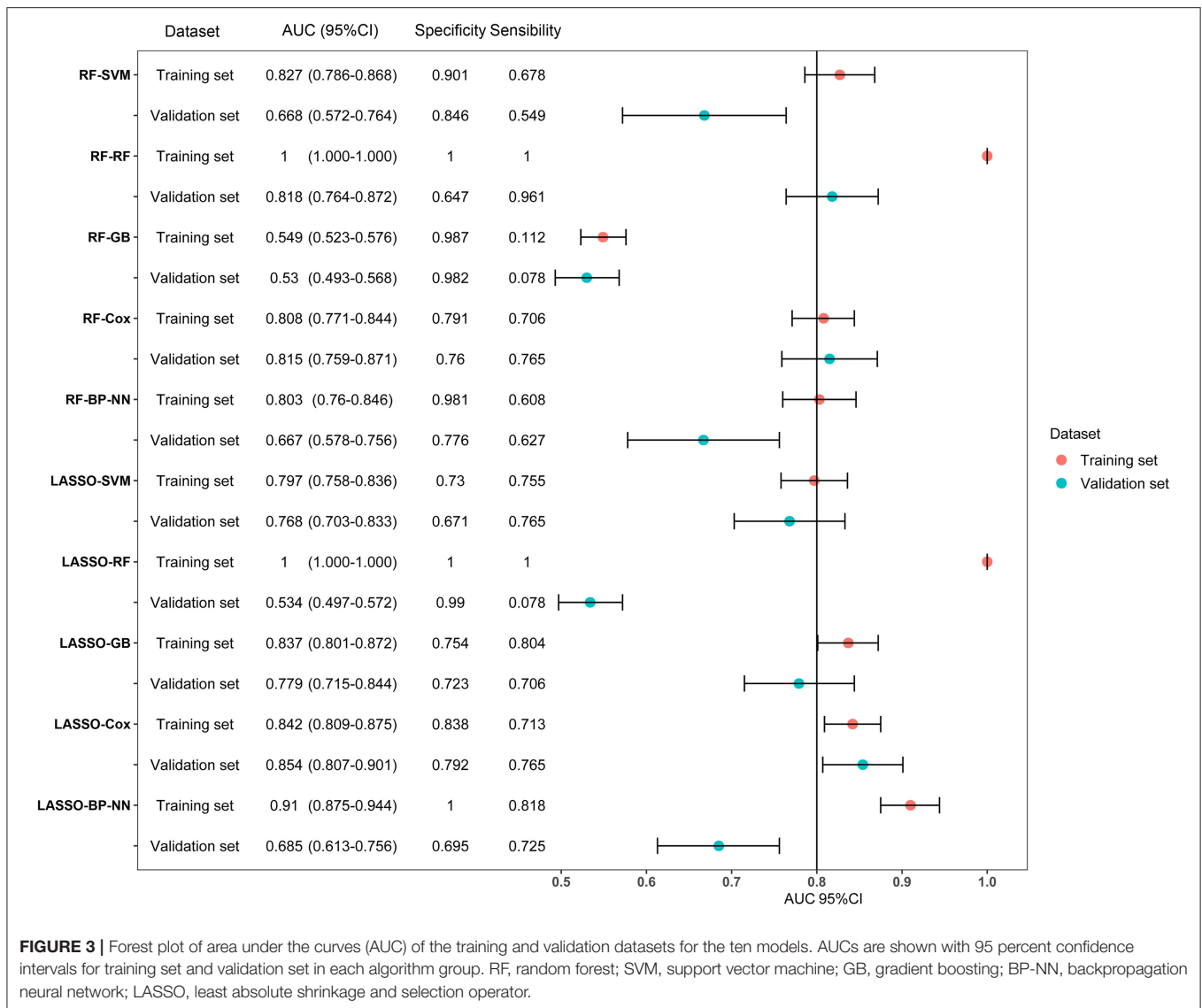
on the traditional CHA₂DS₂-VASc and HAS-BLED models. These results indicate that our nomogram has greater potential for accurately predicting prognosis compared to the traditional models. DCA was performed to compare the net benefit of the nomogram with that of the traditional CHA₂DS₂-VASc and HAS-BLED scores. Compared to the CHA₂DS₂-VASc and HAS-BLED scores, the curve of our nomogram showed larger net benefit (**Figure 6**). We further converted the nomogram to a web calculator for the clinician's convenience (<https://afnom.shinyapps.io/DynNomapp/>).

In addition, the optimal cut-off point was determined using the X-tile program to accomplish risk stratification. As shown in **Supplementary Figure 3**, the optimal cut-off point was 0.8. Thus, we stratified the AF patients into a low-risk group (≤ 0.8) and high-risk group (> 0.8). Kaplan–Meier curves showed that the high-risk group exhibited poorer survival than the low-risk group in both the training and validation sets (**Supplementary Figure 4**).

DISCUSSION

This study investigated a novel LASSO-Cox model for the prediction of all-cause mortality in patients with AF to identify AF patients at high risk and to provide personalized treatment using a data-driven approach. Several important findings were identified. First, eight independent risk factors predicted all-cause mortality, including CHA₂DS₂-VASc score, CHE, KET, BUN, MAO, stroke, statin medication use, and cancer. Second, a LASSO-Cox model for 30-, 180-, and 365-day risk prediction was established and validated. Third, the use of the nomogram and risk stratification enables the prediction of mortality for AF patients.

Machine learning can identify non-linear associations and identify interactions in complex and multidimensional variables. The use of the LASSO ML algorithm for variable selection is a well-established method that has been previously utilized for



cancer, heart failure, and AF populations (34–36). The advantages of the LASSO algorithm are high accuracy and stability. Cox proportional hazards regression is a traditional model, that is mainly used to analyze the prognosis of cancer and other chronic diseases. Indeed, our LASSO-Cox model was robust and displayed good discriminatory power in predicting all-cause mortality both in the training and the validation dataset.

There is growing evidence that AF significantly worsens the mortality rate (37–39). Furthermore, AF is an independent risk factor for higher risk of mortality (11). While worse outcomes among AF patients have been confirmed in various studies from Europe and North America, data from East Asia is limited.

Traditional guidelines in AF have focused on identifying patients with different risks of stroke and major bleeding. Several studies have developed and examined prediction models or risk scores in AF patients for stroke, major bleeding,

or composite outcomes, although not exclusively for death outcomes (19, 23, 40). Recently, a death risk score based on age, biomarkers, and clinical history (ABC) was developed and performed well in two large independent clinical trial cohorts (41). However, the detection of novel biomarkers such as GDF-15 are not easily performed in developing countries and regions.

In this LASSO-Cox model, not taking statins is an independent risk factor for AF-associated death. As recently reported, the levels of total cholesterol (TC) are non-linearly associated with all-cause mortality, as well as cancer and cardiovascular disease mortality, in the American population (42). Thus, it is necessary to maintain TC in a moderate range by statin medication. The GARFIELD-AF and ROCKET AF studies have shown that heart failure and sudden cardiac death are the major reasons for death of AF patients taking oral anticoagulant

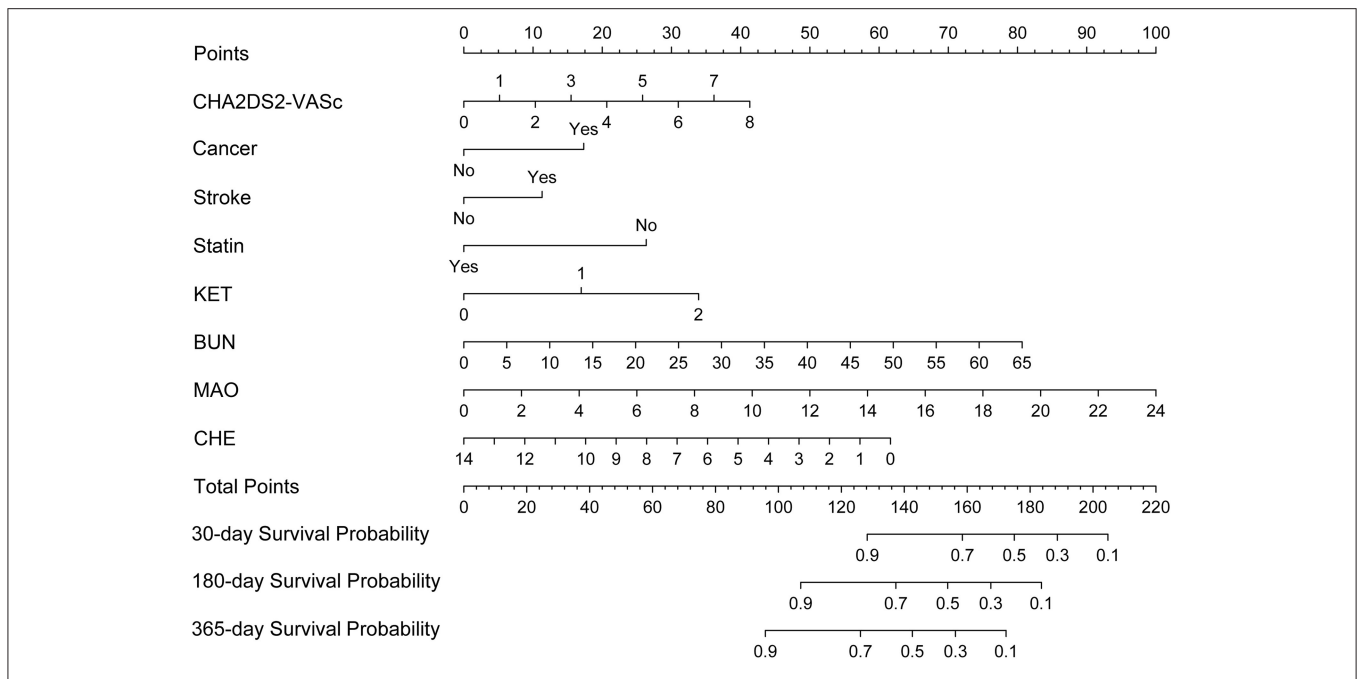


FIGURE 4 | Nomogram for predicting 30-, 180-, and 365-day survival probabilities for AF patients. To calculate patient survival probabilities, obtain points for each covariate value by dropping a vertical line from the points axis to the value of each covariate, calculate the total points obtained from all eight covariate values, and then drop a vertical line from the total points axis to locate the associated 30-, 180-, or 365-day survival probability. KET, urine ketone body; BUN, blood urea nitrogen; CHE, cholinesterase; MAO, monoamine oxidase.

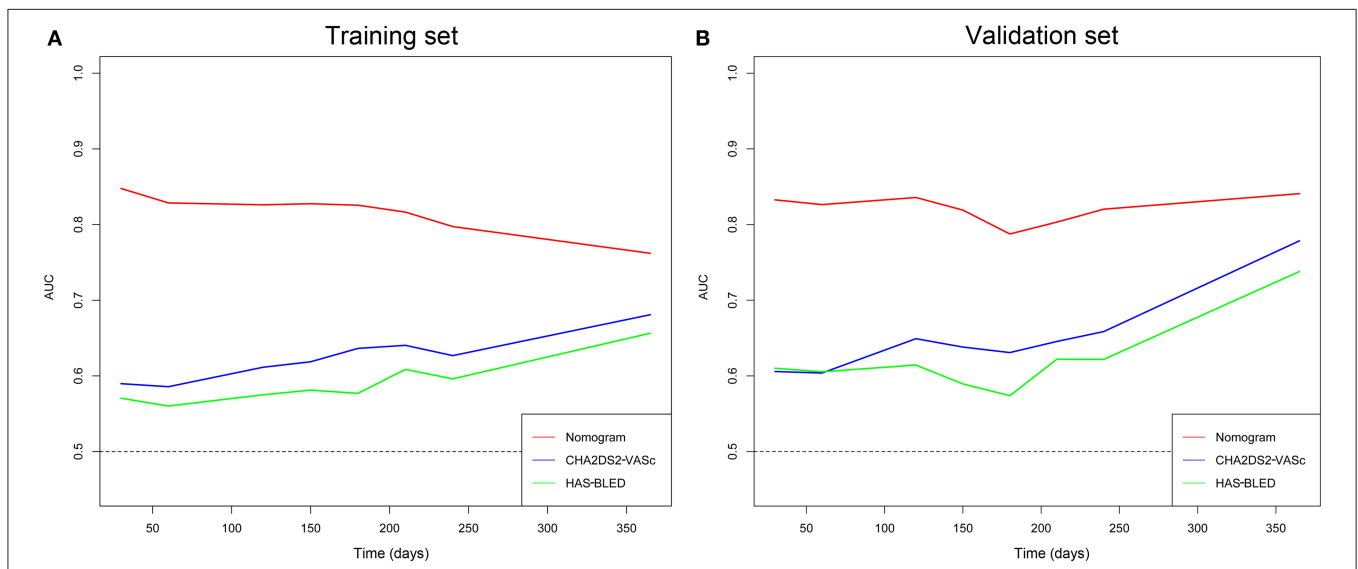
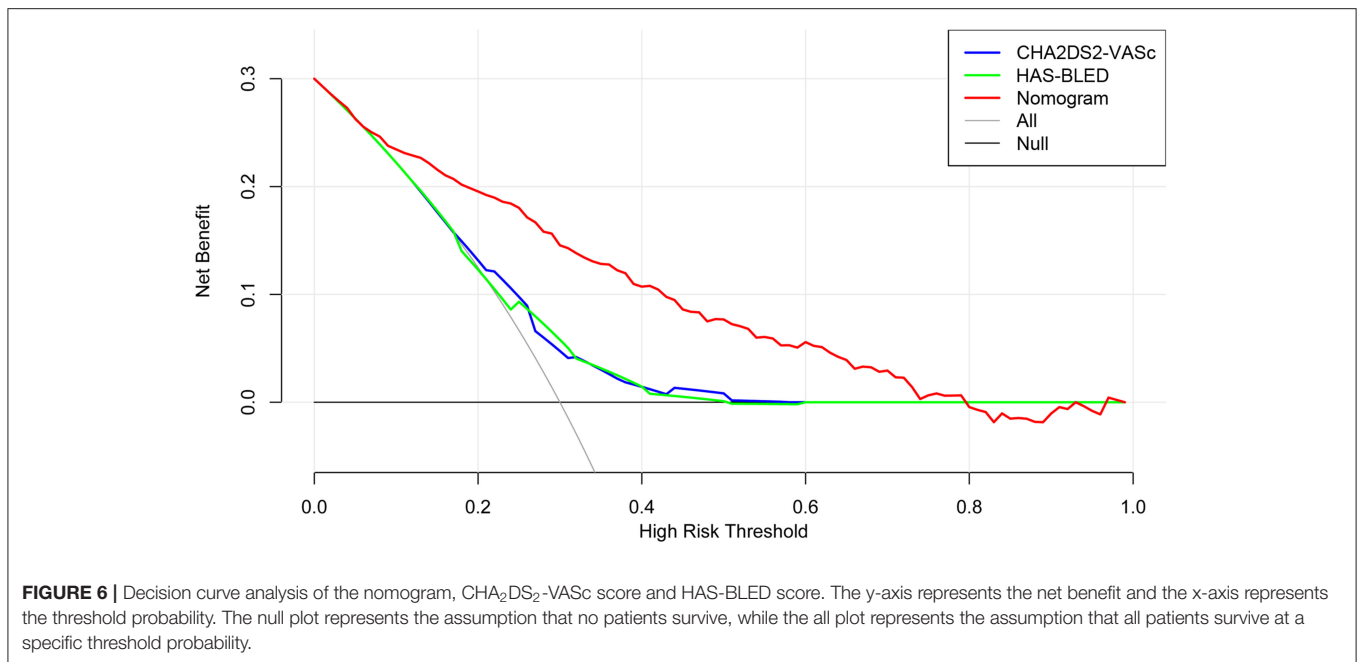


FIGURE 5 | Time-dependent ROC of the nomogram compared with CHA₂DS₂-VASc and HAS-BLED models in the training and validation sets. **(A)** Training set. **(B)** Validation set.

medication (38, 43). Death risk prediction in these patients may give rise to more intense management of risk factors, such as valvular heart disease, myocardial dysfunction, and coronary heart disease.

Among the independent risk factors of death, the four common laboratory examination indicators, including MAO, BUN, CHE, and KET, are strongly associated with mortality. Contemporary AF trials show that cardiac-related deaths account



for the vast majority of all deaths, whereas stroke and bleeding represent a small fraction (44). In our study, MAO is recognized as the most important mortality risk factor in AF patients. Elevated MAO is known to be associated with liver cirrhosis and chronic congestive heart failure. Recent studies show that MAO is a major source of deleterious reactive oxygen species (ROS), regulating cardiomyocyte aging or death (45, 46). Myocardial ROS are involved in the pathophysiology of cardiovascular diseases such as hypertension and heart failure (47, 48), and are important markers of atrial fibrillation in patients after cardiac surgery (49). Thus, MAO inhibition therapy is protective in several settings of cardiac stresses such as pressure overload heart failure, diabetic cardiomyopathy and chronic ischemic heart disease (47). Further studies exploring the potential relationship between AF and ROS are needed.

Increased BUN levels are mainly triggered by impaired renal function, which might be highly related to the occurrence of ischemic stroke in AF patients despite adequate therapeutic warfarin anticoagulation (50). A Swedish study showed that neoplastic disease and renal failure contribute to the increased risk of all-cause mortality in AF patients, which is consistent with our result (11). Declination of cholinesterase is associated with the advanced liver cirrhosis, hepatic failure, and myocardial infarction. Inhibition of CHE has been reported to directly affect the intrinsic cardiac nervous system (51). In addition, increased levels of KET reflects the severity of diabetes, and AF patients with diabetes mellitus have a higher mortality rate (52–54). Collectively, the above risk factors suggest a renewed emphasis on the management of comorbidities such as liver cirrhosis, renal dysfunction, heart failure, and diabetes mellitus, is essential to improve the overall survival and quality of life in AF patients.

The nomogram could provide clinicians with the opportunity to assess risk of all-cause mortality by using a data-driven approach. An additional strength of the LASSO-Cox model is that the eight predictive factors in this nomogram are widely and easily available internationally. In order to facilitate medical use, the clinical implementation of the LASSO-Cox model can either be based on the nomogram, or preferably an online tool.

Limitations

Several limitations of this LASSO-Cox model should be considered. First, validation of this model was performed using a dataset generated from a single center. The performance of our LASSO-Cox model in external datasets needs to be tested by data from other institutions. Second, the LASSO-Cox model did not include information about biomarkers, such as NT-proBNP and hs-cTnT. However, considering that these biomarkers often require additional examination, thus increasing the difficulty of acquisition, our model has good accuracy and ease of application. Third, multiple imputation for the missing values is a potential source of bias. Nevertheless, multiple-imputation is a commonly used rigorous technique for imputation (55).

CONCLUSION

A new LASSO-Cox model for predicting risk of all-cause mortality in patients with AF was successfully developed, and internally validated. The LASSO-Cox model using CHA₂DS₂-VASc score, statin medication, medical history (stroke, cancer), and four clinical examination parameters (KET, BUN, MAO, and CHE), performed well and

may assist physicians in decision-making when treating AF patients.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

This study complied with the principles of the Declaration of Helsinki and was approved by the Ethics Committee of the First Affiliated Hospital of Shantou University Medical College. All participants provided written informed consent to participate in this study. All procedures were performed in conformity with European society of cardiology guidelines. All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki declaration of 1975, as revised in 2000. Informed consent was obtained from all patients for being included in the study.

AUTHOR CONTRIBUTIONS

SW, YQC, and XT: concept and design, data analysis and interpretation, critical revision of article, and approval. YC, MW, ZX, XN, and BW: statistics, data analysis, and drafting of article. YC, SW, JY, CC, YQC, and RL: data collection, data analysis, critical revision of article, and approval. All authors read and approved the final manuscript.

REFERENCES

- Lippi G, Sanchis-Gomar F, Cervellin G. Global epidemiology of atrial fibrillation: an increasing epidemic and public health challenge. *Int J Stroke*. (2021) 16:217–21. doi: 10.1177/1747493019897870
- Bai Y, Wang YL, Shantsila A, Lip GYH. The global burden of atrial fibrillation and stroke: a systematic review of the clinical epidemiology of atrial fibrillation in asia. *Chest*. (2017) 152:810–20. doi: 10.1016/j.chest.2017.03.048
- Jelavic MM, Krstacic G, Pintaric H. Usage and safety of direct oral anticoagulants at patients with atrial fibrillation and planned diagnostic procedures, interventions, and surgery. *Heart and Mind*. (2019) 3:1. doi: 10.4103/hm.hm_61_19
- Kirchhof P. The future of atrial fibrillation management: integrated care and stratified therapy. *Lancet*. (2017) 390:1873–87. doi: 10.1016/S0140-6736(17)31072-3
- Krijthe BP, Kunst A, Benjamin EJ, Lip GY, Franco OH, Hofman A, et al. Projections on the number of individuals with atrial fibrillation in the European Union, from 2000 to 2060. *Eur Heart J*. (2013) 34:2746–51. doi: 10.1093/eurheartj/ehs280
- Patel NJ, Deshmukh A, Pant S, Singh V, Patel N, Arora S, et al. Contemporary trends of hospitalization for atrial fibrillation in the United States, 2000 through 2010: implications for healthcare planning. *Circulation*. (2014) 129:2371–9. doi: 10.1161/CIRCULATIONAHA.114.008201

FUNDING

This work was supported by projects from Grant for Key Disciplinary Project of Clinical Medicine under the High-level University Development Program (2020), Innovation Team Project of Guangdong Universities (2019KCXTD003), Li Ka Shing Foundation Cross-Disciplinary Research Grant (2020LKSFG19B), Funding for Guangdong Medical Leading Talent (2019-2022), and National Natural Science Foundation of China (82073659).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcvm.2021.730453/full#supplementary-material>

Supplementary Figure 1 | Receiver operating characteristic (ROC) curves for 30-, 180-, and 365-day survival rates. **(A)** Training set. **(B)** Validation set.

Supplementary Figure 2 | Calibration curves for the nomogram in the training and validation sets. Y-axis represents the actual survival rate, while the X-axis represents the nomogram-predicted survival rate. The blue dotted line indicates perfect prediction by an ideal model. **(A–C)** 30-, 180-, and 365-day survival rates in the training set. **(D–F)** 30-, 180-, and 365-day survival rates in the validation set.

Supplementary Figure 3 | Determination of the cut-off score, for the mortality risk stratification, using the X-tile program. A cut-off score ≤ 0.8 indicates low-risk, and > 0.8 indicates high-risk.

Supplementary Figure 4 | Kaplan-Meier curves for the high-risk group and low-risk group in the training and validation sets. **(A)** Training set. **(B)** Validation set.

Supplementary Table 1 | The 122 variables and the missing rates collected in the dataset.

Supplementary Table 2 | Top-20 variables selected by the RF algorithm.

Supplementary Table 3 | Hazard ratios and 95% confidence intervals of 8 variables in the Cox proportional hazards model.

- Zoni-Berisso M, Lercari F, Carazza T, Domenicucci S. Epidemiology of atrial fibrillation: European perspective. *Clin Epidemiol*. (2014) 6:213–20. doi: 10.2147/CLEP.S47385
- Odutayo A, Wong CX, Hsiao AJ, Hopewell S, Altman DG, Emdin CA. Atrial fibrillation and risks of cardiovascular disease, renal disease, and death: systematic review and meta-analysis. *BMJ*. (2016) 354:i4482. doi: 10.1136/bmj.i4482
- Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham study. *Stroke*. (1991) 22:983–8. doi: 10.1161/01.STR.22.8.983
- Mukherjee K, Kamal KM. Impact of atrial fibrillation on inpatient cost for ischemic stroke in the USA. *Int J Stroke*. (2019) 14:159–66. doi: 10.1177/1747493018765491
- Andersson T, Magnuson A, Bryngelsson IL, Fröbert O, Henriksson KM, Edvardsson N, et al. All-cause mortality in 272,186 patients hospitalized with incident atrial fibrillation 1995-2008: a Swedish nationwide long-term case-control study. *Eur Heart J*. (2013) 34:1061–7. doi: 10.1093/eurheartj/ehs469
- Wang Z, Chen Z, Wang X, Zhang L, Li S, Tian Y, et al. The disease burden of atrial fibrillation in china from a national cross-sectional survey. *Am J Cardiol*. (2018) 122:793–8. doi: 10.1016/j.amjcard.2018.05.015
- Chen LY, Chung MK, Allen LA, Ezekowitz M, Furie KL, McCabe P, et al. Atrial fibrillation burden: moving beyond atrial fibrillation as a binary entity:

- a scientific statement from the American heart association. *Circulation*. (2018) 137:e623–e44. doi: 10.1161/CIR.0000000000000568
14. Lip GY, Nieuwlaat R, Pisters R, Lane DA, Crijns HJ. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. *Chest*. (2010) 137:263–72. doi: 10.1378/chest.09-1584
 15. Gage BF, Waterman AD, Shannon W, Boehler M, Rich MW, Radford MJ. Validation of clinical classification schemes for predicting stroke: results from the national registry of atrial fibrillation. *JAMA*. (2001) 285:2864–70. doi: 10.1001/jama.285.22.2864
 16. Pisters R, Lane DA, Nieuwlaat R, de Vos CB, Crijns HJ, Lip GY, et al. Novel user-friendly score (HAS-BLED) to assess 1-year risk of major bleeding in patients with atrial fibrillation: the Euro heart survey. *Chest*. (2010) 138:1093–100. doi: 10.1378/chest.10-0134
 17. Camm AJ, Kirchhof P, Lip GY, Schotten U, Savelieva I, Ernst S, et al. Guidelines for the management of atrial fibrillation: the task force for the management of atrial fibrillation of the European society of cardiology (ESC). *Eur Heart J*. (2010) 31:2369–429. doi: 10.1093/eurheartj/ehq278
 18. Deo RC. Machine learning in medicine. *Circulation*. (2015) 132:1920–30. doi: 10.1161/CIRCULATIONAHA.115.001593
 19. Fox KAA, Lucas JE, Pieper KS, Bassand JP, Camm AJ, Fitzmaurice DA, et al. Improved risk stratification of patients with atrial fibrillation: an integrated GARFIELD-AF tool for the prediction of mortality, stroke and bleed in patients with and without anticoagulation. *BMJ Open*. (2017) 7:e017157. doi: 10.1136/bmjopen-2017-017157
 20. Hijazi Z, Granger CB, Hohnloser SH, Westerbergh J, Lindbäck J, Alexander JH, et al. Association of different estimates of renal function with cardiovascular mortality and bleeding in atrial fibrillation. *J Am Heart Assoc*. (2020) 9:e017155. doi: 10.1161/JAHA.120.017155
 21. Goto S, Goto S, Pieper KS, Bassand JP, Camm AJ, Fitzmaurice DA, et al. New artificial intelligence prediction model using serial prothrombin time international normalized ratio measurements in atrial fibrillation patients on vitamin K antagonists: GARFIELD-AF. *Eur Heart J Cardiovasc Pharmacother*. (2020) 6:301–9. doi: 10.1093/ehjcvp/pvz076
 22. Goto S, Kimura M, Katsumata Y, Goto S, Kamatani T, Ichihara G, et al. Artificial intelligence to predict needs for urgent revascularization from 12-lead electrocardiography in emergency patients. *PLoS ONE*. (2019) 14:e0210103. doi: 10.1371/journal.pone.0210103
 23. Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, et al. 2016 ESC guidelines for the management of atrial fibrillation developed in collaboration with EACTS. *Eur Heart J*. (2016) 37:2893–962. doi: 10.1093/eurheartj/ehw210
 24. Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med*. (1997) 16:385–95. doi: 10.1002/(SICI)1097-0258(19970228)16:4<385::AID-SIM380>3.0.CO;2-3
 25. Svetnik V, Liaw A, Tong C, Culbertson JC, Sheridan RP, Feuston BP. Random forest: a classification and regression tool for compound classification and QSAR modeling. *J Chem Inf Comput Sci*. (2003) 43:1947–58. doi: 10.1021/ci034160g
 26. Ambale-Venkatesh B, Yang X, Wu CO, Liu K, Hundley WG, McClelland R, et al. Cardiovascular event prediction by machine learning: the multi-ethnic study of atherosclerosis. *Circ Res*. (2017) 121:1092–101. doi: 10.1161/CIRCRESAHA.117.311312
 27. Byvatov E, Schneider G. Support vector machine applications in bioinformatics. *Appl Bioinformatics*. (2003) 2:67–77.
 28. Basheer IA, Hajmeer M. Artificial neural networks: fundamentals, computing, design, and application. *J Microbiol Methods*. (2000) 43:3–31. doi: 10.1016/S0167-7012(00)00201-3
 29. Natekin A, Knoll A. Gradient boosting machines, a tutorial. *Front Neurobot*. (2013) 7:21. doi: 10.3389/fnbot.2013.00021
 30. Camp RL, Dolled-Filhart M, Rimm DL. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin Cancer Res*. (2004) 10:7252–9. doi: 10.1158/1078-0432.CCR-04-0713
 31. Zhang Z. Multiple imputation with multivariate imputation by chained equation (MICE) package. *Ann Transl Med*. (2016) 4:30. doi: 10.3978/j.issn.2305-5839.2015.12.63
 32. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw*. (2010) 33:1–22. doi: 10.18637/jss.v033.i01
 33. Angraal S, Mortazavi BJ, Gupta A, Khera R, Ahmad T, Desai NR, et al. Machine learning prediction of mortality and hospitalization in heart failure with preserved ejection fraction. *JACC Heart Fail*. (2020) 8:12–21. doi: 10.1016/j.jchf.2019.06.013
 34. Mavaddat N, Michailidou K, Dennis J, Lush M, Fachal L, Lee A, et al. Polygenic risk scores for prediction of breast cancer and breast cancer subtypes. *Am J Hum Genet*. (2019) 104:21–34. doi: 10.1016/j.ajhg.2018.11.002
 35. Lanfear DE, Gibbs JJ, Li J, She R, Petucci C, Culver JA, et al. Targeted metabolomic profiling of plasma and survival in heart failure patients. *JACC Heart Fail*. (2017) 5:823–32. doi: 10.1016/j.jchf.2017.07.009
 36. Hill NR, Ayoubkhani D, McEwan P, Sugrue DM, Farooqui U, Lister S, et al. Predicting atrial fibrillation in primary care using machine learning. *PLoS ONE*. (2019) 14:e0224582. doi: 10.1371/journal.pone.0224582
 37. Ruddox V, Sandven I, Munkhaugen J, Skattebu J, Edvardsen T, Otterstad JE. Atrial fibrillation and the risk for myocardial infarction, all-cause mortality and heart failure: a systematic review and meta-analysis. *Eur J Prev Cardiol*. (2017) 24:1555–66. doi: 10.1177/204748731715769
 38. Pokorney SD, Piccini JP, Stevens SR, Patel MR, Pieper KS, Halperin JL, et al. Cause of death and predictors of all-cause mortality in anticoagulated patients with nonvalvular atrial fibrillation: data from ROCKET AF. *J Am Heart Assoc*. (2016) 5:e002197. doi: 10.1161/JAHA.115.002197
 39. Fauchier L, Villejoubert O, Clementy N, Bernard A, Pierre B, Angoulvant D, et al. Causes of death and influencing factors in patients with atrial fibrillation. *Am J Med*. (2016) 129:1278–87. doi: 10.1016/j.amjmed.2016.06.045
 40. January CT, Wann LS, Alpert JS, Calkins H, Cigarroa JE, Cleveland JC Jr, et al. 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American college of cardiology/American heart association task force on practice guidelines and the heart rhythm society. *J Am Coll Cardiol*. (2014) 64:e1–76. doi: 10.1016/j.jacc.2014.03.022
 41. Hijazi Z, Oldgren J, Lindbäck J, Alexander JH, Connolly SJ, Eikelboom JW, et al. A biomarker-based risk score to predict death in patients with atrial fibrillation: the ABC (age, biomarkers, clinical history) death risk score. *Eur Heart J*. (2018) 39:477–85. doi: 10.1093/eurheartj/ehx584
 42. He GD, Liu XC, Liu L, Yu YL, Chen CL, Huang JY, et al. A nonlinear association of total cholesterol with all-cause and cause-specific mortality. *Nutr Metab*. (2021) 18:25. doi: 10.1186/s12986-021-00548-1
 43. Bassand JP, Accetta G, Camm AJ, Cools F, Fitzmaurice DA, Fox KA, et al. Two-year outcomes of patients with newly diagnosed atrial fibrillation: results from GARFIELD-AF. *Eur Heart J*. (2016) 37:2882–9. doi: 10.1093/eurheartj/ehw233
 44. Gómez-Outes A, Lagunar-Ruiz J, Terleira-Fernández AI, Calvo-Rojas G, Suárez-Gea ML, Vargas-Castrillón E. Causes of death in anticoagulated patients with atrial fibrillation. *J Am Coll Cardiol*. (2016) 68:2508–21. doi: 10.1016/j.jacc.2016.09.944
 45. Manzella N, Santini Y, Maggiorani D, Martini H, Douin-Echinard V, Passos JF, et al. Monoamine oxidase-A is a novel driver of stress-induced premature senescence through inhibition of parkin-mediated mitophagy. *Aging Cell*. (2018) 17:e12811. doi: 10.1111/acel.12811
 46. Santini Y, Sicard P, Vigneron F, Guilbeau-Frugier C, Dutaur M, Lairez O, et al. Oxidative stress by monoamine oxidase-a impairs transcription factor EB activation and autophagosome clearance, leading to cardiomyocyte necrosis and heart failure. *Antioxid Redox Signal*. (2016) 25:10–27. doi: 10.1089/ars.2015.6522
 47. Mialet-Perez J, Parini A. Cardiac monoamine oxidases: at the heart of mitochondrial dysfunction. *Cell Death Dis*. (2020) 11:54. doi: 10.1038/s41419-020-2251-4
 48. Phoswa WN. Dopamine in the pathophysiology of preeclampsia and gestational hypertension: monoamine oxidase (MAO) and catechol-O-methyl transferase (COMT) as possible mechanisms. *Oxid Med Cell Longev*. (2019) 2019:3546294. doi: 10.1155/2019/3546294
 49. Anderson EJ, Efrid JT, Davies SW, O'Neal WT, Darden TM, Thayne KA, et al. Monoamine oxidase is a major determinant of redox balance in human atrial myocardium and is associated with postoperative atrial fibrillation. *J Am Heart Assoc*. (2014) 3:e000713. doi: 10.1161/JAHA.113.000713

50. Aachi RV, Birnbaum LA, Topel CH, Seifi A, Hafeez S, Behrouz R. Laboratory characteristics of ischemic stroke patients with atrial fibrillation on or off therapeutic warfarin. *Clin Cardiol.* (2017) 40:1347–51. doi: 10.1002/clc.22838
51. Darvesh S, Arora RC, Martin E, Magee D, Hopkins DA, Armour JA. Cholinesterase inhibitors modify the activity of intrinsic cardiac neurons. *Exp Neurol.* (2004) 188:461–70. doi: 10.1016/j.expneurol.2004.05.002
52. January CT, Wann LS, Calkins H, Chen LY, Cigarroa JE, Cleveland JC Jr, et al. 2019 AHA/ACC/HRS focused update of the 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American college of cardiology/American heart association task force on clinical practice guidelines and the heart rhythm society. *J Am Coll Cardiol.* (2019) 74:104–32.
53. Domek M, Li YG, Gumprecht J, Asaad N, Rashed W, Alsheikh-Ali A, et al. One-year all-cause mortality risk among atrial fibrillation patients in Middle East with and without diabetes: the gulf SAFE registry. *Int J Cardiol.* (2020) 302:47–52. doi: 10.1016/j.ijcard.2019.12.061
54. Echouffo-Tcheugui JB, Shrader P, Thomas L, Gersh BJ, Kowey PR, Mahaffey KW, et al. Care patterns and outcomes in atrial fibrillation patients with and without diabetes: ORBIT-AF registry. *J Am Coll Cardiol.* (2017) 70:1325–35. doi: 10.1016/j.jacc.2017.07.755
55. Park SY, Freedman ND, Haiman CA, Le Marchand L, Wilkens LR, Setiawan VW. Association of coffee consumption with total and cause-specific mortality among nonwhite populations. *Ann Intern Med.* (2017) 167:228–35. doi: 10.7326/M16-2472

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Chen, Wu, Ye, Wu, Xiao, Ni, Wang, Chen, Chen, Tan and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.