

# Leveraging machine learning for enhanced mortality risk prediction in atrial fibrillation: A step towards precision medicine?

Yang Chen<sup>1,2</sup>, Yalin Zheng<sup>1,3</sup>, Gregory YH Lip<sup>1,4</sup>

<sup>1</sup>Liverpool Centre for Cardiovascular Science at the University of Liverpool, Liverpool John Moores University, and Liverpool Heart & Chest Hospital, Liverpool, United Kingdom

<sup>2</sup>Department of Cardiovascular and Metabolic Medicine, Institute of Life Course and Medical Sciences, University of Liverpool, Liverpool, United Kingdom

<sup>3</sup>Department of Eye and Vision Sciences, Institute of Life Course and Medical Sciences, University of Liverpool, Liverpool, United Kingdom

<sup>4</sup>Danish Centre for Health Services Research, Department of Clinical Medicine, Aalborg University, Aalborg, Denmark

## Related article

by Wang et al.

### Correspondence to:

Prof. Gregory Y H Lip, MD, PhD,  
Liverpool Centre  
for Cardiovascular Science  
at the University of Liverpool,  
Liverpool John Moores University  
and Liverpool Heart & Chest  
Hospital,  
Thomas Drive L14 3PE Liverpool,  
United Kingdom,  
phone: 0151 794 91 66,  
e-mail:  
gregory.lip@liverpool.ac.uk

Copyright by the Author(s), 2024

DOI: 10.33963/v.phj.102455

**Received:** September 5, 2024

**Accepted:** September 5, 2024

**Early publication date:**  
September 5, 2024

Despite significant efforts in investigating the pathophysiology of atrial fibrillation (AF) and improving treatments, AF remains the most common cardiac arrhythmia. To date, limited studies have assessed trends of all-cause mortality in contemporary AF patients. These data have shown a small increase in age-standardized all-cause mortality in AF patients (increased by 2.0% between 1990 and 2019) [1] and no significant improvement in 1-year all-cause mortality in AF patients recruited in 2007 and 2016, respectively (8.0% in 2007 vs. 7.8% in 2016) [2]. Such trends highlight that risk management of mortality in patients with AF remains a concern.

With the development of computational technologies, machine learning (ML) approaches are increasingly being applied in AF-related fields. Compared to traditional regression models, ML models have the ability to handle a large number of variables, even if there is an intrinsic correlation between these variables [3]. This enables ML models to identify some non-traditional or previously unidentified risk factors and accurately assess their relative importance in predicting outcomes. However, despite the growing interest in ML, there is a notable lack of models tailored to 1-year all-cause mortality in AF patients.

In this edition of the *Polish Heart Journal*, Wang et al. [4] developed a risk-scoring system to predict 1-year all-cause mortality in AF patients, the CRAMB score, using the eXtreme

Gradient Boosting (XGBoost) model. The study included 26 365 AF patients from the Medical Information Mart for Intensive Care IV (MIM-IC-IV, version 2.2) database as a derivation cohort, which was randomly allocated to training and test cohorts in a 3:1 ratio. In the training cohort, features that were important for the endpoint were pre-screened using the XGBoost model. Charlson Comorbidity Index (CCI), readmission status, age, metastatic solid tumor, and maximum value of blood urea nitrogen were finally imported to construct the XGBoost model. Subsequently, the performance of this model was evaluated in the test cohort and the external validation cohort (from the Chinese intensive care unit database), respectively. Then, the contribution of the different features for predicting the target endpoint was assessed. The results showed that the CRAMB score had an area under the curve (AUC) of 0.765 (95% confidence interval [CI], 0.753–0.776) in the test cohort, which was significantly better than the CCI (AUC 0.733; 95% CI, 0.720–0.746) and CHA<sub>2</sub>DS<sub>2</sub>-VASc (AUC 0.617; 95% CI, 0.603–0.631). In the external validation cohort, the CRAMB score had an AUC of 0.582 (95% CI, 0.502–0.657), which was also superior to the CHA<sub>2</sub>DS<sub>2</sub>-VASc. In addition, decision curve analysis showed that the net benefit of the CRAMB score exceeded that of CHA<sub>2</sub>DS<sub>2</sub>-VASc. This suggests a potential advantage of the CRAMB score in clinical applications for more accurate assessment of

the 1-year risk of all-cause mortality in AF patients than the CHA<sub>2</sub>DS<sub>2</sub>-VASc score.

While traditional risk assessment scores have helped clinicians identify the long-term risk of all-cause mortality in AF patients, the article by Wang et al. provides a good opportunity to discuss whether ML models, like CRAMB score, are more effective compared to conventional clinical risk scores. It may not be completely fair to compare the most widely used clinical scores, CHADS<sub>2</sub> and CHA<sub>2</sub>DS<sub>2</sub>-VASc scores, which were initially used to predict thromboembolic risk in AF patients.

However, other scores have been proposed combining CHADS<sub>2</sub> and CHA<sub>2</sub>DS<sub>2</sub>-VASc with estimated glomerular filtration rate and creatinine clearance have shown some promise, albeit with the best-performing one only having an AUC of 0.734 in predicting 1-year all-cause mortality in AF patients [5].

In addition, the GARFIELD-AF risk score had an AUC of 0.77 for identifying 1-year all-cause mortality [6]. However, the complexity of GARFIELD-AF requires more comprehensive data inputs, which, on the one hand, may limit the application of the model in real clinical settings and, on the other hand, may lead to instability in model predictions, especially when confronted with unbalanced data. This highlights the necessity for simpler yet effective predictive tools that are easily implemented in routine practice to enhance clinical decision-making.

In contrast, the CRAMB score offers a straightforward approach with only five key features. Its simplicity reduces the burden of data collection and potentially makes it more practical for clinical applications. Notably, the CRAMB score includes predictors such as CCI and readmission status, which are not part of CHA<sub>2</sub>DS<sub>2</sub>-VASc, providing additional insights that could guide future model enhancements. The CRAMB score also demonstrated strong performance in the test cohort, achieving an AUC of 0.765, which surpassed that of CHA<sub>2</sub>DS<sub>2</sub>-VASc. This underscores the potential of ML models to not only simplify prediction but also deliver predictive accuracy on par with, or even exceeding, traditional clinical scores.

Although the CRAMB score shows potential, its use in clinical practice still has a long way to go. One of the main issues is the very poor performance of the CRAMB score in external validation from a different ethnicity cohort, with an AUC of only 0.582. This suggests that the model may not generalize well across populations, particularly those with different ethnic backgrounds compared to the derived cohort.

Ethnicity is now generally recognized as an important factor in the epidemiology as well as in the management of AF, with recognized ethnic differences in AF-related complications [7, 8]. Increasing focus has also been on sex differences in risks, with implications for clinical risk stratification [9–11]. These issues highlight the possible need for ethnicity-specific ML models for AF mortality to ensure their reliability and utility in real-world settings [12].

Another limitation is that the CRAMB score relies on the XGBoost algorithm for both feature selection and model construction. While XGBoost is a powerful ML algorithm known for its high accuracy and efficiency, relying on it alone may bias feature selection and may miss opportunities to compare performance with other ML algorithms (e.g., neural network, random forest) that may have complementary strengths or show different insights [13]. The reliance on metastatic solid tumors and maximum blood urea nitrogen that are not easy to obtain may also limit the usability of the CRAMB score. Finally, risk is not static but dynamic, changing with age and incident comorbidities, and ML models need to account for the dynamic nature of changing risk factors [14].

Overall, the CRAMB score presents a simplified ML model for predicting 1-year all-cause mortality in AF patients with only five key features. However, the poor external validation performance of this method highlights the problem of its generalizability, especially in ethnically diverse populations (as we previously reported [15]). Future studies should aim to validate these models in ethnically diverse cohorts and explore a wider range of ML algorithms, including recent foundation models, to improve predictive accuracy and clinical relevance in AF management.

## Article information

**Conflict of interest:** None declared.

**Funding:** None.

**Open access:** This article is available in open access under Creative Commons Attribution-Non-Commercial-No Derivatives 4.0 International (CC BY-NC-ND 4.0) license, which allows downloading and sharing articles with others as long as they credit the authors and the publisher, but without permission to change them in any way or use them commercially. For commercial use, please contact the journal office at [polishheartjournal@ptkardio.pl](mailto:polishheartjournal@ptkardio.pl)

## REFERENCES

- Dong XJ, Wang BB, Hou FF, et al. Global burden of atrial fibrillation/atrial flutter and its attributable risk factors from 1990 to 2019. *Europace*. 2023; 25(3): 793–803, doi: [10.1093/europace/euac237](https://doi.org/10.1093/europace/euac237), indexed in PubMed: [36603845](https://pubmed.ncbi.nlm.nih.gov/36603845/).
- Singh SM, Abdel-Qadir H, Pang A, et al. Population trends in all-cause mortality and cause specific-death with incident atrial fibrillation. *J Am Heart Assoc*. 2020;9(19):e016810, doi: [10.1161/JAHA.120.016810](https://doi.org/10.1161/JAHA.120.016810), indexed in PubMed: [32924719](https://pubmed.ncbi.nlm.nih.gov/32924719/).
- Haug CJ, Drazen JM. Artificial intelligence and machine learning in clinical medicine, 2023. *N Engl J Med*. 2023;388(13): 1201–1208, doi: [10.1056/NEJMra2302038](https://doi.org/10.1056/NEJMra2302038), indexed in PubMed: [36988595](https://pubmed.ncbi.nlm.nih.gov/36988595/).
- Wang B, Jin F, Cao H, et al. Development of a risk score for predicting one-year mortality in patients with atrial fibrillation using XGBoost-assisted feature selection. *Pol Heart J*. 2024;82(10): 941–948, doi: [10.33963/v.phj.101842](https://doi.org/10.33963/v.phj.101842), indexed in PubMed: [39140655](https://pubmed.ncbi.nlm.nih.gov/39140655/).
- Cemin R, Maggioni AP, Gonzini L, et al. Simple scores to predict 1-year mortality in atrial fibrillation. *J Cardiovasc Med (Hagerstown)*. 2024; 25(4): 271–279, doi: [10.2459/JCM.0000000000001602](https://doi.org/10.2459/JCM.0000000000001602), indexed in PubMed: [38488064](https://pubmed.ncbi.nlm.nih.gov/38488064/).
- Fox KAA, Lucas JE, Pieper KS, 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(12): e017157, doi: [10.1136/bmjopen-2017-017157](https://doi.org/10.1136/bmjopen-2017-017157), indexed in PubMed: [29273652](https://pubmed.ncbi.nlm.nih.gov/29273652/).

7. Kang DS, Yang PS, Kim D, et al. Racial differences in bleeding risk: An ecological epidemiological study comparing Korea and United Kingdom subjects. *Thromb Haemost.* 2024; 124(9): 842–851, doi: [10.1055/a-2269-1123](https://doi.org/10.1055/a-2269-1123), indexed in Pubmed: [38359877](https://pubmed.ncbi.nlm.nih.gov/38359877/).
8. Kang DS, Yang PS, Kim D, et al. Racial differences in ischemic and hemorrhagic stroke: An ecological epidemiological study. *Thromb Haemost.* 2024; 124(9): 883–892, doi: [10.1055/a-2278-8769](https://doi.org/10.1055/a-2278-8769), indexed in Pubmed: [38423097](https://pubmed.ncbi.nlm.nih.gov/38423097/).
9. Teppo K, Airaksinen KE, Jaakkola J, et al. Ischaemic stroke in women with atrial fibrillation: temporal trends and clinical implications. *Eur Heart J.* 2024; 45(20): 1819–1827, doi: [10.1093/eurheartj/ehae198](https://doi.org/10.1093/eurheartj/ehae198), indexed in Pubmed: [38606837](https://pubmed.ncbi.nlm.nih.gov/38606837/).
10. Teppo K, Lip GY, Airaksinen KE, et al. Comparing CHA2DS2-VA and CHA2DS2-VASc scores for stroke risk stratification in patients with atrial fibrillation: A temporal trends analysis from the retrospective Finnish AntiCoagulation in Atrial Fibrillation (FinACAF) cohort. *Lancet Reg Health Eur.* 2024; 43: 100967, doi: [10.1016/j.lanep.2024.100967](https://doi.org/10.1016/j.lanep.2024.100967), indexed in Pubmed: [39171253](https://pubmed.ncbi.nlm.nih.gov/39171253/).
11. Corica B, Lobban T, True Hills M, et al. Sex as a risk factor for atrial fibrillation-related stroke. *Thromb Haemost.* 2024; 124(4): 281–285, doi: [10.1055/s-0043-1776394](https://doi.org/10.1055/s-0043-1776394), indexed in Pubmed: [37871631](https://pubmed.ncbi.nlm.nih.gov/37871631/).
12. Norby FL, Benjamin EJ, Alonso A, et al. Racial and ethnic considerations in patients with atrial fibrillation: JACC focus seminar 5/9. *J Am Coll Cardiol.* 2021; 78(25): 2563–2572, doi: [10.1016/j.jacc.2021.04.110](https://doi.org/10.1016/j.jacc.2021.04.110), indexed in Pubmed: [34887142](https://pubmed.ncbi.nlm.nih.gov/34887142/).
13. Suri JS, Bhagawati M, Paul S, et al. Understanding the bias in machine learning systems for cardiovascular disease risk assessment: The first of its kind review. *Comput Biol Med.* 2022; 142: 105204, doi: [10.1016/j.combiomed.2021.105204](https://doi.org/10.1016/j.combiomed.2021.105204), indexed in Pubmed: [35033879](https://pubmed.ncbi.nlm.nih.gov/35033879/).
14. Lip GYH, Genaidy A, Tran G, et al. Improving stroke risk prediction in the general population: A comparative assessment of common clinical rules, a new multimorbid index, and machine-learning-based algorithms. *Thromb Haemost.* 2022; 122(1): 142–150, doi: [10.1055/a-1467-2993](https://doi.org/10.1055/a-1467-2993), indexed in Pubmed: [33765685](https://pubmed.ncbi.nlm.nih.gov/33765685/).
15. Chen Y, Gue Y, Calvert P, et al. Predicting stroke in Asian patients with atrial fibrillation using machine learning: A report from the KERALA-AF registry, with external validation in the APHRS-AF registry. *Curr Probl Cardiol.* 2024; 49(4): 102456, doi: [10.1016/j.cpcardiol.2024.102456](https://doi.org/10.1016/j.cpcardiol.2024.102456), indexed in Pubmed: [38346609](https://pubmed.ncbi.nlm.nih.gov/38346609/).