Development of a risk score for predicting one-year mortality in patients with atrial fibrillation using XGBoost-assisted feature selection

Bin Wang[1]*, Feifei Jin[2, 3, 4]*, Han Cao[5]*, Qing Li[6], Ping Zhang[6]

[1]School of Clinical Medicine, Tsinghua University, Beijing, China

[2]Trauma Medicine Center, Peking University People's Hospital, Beijing, China

[3]Key Laboratory of Trauma Treatment and Neural Regeneration, Peking University, Ministry of Education, Beijing, China

[4]National Center for Trauma Medicine of China, Beijing, China

[5]Medical Data Science Center, Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua University, Beijing, China

[6]Department of Cardiology, Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua University, Beijing, China

*These authors equally contributed to the study.

Editorial
by Chen et al.

Correspondence to:
Ping Zhang, MD,
Department of Cardiology,
Beijing Tsinghua Changgung
Hospital,
School of Clinical Medicine,
Tsinghua University,
No. 168 Litang Road, Changping
District, Beijing, 102218, China,
phone: +86 010 561 189 15,
e-mail: zhpdoc@126.com

ABSTRACT

Background: There are no tools specifically designed to assess mortality risk in patients with atrial fibrillation (AF).

Aims: This study aimed to utilize machine learning methods to identify pertinent variables and develop an easily applicable prognostic score to predict 1-year mortality in AF patients.

Methods: This study, based on the Medical Information Mart for Intensive Care-IV (MIMIC-IV) database, focused on patients aged 18 years and older with AF. A critical care database from China was the external validation set. The importance of variables from XGBoost guided the development of a logistic model, forming the basis for an AF scoring model.

Results: Records of of 26 365 AF patients were obtained from the MIMIC-IV database. The external validation dataset included 231 AF patients. The CRAMB score (Charlson comorbidity index, readmission, age, metastatic solid tumor, and maximum blood urea nitrogen concentration) outperformed the CCI and CHA_2DS_2-VASc scores, demonstrating superior predictive value for 1-year mortality. In the test set, the area under the receiver operating characteristic (AUC) for the CRAMB score was 0.765 (95% confidence interval [CI], 0.753–0.776), while in the external validation set, it was 0.582 (95% CI, 0.502–0.657).

Conclusions: The simplicity of the CRAMB score makes it user-friendly, allowing for coverage of a broader and more heterogeneous AF population.

Key words: atrial fibrillation, machine learning, mortality, predictive model, risk score

INTRODUCTION

Atrial fibrillation (AF) is a prevalent cardiac arrhythmia linked to considerable morbidity and mortality. It is characterized by an irregular and often rapid heart rate, resulting in compromised blood flow and potential complications such as stroke, heart failure, and other cardiovascular events [1]. AF has a broad impact on cardiac function, functional status, and quality of life and is also a risk factor for stroke [2]. AF becomes more prevalent with age, affecting more than 2 million individuals in the United States, 14% to 17% of whom are aged 65 years and older [3]. The prevalence of AF in the Polish population ≥65 years was estimated as 19.2% [4]. AF represents a signif-

WHAT'S NEW?

Atrial fibrillation (AF) represents a significant public health issue due to its considerable impact on morbidity and mortality as well as its economic strain on healthcare systems. Nevertheless, tools specifically designed to assess mortality risk in patients with AF are lacking. This study aimed to utilize machine learning methods for identifying pertinent variables and developing an easily applicable prognostic score to predict 1-year mortality in AF patients. By leveraging a large population dataset and employing XGBoost models for predictor screening, the CRAMB (Charlson comorbidity index, readmission, age, metastatic solid tumor, and blood urea nitrogen maximum) score was developed. The simplicity of the CRAMB score makes it user-friendly, allowing for coverage of a large and heterogeneous AF population. Moreover, the proposed model has better predictive performance than that of the clinically used CHA_2DS_2-VASc risk score for 1-year mortality among AF patients.

icant public health issue due to its considerable impact on morbidity and mortality as well as its economic strain on healthcare systems.

The assessment tool for evaluating the risk of stroke in AF patients known as CHA_2DS_2-VASc score (congestive heart failure, hypertension, age, diabetes mellitus, prior stroke or transient ischemic attack or thromboembolism, vascular disease, age, sex) [5] has been associated with cardiovascular events and mortality in diverse patient groups, including those without AF [6]. Nevertheless, tools specifically designed to assess mortality risk in AF patients are lacking. Although recent studies have introduced new AF risk scores [7, 8], these scores were developed based on data from clinical trials, limiting their applicability to the broader AF population.

Consequently, further research is necessary to identify potential models for scoring AF risk. The objective of this study was to employ machine learning methods to identify relevant variables and create an easily applicable prognostic score for predicting 1-year mortality in AF patients.

METHODS

Study design and setting

The data used in this research originated from the Medical Information Mart for Intensive Care-IV (MIMIC-IV version 2.2) database [9, 10]. Over the period from 2008 to 2019, the intensive care unit (ICU) at Beth Israel Deaconess Medical Center admitted more than 50 000 critically ill patients, as documented in MIMIC-IV. Approval for the MIMIC-IV database was granted by the Massachusetts Institute of Technology (Cambridge, MA, US) and Beth Israel Deaconess Medical Center (Boston, MA, US), with consent obtained for the initial data collection. The critical care database from China comprises comprehensive information on 2790 ICU patients, predominantly with pneumonia, admitted from January 2019 to December 2020 [11]. The database was approved by the Ethics Committee of Zigong Fourth People's Hospital (Approval Number: 2021-014) and can be accessed through the online repository "PhysioNet" with the requisite credentials [12]. The MIMIC-IV database was used for model development and testing, while the Chinese hospital ICU database was used for external validation of the model.

Study population

The study population included patients aged 18 years and older with a discharge diagnosis of AF. AF patients were identified by searching International Classification of Diseases diagnostic terminology in the MIMIC-IV database and the external validation database by matching the keyword "atrial fibrillation". The types of queried AF diagnostic terms were manually reviewed to ensure compliance. The exclusion criterion was lack of patient data on survival outcomes. In the case of MIMIC-IV records containing the same patient ID, only one record was retained with the smallest hospitalization sequence.

Study variables

The variables examined in the research included the characteristics of the study population, complications, various scores (such as the Charlson Comorbidity Index and the CHA_2DS_2-VASc score), vital signs, and an array of laboratory tests (including routine blood tests, blood biochemistry, coagulation, blood lipids, cardiac markers, etc.). Additionally, the investigation considered the use of vasopressors (norepinephrine, epinephrine, phenylephrine, dopamine, dobutamine, vasopressin, and milrinone), antithrombotic agents (heparin, enoxaparin, warfarin, aspirin, clopidogrel, ticagrelor, rivaroxaban, edoxaban, dabigatran etexilate, fondaparinux sodium, prasugrel, and apixaban), beta-blockers (propranolol, metoprolol, bisoprolol, carvedilol, labetalol, atenolol, and nebivolol) and various other data points. For laboratory tests, summary statistics, including minimum and maximum values during hospitalization, were used to derive variables. An indicator column for the respective drug was generated based on whether the drug was used during hospitalization. The variable "readmission" was derived from the variable "hospital stay sequence" for convenient clinical application. If the number of hospital admissions was greater than 1, "readmission" was assigned the value "Yes"; otherwise, it was assigned the value "No".

Outcome variable

The primary outcome measured was 1-year mortality. Survival time was calculated by using the date of death available in the MIMIC-IV database and external validation database restricted to a 1-year timeframe.

Machine learning model development and validation

The derivation dataset was randomly partitioned into training and test samples at a 3:1 ratio. To prevent model overfitting, tenfold cross-validation and model calibration techniques were applied. To accommodate varying degrees of missing values in dataset variables, the mainstream machine learning model XGBoost was employed due to its ability to handle missing data. The discriminative performance of the models was assessed using the area under the receiver operating characteristic curve. Feature scaling was deemed unnecessary before inputting the data into the model. A total of 174 candidate variables were incorporated into the model training process. Furthermore, a calibration curve was used as a graphical representation to evaluate the concordance between the predicted probabilities and observed outcomes in binary classification models. On the calibration curve, the x-axis denotes the mean predicted probability assigned by the model to a specific class, and the y-axis signifies the observed frequency of positive instances. Ideally, a well-calibrated model produces a calibration curve that closely aligns with the diagonal line (y = x), signifying a perfect correspondence between the predicted probabilities and actual outcomes.

Machine learning model interpretability

SHapley Additive exPlanations (SHAP) is a model-agnostic explainability technique that assigns importance values to features based on their contribution to a model's prediction [13]. SHAP values are grounded in SHapley values from game theory, which fairly distribute payouts based on each player's contribution to the total gain. This method ensures local accuracy, missingness, and consistency, making it versatile and reliable across different model types.

Development of the scoring scale

The XGBoost model assigned importance to predictor variables, and variables with higher importance were selected based on this ranking. These selected variables were subsequently integrated into a logistic model to construct the scoring model. Manual testing was employed to evaluate the impact of introducing or removing variables on the area under the curve (AUC) of the logistic regression model in the test set. After striking a balance between AUC performance and the increase in model complexity associated with the number of variables included, the chosen variables for the AF scoring model were ultimately determined. A nomogram was used to construct the finalized AF scores. Decision curve analysis (DCA) was employed to assess the clinical utility and net benefit of the AF scoring model, CCI, and CHA_2DS_2-VASc scores within the test set [14]. DCA quantified the net benefit of a clinical prediction model at different risk thresholds, avoiding the simplistic assumptions that all patients were at low or high risk. The superior model was identified by the highest net benefit at the chosen threshold. The flowchart of the study is shown in Supplementary material, *Figure S1*.

Data analysis

Python software (version 3.11.5) was used to construct the machine learning models, evaluate the performance, and generate the AUC and calibration curves. R software (version 4.3.2) was used for logistic and Cox regression analyses, forest plot creation, DCA, and nomogram generation. Baseline characteristics are presented as means (standard deviations), medians (IQR), or percentages (%), as determined by the distribution characteristics of the data. The DeLong test was applied to determine whether the AUC of a given prediction significantly differed from that of another prediction [15]. Python was used to make descriptive tables [16] and run the DeLong test. When constructing the original machine learning model, no handling of missing values was conducted. However, during the development of the logistic model for the AF score, missing values were removed from the dataset based on the variables included in the AF score, as logistic models are unable to manage missing values. In all analyses, statistical significance was defined as a two-sided *P*-value <0.05.

During sensitivity analysis, missing values in the original dataset were imputed. The Python library "MIDASpy" was used for data filling [17]. Additionally, hyperparameter tuning was performed on the XGBoost model to evaluate the impact of imputation and model parameter adjustments on performance. A grid search was used for hyperparameter tuning, with values for 'n_estimators' of 50, 100, 150, and 200 and values for 'max_depth' ranging from 3 to 10.

RESULTS

Baseline characteristics

This study enrolled 26 365 individuals diagnosed with AF from the MIMIC-IV database. Among the patients, 56.3% were male. The cohort had a median age of 77.0 years (with an interquartile range [IQR] of 68.0–85.0), a median CHA_2DS_2-VASc score of 4 (IQR 2–5), a median CCI of 5 (IQR 4–7), and a median hospitalization duration of 4 days (IQR 1–7). Additional results are presented in Supplementary material *Tables S1* and *S2*. The external validation dataset included 231 patients with atrial fibrillation, of whom 152 (65.8%) died. Additional findings are detailed in Supplementary material, *Table S3*.

Screening variables using the XGBoost model

The XGBoost model showed an AUC of 95% and a confidence interval (95% CI) of 0.825 (95% CI, 0.816–0.835) for the prediction of 1-year mortality in the test set (Figure 1). Figure 2 illustrates the significance of the predictor variables determined by the XGBoost model. Notably, the CCI and the presence of metastatic solid tumors were identified as the top two variables, with considerably greater importance than other variables. Supplementary material, *Figure*
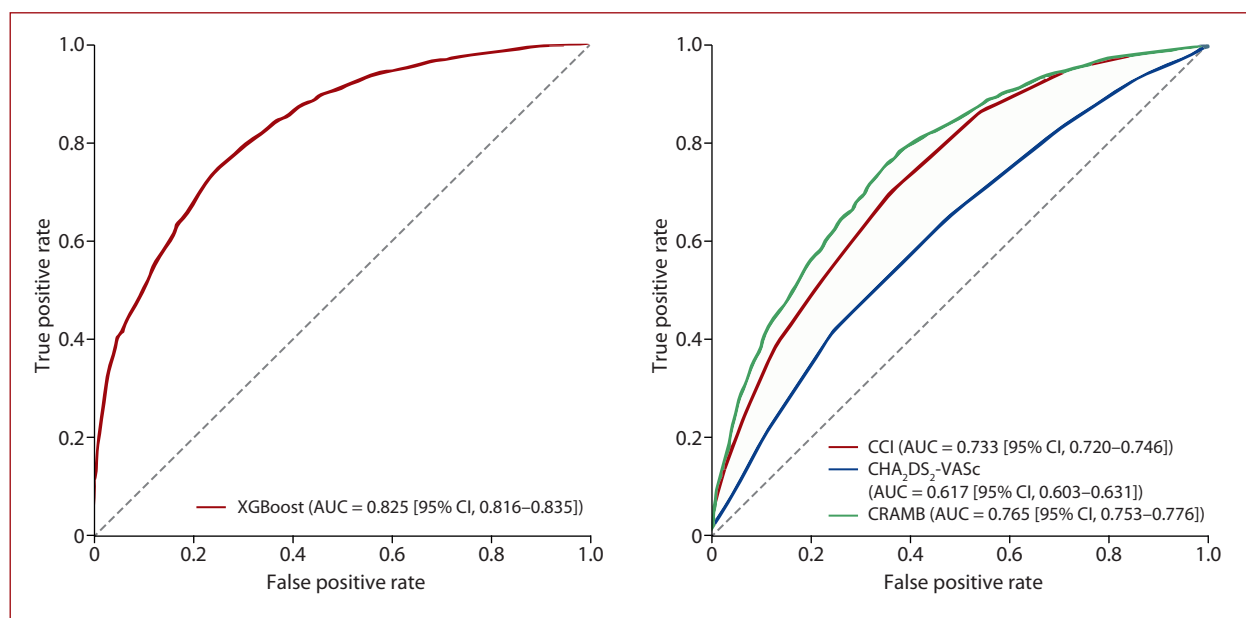
Figure 1. ROC curves for the XGBoost model in the test set and the scoring model in the test set

Abbreviations: CCI, Charlson Comorbidity Index; CHA_2DS_2-VASc score: congestive heart failure, hypertension, age, diabetes mellitus, prior stroke or transient ischemic attack or thromboembolism, vascular disease, age, sex category; CRAMB score: Charlson comorbidity index, readmission, age, metastatic solid tumor, and blood urea nitrogen maximum

S2 shows the predictor importance interpretation based on the SHAP values for the XGBoost model.

Derivation and evaluation of the AF score

The 1-year mortality risk score for AF was calculated as the CRAMB score, which represents the CCI, readmission, age, metastatic solid tumor, and maximum blood urea nitrogen (BUN) (Figure 2). Logistic and Cox regression analyses were employed to assess the predictive value of these five variables for the outcome of death and were expressed as odds ratios (ORs) and hazard ratios (HRs). Both the forest plot of ORs (Figure 3) and the forest plot of HRs (Figure 4) demonstrated that these variables were significantly different.

A nomogram was used to calculate the CRAMB score (Supplementary material, *Figure S3*). In the test set, the AUC for the CRAMB score was 0.765 (95% CI, 0.753–0.776), surpassing the CCI at 0.733 (95% CI, 0.720–0.746) and the CHA_2DS_2-VASc score at 0.617 (95% CI, 0.603–0.631) (Figure 1). The sensitivity analysis showed that hyperparameter adjustment and missing value filling had very little impact on the AUC of XGBoost and the different scoring models (Supplementary material, *Figure S4*). Table 1 displays supplementary performance metrics corresponding to these scores. The DeLong test results comparing the CRAMB score with existing scores (CCI and CHA_2DS_2-VASc) showed statistically significant differences (P <0.001), as indicated in Table 1. The DCA results provided in Figure 5 demonstrate that the CRAMB score consistently exhibited a positive and greater net benefit across the entire threshold range than did the default strategies, assuming either high or low risk, as indicated by the CCI and CHA_2DS_2-VASc scores, and the hypothesis of not using a scoring system. The calibration

plot (Supplementary material, *Figure S5*) for the test set indicated that the CRAMB score was well calibrated.

Model evaluation on the external validation set

In the external validation set, the AUC for the CRAMB score was 0.582 (95% CI, 0.502–0.657), which surpassed that of the CCI (0.542 [95% CI, 0.469–0.618]) and that of the CHA_2DS_2-VASc score (0.511 [95% CI, 0.438–0.586]) (Supplementary material, *Figure S6*). Additional findings are detailed in Supplementary material, *Table S4*. Decision curve analysis showed that the positive return of the CRAMB score exceeded that of the other two scores between the threshold probabilities of 60%–80% (Supplementary material, *Figure S7*).

DISCUSSION

Main findings

This study's primary contribution is establishing a benchmark for using machine learning models in the construction of AF scores for mortality prediction. This study introduces and validates a novel risk score for assessing the 1-year mortality risk in AF patients. By leveraging a large-sample population dataset and employing XGBoost models for predictor screening, we developed the CRAMB score (Charlson comorbidity index, readmission, age, metastatic solid tumor, and blood urea nitrogen maximum). XGBoost excels at variable selection by effectively capturing nonlinear relationships and handling missing data [18]. Its built-in feature importance mechanism automatically identifies key variables, a capability lacking in logistic regression. Furthermore, compared with logistic regression, XGBoost's
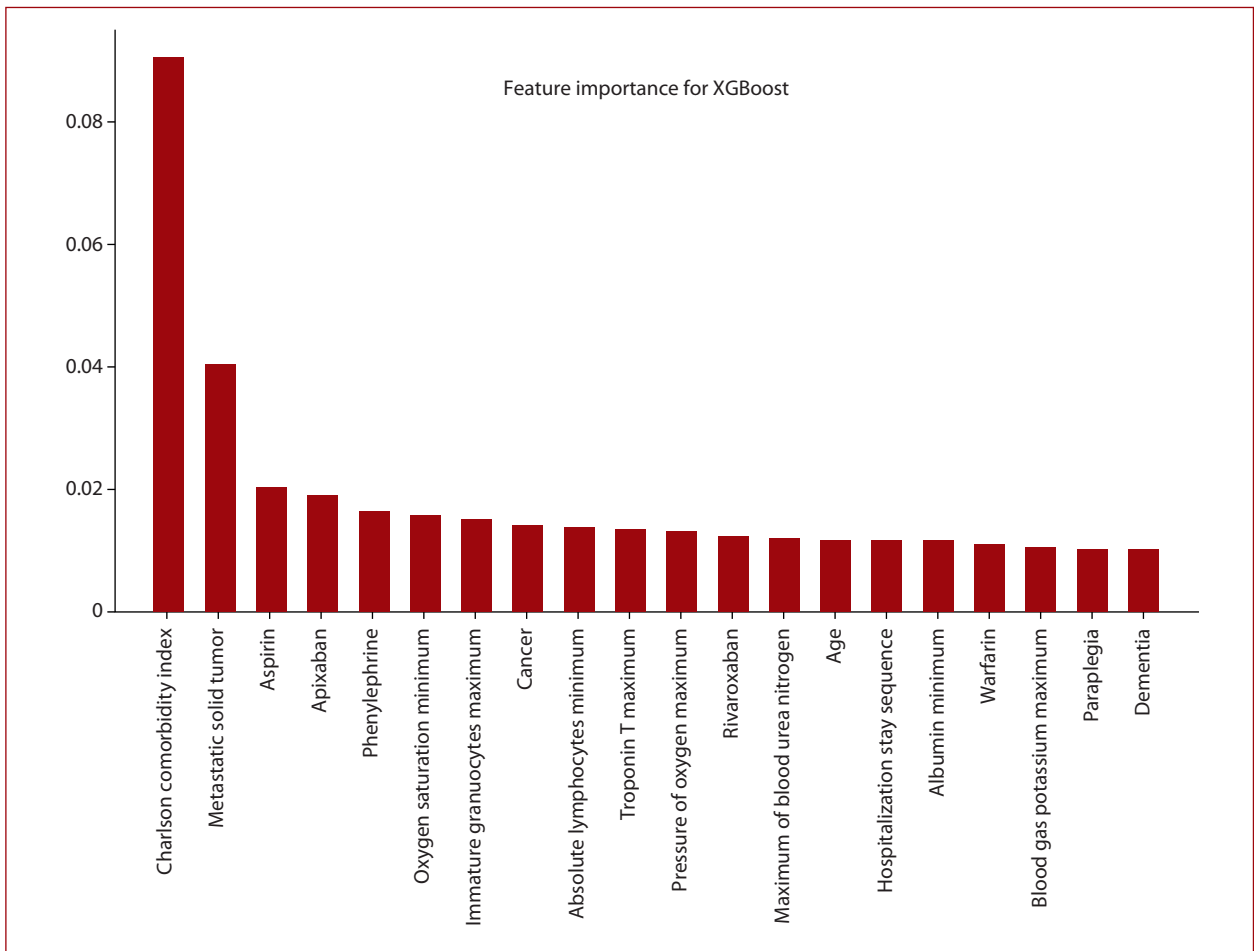
Figure 2. Feature importance values of the XGBoost model in the training set
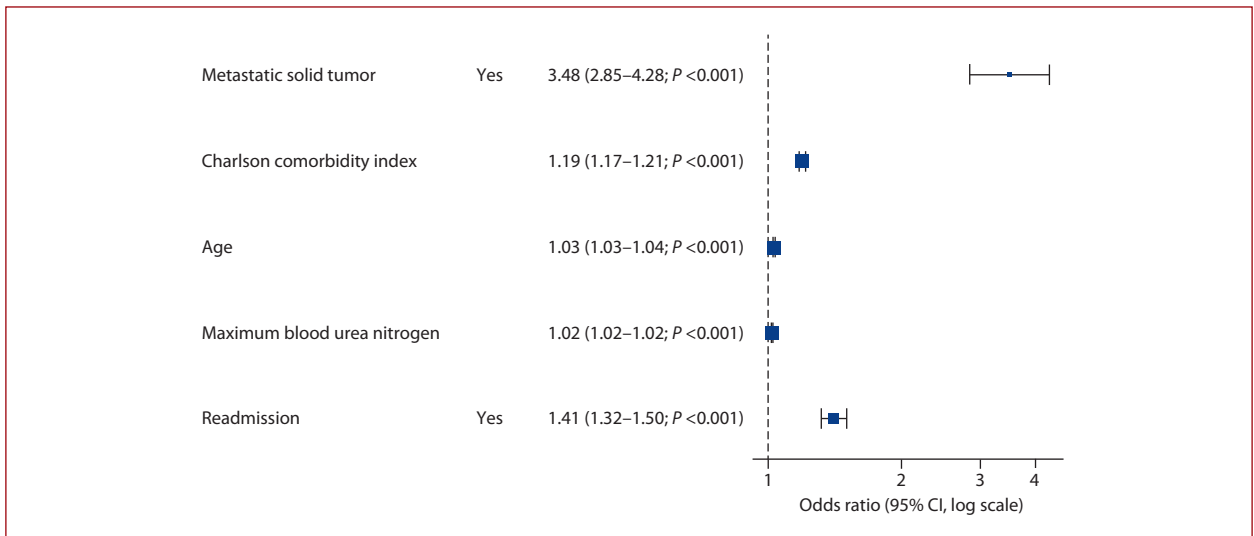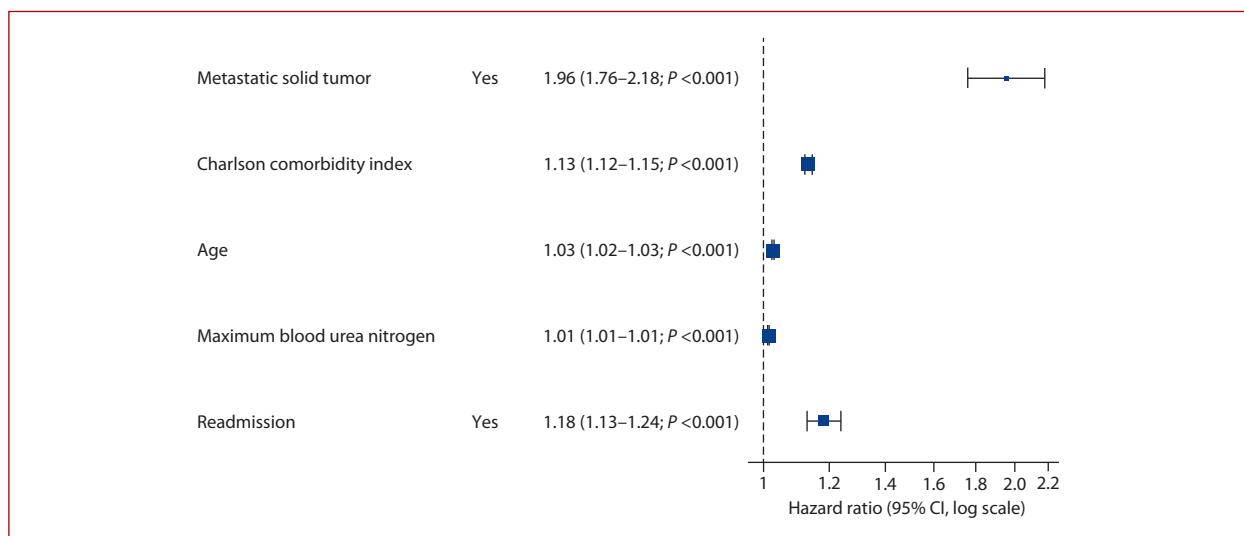

Figure 3. Forest plot of the logistic model (CRAMB score) for predicting 1-year mortality in the training set

Figure 4. Forest plot illustrating the ability of the Cox regression model to predict 1-year mortality in the training cohort stratified by the CRAMB score

Table 1. Comparison of the predictive performance of the scores in the test set and DeLong test. The *P*-value for the DeLong test was obtained by comparing the area under the curve (AUC) of the CRAMB score with that of the corresponding score

Item	Accuracy	Sensitivity	Specificity	ROC AUC	DeLong test *P*-value
Charlson Comorbidity Index	0.692	0.375	0.876	0.733	<0.001
CHA$_2$DS$_2$-VASc	0.635	0.068	0.963	0.617	<0.001
CRAMB	0.715	0.438	0.876	0.765	–

Abbreviations: CHA$_2$DS$_2$-VASc score: congestive heart failure, hypertension, age, diabetes mellitus, prior stroke or transient ischemic attack or thromboembolism, vascular disease, age, sex category; CRAMB score: Charlson comorbidity index, readmission, age, metastatic solid tumor, and maximum blood urea nitrogen
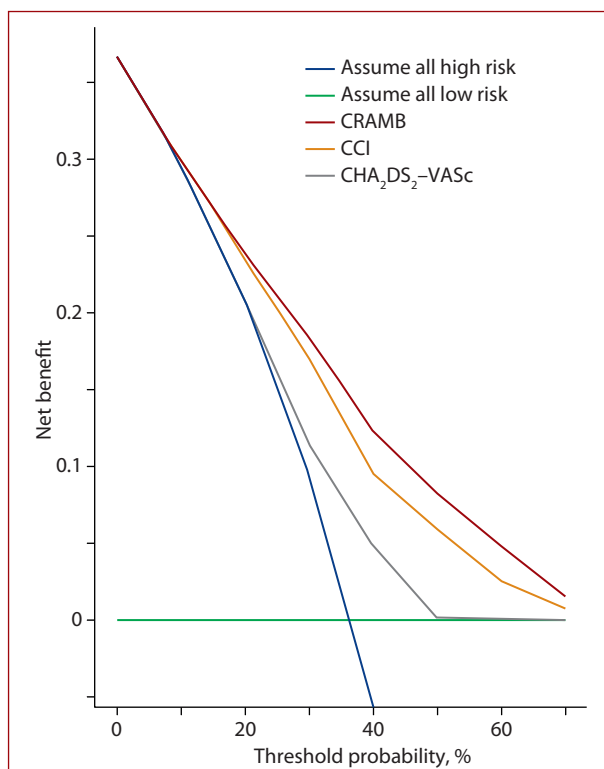
Figure 5. Decision curve analysis of various scores in the test set

Abbreviations: CCI, Charlson comorbidity index; CRAMB score: Charlson comorbidity index, readmission, age, metastatic solid tumor, and blood urea nitrogen maximum

ensemble learning often results in superior predictive performance, and its regularization techniques boost resilience against overfitting, making it a robust choice for predictive modeling and variable selection. The variables incorporated in the CRAMB score were validated through logistic and Cox regression analyses, demonstrating their predictive significance for mortality. The CRAMB score exhibited excellent calibration, and DCA illustrated its clinical utility. Importantly, the findings of this study demonstrated that the CRAMB score outperformed the widely used CHA$_2$DS$_2$-VASc risk score in predicting mortality despite the latter's original focus on predicting ischemic stroke.

Predictors of death in AF patients

Predictors and risk factors for death in AF patients span a broad spectrum of clinical and demographic variables. Hypertension has been identified as a significant risk factor for incident heart failure and all-cause mortality in AF patients [19]. Moreover, patients with chronic kidney disease who develop AF face an increased risk of stroke and death [20], and renal function has been associated with the risk of stroke and bleeding in AF patients [21]. Additionally, age is correlated with elevated risks of stroke and mortality in patients with either AF or sinus rhythm [22]. Proposed factors such as cancer-related inflammation, anticancer treatments, and other comorbidities associated with cancer are believed to influence atrial remodeling, potentially

increasing the susceptibility of cancer patients to AF [23]. Therefore, AF screening is important to reduce the burden of AF-associated stroke [24].

Comparison with similar studies

Compared to the ABC-death (age, biomarkers [N-terminal pro B-type natriuretic peptide, troponin T, growth differentiation factor-15]) risk score [7] and BASIC-AF risk score (biomarkers, age, ultrasound, ventricular conduction delay, and clinical history) [8], the CRAMB score was constructed based on the MIMIC-IV database, leading to significant differences in population characteristics compared to clinical trial populations. Therefore, this study addresses a gap in the development of scoring methods and screening predictor variables within a broader population than previous studies of this nature. Future research on AF scores should focus on the characteristics of the population used for score development, comprehensively considering the importance and applicability of the variables included.

Expanding the clinical application potential of the CRAMB score

For effective integration into clinical workflows, the CRAMB score should be incorporated into electronic health records for automated calculation and routine assessments during admissions and outpatient visits. In addition, the nomogram can also be turned into an online tool for automatic calculation. Training clinical staff on its use, interpretation, and communication with patients is essential. Integrating the score into clinical decision support systems and multidisciplinary team meetings will enhance patient management. Pilot programs and continuous outcome monitoring will refine its application, ensuring robust and effective use, ultimately improving patient care and optimizing use of resources.

Limitations

The main limitation of this study is the limited representativeness of the external validation set. Future studies should validate the model using datasets from multiple medical centers or international sources to enhance generalizability. Conducting a prospective cohort study to assess the predictive power of the CRAMB score would provide stronger evidence of its efficacy, as prospective data collection allows for better control of variables and reduces retrospective biases. Additionally, incorporating a broader set of variables, such as lifestyle factors and detailed medication history, could improve the model's accuracy and relevance. Finally, the CRAMB score was developed using data from a specific period, which may not reflect current medical practices. As healthcare evolves, this could affect the score's relevance. To address this, we should regularly update the model with recent data to maintain its accuracy and relevance. Continuous recalibration and validation will ensure that the CRAMB score reflects current practices and

improves patient care and outcomes in a dynamic healthcare environment.

CONCLUSIONS

This study's primary contribution is establishing a benchmark for using machine learning models in the construction of a score for mortality prediction in AF patients. By leveraging a large-sample population dataset and employing XGBoost models for predictor screening, we developed the CRAMB score (CCI, readmission, age, metastatic solid tumor, and blood urea nitrogen maximum). The simplicity of the CRAMB score makes it user-friendly, allowing for the coverage of a broad and heterogeneous AF population. Moreover, the proposed model has superior predictive performance compared to that of the currently used CHA_2DS_2-VASc risk score for 1-year mortality in AF patients. External validation of the CRAMB score in new datasets has potential value for enhancing clinical practice.

Supplementary material

Supplementary material is available at https://journals. viamedica.pl/polish_heart_journal.

Article information

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): 154–162, doi: 10.1016/S0140-6736(14)61774-8, indexed in Pubmed: 25960110.
2. Westerman S, Wenger N. Gender differences in atrial fibrillation: A review of epidemiology, management, and outcomes. Curr Cardiol Rev. 2019; 15(2): 136–144, doi: 10.2174/1573403X15666181205110624, indexed in Pubmed: 30516110.
3. Jacobs LG. The sin of omission: A systematic review of antithrombotic therapy to prevent stroke in atrial fibrillation. J Am Geriatr Soc. 2001; 49(1): 91–94, doi: 10.1046/j.1532-5415.2001.49016.x, indexed in Pubmed: 11207849.
4. Kalarus Z, Średniawa B, Mitręga K, et al. Prevalence of atrial fibrillation in the 65 or over Polish population. Report of cross-sectional NOMED-AF study. Kardiol Pol. 2023; 81(1): 14–21, doi: 10.33963/kp.a2022.0202, indexed in Pubmed: 36043418.
5. January CT, Wann LS, Calkins H, 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 in Collaboration With the Society of Thoracic Surgeons. Circulation. 2019; 140(2): e125–e151, doi: 10.1161/CIR.0000000000000665, indexed in Pubmed: 30686041.
6. Renda G, Ricci F, Patti G, et al. CHA2DS2VASc score and adverse outcomes in middle-aged individuals without atrial fibrillation. Eur J Prev Cardiol.

2019; 26(18): 1987–1997, doi: 10.1177/2047487319868320, indexed in Pubmed: 31409109.

7. Hijazi Z, Oldgren J, Lindbäck J, 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(6): 477–485, doi: 10.1093/eurheartj/ehx584, indexed in Pubmed: 29069359.

8. Samaras A, Kartas A, Akrivos E, et al. A novel prognostic tool to predict mortality in patients with atrial fibrillation: The BASIC-AF risk score. Hellenic J Cardiol. 2021; 62(5): 339–348, doi: 10.1016/j.hjc.2021.01.007, indexed in Pubmed: 33524615.

9. Johnson AEW, Bulgarelli L, Shen Lu, et al. MIMIC-IV, a freely accessible electronic health record dataset. Sci Data. 2023; 10(1): 1, doi: 10.1038/s41597-022-01899-x, indexed in Pubmed: 36596836.

10. Johnson A, Bulgarelli L, Pollard T, et al. MIMIC-IV (version 2.2). PhysioNet. 2023, doi: 10.13026/6mm1-ek67.

11. Xu P, Chen L, Zhu Y, et al. Critical care database comprising patients with infection. Front Public Health. 2022; 10: 852410, doi: 10.3389/fpubh.2022.852410, indexed in Pubmed: 35372245.

12. Xu P, Chen L, Zhang Z. Critical care database comprising patients with infection at Zigong Fourth People's Hospital (version 1.1). PhysioNet., doi: 10.13026/xpt9-z726.

13. Lundberg SM, Lee SIA. unified approach to interpreting model predictions. Advances in neural information processing systems. ; 2017: 30.

14. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making. 2006; 26(6): 565–574, doi: 10.1177/0272989X06295361, indexed in Pubmed: 17099194.

15. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. Biometrics. 1988; 44(3): 837–845, doi: 10.2307/2531595, indexed in Pubmed: 3203132.

16. Pollard TJ, Johnson AEW, Raffa JD, et al. An open source Python package for producing summary statistics for research papers. JAMIA Open.

2018; 1(1): 26–31, doi: 10.1093/jamiaopen/ooy012, indexed in Pubmed: 31984317.

17. Lall R, Robinson T. Efficient multiple imputation for diverse data in Python and: MIDASpy and rMIDAS. J Statistical Software. 2023; 107(9): 1–38, doi: 10.18637/jss.v107.i09.

18. Chen T, Guestrin C. XGBoost. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Association for Computing Machinery 2016: 785–794.

19. Middeldorp ME, Ariyaratnam JP, Kamsani SH, et al. Hypertension and atrial fibrillation. J Hypertens. 2022; 40(12): 2337–2352, doi: 10.1097/HJH.0000000000003278, indexed in Pubmed: 36204994.

20. Carrero JJ, Trevisan M, Sood MM, et al. Incident atrial fibrillation and the risk of stroke in adults with chronic kidney disease: the stockholm creatinine measurements (SCREAM) project. Clin J Am Soc Nephrol. 2018; 13(9): 1314–1320, doi: 10.2215/CJN.04060318, indexed in Pubmed: 30030271.

21. Bonde AN, Lip GYH, Kamper AL, et al. Renal function and the risk of stroke and bleeding in patients with atrial fibrillation: An observational cohort study. Stroke. 2016; 47(11): 2707–2713, doi: 10.1161/STROKEAHA.116.014422, indexed in Pubmed: 27758943.

22. Xing Y, Sun Y, Li H, et al. CHADS-VASc score as a predictor of long-term cardiac outcomes in elderly patients with or without atrial fibrillation. Clin Interv Aging. 2018; 13: 497–504, doi: 10.2147/CIA.S147916, indexed in Pubmed: 29636604.

23. Chu G, Versteeg HH, Verschoor AJ, et al. Atrial fibrillation and cancer — An unexplored field in cardiovascular oncology. Blood Rev. 2019; 35: 59–67, doi: 10.1016/j.blre.2019.03.005, indexed in Pubmed: 30928168.

24. Boriani G, Imberti JF, Vitolo M. Screening for atrial fibrillation: Different approaches targeted to reduce ischemic stroke. Kardiol Pol. 2023; 81(1): 1–3, doi: 10.33963/KP.a2022.0281, indexed in Pubmed: 36475515.