

Development of a Novel Risk Score for Predicting One-Year Mortality Risk in Patients with Atrial Fibrillation using XGBoost-Assisted Feature Selection

Bin Wang^{1,#}, MD; Feifei Jin^{2,3,4,#}, MD; Han Cao^{5,#}, MD; Qing Li¹, MD; Ping Zhang^{1,*}, MD

¹Department of Cardiology, Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua University, Beijing, China.

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

³Key Laboratory of Trauma treatment and Neural Regeneration, Peking University, Ministry of Education, Beijing, China.

⁴National Center for Trauma Medicine of China, Beijing, China.

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

These authors contributed equally.

*** Corresponding Author:**

Ping Zhang

Department of Cardiology, Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua University, No. 168 Litang Road, Changping District, Beijing, 102218, China

Email: zhpdoc@126.com

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

Short title

A Novel Risk Score for Atrial Fibrillation

ORCID

Bin Wang <http://orcid.org/0000-0003-0012-9835>

Feifei Jin <https://orcid.org/0000-0002-4991-0158>

Han Cao <https://orcid.org/0000-0002-9491-6188>

Qing Li <https://orcid.org/0000-0001-8218-7929>

Word count: 5572

Total number of words in the manuscript: 2981

Number of words in the abstract: 380

Number of figures: 8

Number of tables: 2

1 **Development of a Novel Risk Score for Predicting One-Year Mortality Risk in**
2 **Patients with Atrial Fibrillation using XGBoost-Assisted Feature Selection**

3

4 **ABSTRACT**

5 **Background:** There is a lack of tools specifically designed to assess mortality risk in
6 patients with atrial fibrillation (AF). The aim of this study was to utilize machine
7 learning methods for identifying pertinent variables and developing an easily
8 applicable prognostic score to predict 1-year mortality in AF patients.

9 **Methods:** This single-center retrospective cohort study based on the Medical
10 Information Mart for Intensive Care-IV (MIMIC-IV) database focused on patients
11 aged 18 years and older with AF. The study thoroughly scrutinized patient data to
12 identify and analyze variables, encompassing demographic variables, comorbidities,
13 scores, vital signs, laboratory test results, and medication usage. The variable
14 importance from XGBoost guided the development of a logistic model, forming the
15 basis for an AF scoring model. Decision curve analysis was used to compare the AF
16 score with other scores. Python and R software were used for data analysis.

17 **Results:** A cohort of 59,595 AF patients was obtained from the MIMIC-IV database;
18 these patients were predominantly elderly (median age 77.3 years) and male (55.6%).
19 The XGBoost model effectively predicted 1-year mortality (Area under the curve
20 (AUC): 0.833; 95% confidence intervals: 0.826-0.839), underscoring the significance
21 of the Charlson Comorbidity Index (CCI) and the presence of metastatic solid tumors.

22 The CRAMB score (Charlson comorbidity index, readmission, age, metastatic solid
23 tumor, and blood urea nitrogen maximum) outperformed the CCI and
24 CHA2DS2-VASc scores, demonstrating superior predictive value for 1-year mortality.
25 In the test set, the area under the ROC curve (AUC) for the CRAMB score was 0.756
26 (95% confidence intervals: 0.748-0.764), surpassing the CCI score of 0.720 (95%
27 confidence intervals: 0.712-0.728) and the CHA2DS2-VASc score of 0.609 (95%
28 confidence intervals: 0.600-0.618). Decision curve analysis revealed that the CRAMB
29 score had a consistently positive effect and greater net benefit across the entire
30 threshold range than did the default strategies and other scoring systems. The
31 calibration plot for the test set indicated that the CRAMB score was well calibrated.

32 **Conclusions:** This study's primary contribution is the establishment of a benchmark
33 for utilizing machine learning models in construction of a score for mortality
34 prediction in AF. The CRAMB score was developed by leveraging a large-sample
35 population dataset and employing XGBoost models for predictor screening. The
36 simplicity of the CRAMB score makes it user friendly, allowing for coverage of a
37 broader and more heterogeneous AF population.

38

39 **Keywords:** Atrial fibrillation; Mortality; Risk score; Machine learning; Risk factors

40

41

42

43 **List of abbreviations**

44 **ABC-death:** Age, biomarkers (N-terminal pro B-type natriuretic peptide, troponin T,

45 growth differentiation factor-15), and clinical history of heart failure

46 **AF:** Atrial fibrillation

47 **AUC:** Area under the curve

48 **BASIC-AF risk score:** Biomarkers, age, ultrasound, intraventricular conduction

49 delay, and clinical history

50 **BUN:** blood urea nitrogen

51 **CCI:** Charlson Comorbidity Index

52 **CHA2DS2-VASc:** congestive heart failure, hypertension, age, diabetes mellitus, prior

53 stroke or TIA or thromboembolism, vascular disease, age, sex category

54 **CI:** Confidence interval

55 **CRAMB:** Charlson comorbidity index, readmission, age, metastatic solid tumor, and

56 maximum blood urea nitrogen

57 **DCA:** Decision curve analysis

58 **HR:** hazard ratio

59 **ICU:** Intensive care unit

60 **MIMIC-IV:** Medical Information Mart for Intensive Care-IV

61 **OR:** Odds ratio

62 **ROC:** Receiver operating characteristic

63 **Introduction**

64 Atrial fibrillation (AF) is a prevalent cardiac arrhythmia linked to considerable
65 morbidity and mortality. It is characterized by an irregular and often rapid heart rate,
66 resulting in compromised blood flow and potential complications such as stroke, heart
67 failure, and other cardiovascular events [1]. AF has a broad impact on cardiac
68 function, functional status, and quality of life and is also a risk factor for stroke [2].
69 Additionally, AF is a significant risk factor for stroke and becomes more prevalent
70 with age, affecting more than 2 million individuals in the United States, 14% to 17%
71 of whom are aged 65 years and older [3]. Individuals diagnosed with AF experienced
72 a 3.7 times greater likelihood of mortality from any cause than the general population
73 [4]. AF represents a significant public health issue due to its considerable impact on
74 morbidity and mortality as well as its economic strain on healthcare systems.

75 The assessment tool for evaluating the risk of stroke in patients with AF, known as
76 the congestive heart failure, hypertension, age, diabetes mellitus, prior stroke or TIA
77 or thromboembolism, vascular disease, age, sex category (CHA₂DS₂-VASc) score [5],
78 has been associated with cardiovascular events and mortality in diverse patient groups,
79 including those with diabetes [6] and individuals without AF [7]. Nevertheless, tools
80 specifically designed to assess mortality risk in patients with AF are lacking. Although
81 recent studies have introduced new AF risk scores [8, 9], these scores were developed
82 based on data from clinical trials, limiting their applicability to the broader AF
83 population.

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

88

89

90 **Methods**

91 **Study design and setting**

92 This was a single-center retrospective cohort study. The data utilized in this research
93 originated from the Medical Information Mart for Intensive Care-IV (MIMIC-IV
94 version 2.2) database [10, 11]. Over the period from 2008 to 2019, the intensive care
95 unit (ICU) at Beth Israel Deaconess Medical Center admitted more than 50,000
96 critically ill patients, as documented in MIMIC-IV. Approval for the MIMIC-IV
97 database was granted by the Massachusetts Institute of Technology (Cambridge, MA)
98 and Beth Israel Deaconess Medical Center (Boston, MA), with consent obtained for
99 the initial data collection.

100

101 **Study population**

102 The study population included patients aged 18 years and older with a discharge
103 diagnosis of AF. Patients with AF were identified by searching International
104 Classification of Diseases diagnostic terminology in the MIMIC-IV database by

105 matching the keyword "atrial fibrillation". The types of queried AF diagnostic terms
106 were manually reviewed to ensure compliance. The exclusion criterion was lack of
107 patient data on survival outcomes.

108

109 **Study variables**

110 The variables examined in the research included the characteristics of the study
111 population, complications, various scores (such as Charlson Comorbidity Index and
112 the CHA2DS2-VASc score), vital signs, and an array of laboratory tests (including
113 routine bloodwork, blood biochemistry, coagulation, blood lipids, cardiac markers,
114 etc.). Additionally, the investigation considered use of vasopressors (norepinephrine,
115 epinephrine, phenylephrine, dopamine, dobutamine, vasopressin, and milrinone),
116 antithrombotic agents (heparin, enoxaparin, warfarin, aspirin, clopidogrel, ticagrelor,
117 rivaroxaban, edoxaban, dabigatran etexilate, fondaparinux sodium, prasugrel, and
118 apixaban), and beta blockers (propranolol, metoprolol, bisoprolol, carvedilol, labetalol,
119 atenolol, and nebivolol) and various other data points. For laboratory test items,
120 summary statistics, including minimum and maximum values during hospitalization,
121 were utilized to derive variables. An indicator column for the respective drug was
122 generated based on whether the drug was used during hospitalization.

123

124 **Outcome variable**

125 The primary outcome measured was 1-year mortality. Survival time was calculated by
126 utilizing the date of death information available in the MIMIC-IV database restricted
127 to a 1-year timeframe.

128

129

130

131 **Machine learning model development and validation**

132 The dataset was randomly partitioned into training and test samples at a 3:1 ratio. To
133 prevent model overfitting, tenfold cross-validation and model calibration techniques
134 were applied. To accommodate varying degrees of missing values in dataset variables,
135 the mainstream machine learning model XGBoost was employed due to its ability to
136 handle missing data. The discriminative performance of the models was assessed
137 using the area under the receiver operating characteristic (ROC) curve. Feature
138 scaling was deemed unnecessary before inputting the data into the model. A total of
139 164 candidate variables were incorporated into the model training process.

140 Furthermore, a calibration curve was utilized as a graphical representation to evaluate
141 the concordance between the predicted probabilities and observed outcomes in binary
142 classification models. On the calibration curve, the x-axis denotes the mean predicted
143 probability assigned by the model to a specific class, and the y-axis signifies the
144 observed frequency of positive instances. Ideally, a well-calibrated model produces a

145 calibration curve that closely aligns with the diagonal line ($y=x$), signifying a perfect
146 correspondence between the predicted probabilities and actual outcomes.

147

148 **Development of the scoring scale**

149 The XGBoost model assigned importance to predictor variables, and variables with
150 higher importance were selected based on this ranking. These selected variables were
151 subsequently integrated into a logistic model to construct the scoring model. Manual
152 testing was employed to evaluate the impact of introducing or removing variables on
153 the area under the curve (AUC) of the logistic model in the test set. After striking a
154 balance between AUC performance and the increase in model complexity associated
155 with the number of variables included, the chosen variables for the AF scoring model
156 were ultimately determined. A nomogram was used to construct the finalized AF
157 scores. Decision curve analysis (DCA) was employed to assess the clinical utility and
158 net benefit of the AF scoring model, CCI, and CHA2DS2-VASc scores within the test
159 set [12, 13]. DCA quantifies the net benefit of a clinical prediction model at different
160 risk thresholds, avoiding the simplistic assumptions of all patients being at low or
161 high risk. The superior model is identified by the highest net benefit at the chosen
162 threshold. The flow chart of the study is shown in **Figure 1**.

163

164 **Data analysis**

165 Python software (version 3.11.5) was used to construct machine learning models,
166 evaluate performance, and generate AUCs and calibration curves. R software (version
167 4.3.2) was used for logistic and Cox regression analyses, forest plot creation, DCA,
168 and nomogram generation. Baseline characteristics are presented as the mean
169 (standard deviation), median (Q1, Q3), or percentage (%), as determined by the
170 distribution characteristics of the data. The DeLong test was applied to determine
171 whether the AUC of a given prediction significantly differed from that of another
172 prediction [14]. Python was used for descriptive tables [15] and DeLong tests. When
173 constructing the original machine learning model, no handling of missing values was
174 conducted. However, during the development of the logistic model for the AF score,
175 missing values were removed from the dataset based on the variables included in the
176 AF score, as logistic models are unable to manage missing values. In all analyses,
177 statistical significance was defined as a two-sided p value < 0.05.

178

179

180 **Results**

181 **Baseline characteristics**

182 This study enrolled 59,595 individuals diagnosed with AF from the MIMIC-IV
183 database. Among them, 55.6% were male. The cohort had a median age of 77.3 years
184 (with an interquartile range of 68.1-85.3), a median CHA2DS2-VASc score of 4
185 (interquartile range, 3-5), a median CCI of 6 (interquartile range, 4-8), and a median

186 hospitalization duration of 4 days (interquartile range, 1-7). Additional results are
187 presented in **Table 1**.

188

189 **Screening variables using the XGBoost model**

190 The XGBoost model showed an AUC and 95% confidence intervals (95% CI) of
191 0.833 (95% CI, 0.826-0.839) when predicting 1-year mortality in the test set (**Figure**
192 **2**). **Figure 3** illustrates the significance of the predictor variables determined by the
193 XGBoost model. Notably, CCI and the presence of metastatic solid tumors were
194 identified as the top two variables, with considerably greater importance than the
195 other variables.

196

197 **Derivation and evaluation of the AF score**

198 The 1-year mortality risk score for AF was calculated as CRAMB, which represents
199 CCI, readmission, age, metastatic solid tumor, and maximum blood urea nitrogen
200 (BUN) (variable "bun_max") (**Figure 3**). The variable "readmission" was derived
201 from the variable "hospstay_seq" (hospital stay sequence, indicating the number of
202 hospitalizations) for convenient clinical application. Logistic and Cox regression
203 analyses were employed to assess the predictive value of these five variables for the
204 outcome of death, as expressed as odds ratios (ORs). Both the forest plot (**Figure 4**)
205 and the forest plot of hazard ratios (HRs) (**Figure 5**) revealed that these variables
206 were significantly different.

207 A nomogram was used to calculate the CRAMB score (**Figure 6**). In the test set,
208 the area under the curve (AUC) for the CRAMB score was 0.756 (95%
209 CI=0.748-0.764), surpassing CCI, at 0.720 (95% CI=0.712-0.728), and the
210 CHA2DS2-VASc score, at 0.609 (95% CI=0.600-0.618). **Table 2** displays
211 supplementary performance metrics corresponding to these scores. The DeLong test
212 results comparing the CRAMB score with existing scores (CCI and CHA2DS2-VASc)
213 revealed statistically significant differences ($p < 0.001$), as indicated in **Table 2**. The
214 DCA results provided in **Figure 7** demonstrate that the CRAMB score consistently
215 exhibited a positive and greater net benefit across the entire threshold range than the
216 default strategies, assuming either high or low risk, as indicated by CCI and
217 CHA2DS2-VASc scores, and the hypothesis of not using a scoring system. The
218 calibration plot (**Figure 8**) for the test set indicated that the CRAMB score was well
219 calibrated.

220

221

222 **Discussion**

223 **Main findings**

224 This study's primary contribution is the establishment of a benchmark for utilizing
225 machine learning models in the construction of AF scores for mortality prediction.
226 This study introduces and validates a novel risk score for assessing the 1-year
227 mortality risk in patients with AF. By leveraging a large-sample population dataset

228 and employing XGBoost models for predictor screening, the CRAMB (Charlson
229 comorbidity index, readmission, age, metastatic solid tumor, and blood urea nitrogen
230 maximum) score was developed. XGBoost excels in variable selection by effectively
231 capturing nonlinear relationships and handling missing data [16]. Its built-in feature
232 importance mechanism automatically identifies key variables, a capability lacking in
233 logistic regression. Furthermore, compared with logistic regression, XGBoost's
234 ensemble learning often results in superior predictive performance, and its
235 regularization techniques boost resilience against overfitting, making it a robust
236 choice for predictive modeling and variable selection. The variables incorporated in
237 the CRAMB score were validated through logistic and Cox regression analyses,
238 demonstrating their predictive significance for mortality. The CRAMB score
239 exhibited excellent calibration, and DCA illustrated its clinical utility. Importantly, the
240 findings of this study revealed that the CRAMB score outperformed the widely used
241 CHA2DS2-VASc risk score in predicting mortality, despite the latter's original focus
242 on predicting ischemic stroke.

243

244 **Predictors of death in patients with AF**

245 Predictors and risk factors for death in patients with AF span a broad spectrum of
246 clinical and demographic variables. Hypertension has been identified as a significant
247 risk factor for incident heart failure and all-cause mortality in AF patients [17].
248 Moreover, an independent association has been observed between red cell distribution

249 width and the risk of all-cause mortality in AF patients, with elevated red cell
250 distribution width linked to increased mortality risk [18].

251 Left ventricular hypertrophy has been confirmed to be an independent risk factor
252 for stroke and death in AF patients [19]. Moreover, patients with chronic kidney
253 disease who develop AF face an increased risk of stroke and death [20], and renal
254 function has been associated with the risk of stroke and bleeding in AF patients [21].
255 Additionally, age correlates with elevated risks of stroke and mortality in patients with
256 either AF or sinus rhythm [22]. Those undergoing hemodialysis with new-onset AF
257 exhibit heightened risks of death and stroke [23], and chronic kidney disease and
258 hemodialysis impact mortality, length of stay, and total cost of hospitalization in
259 patients admitted with a primary diagnosis of paroxysmal AF [24]. Furthermore, the
260 biatrial volume predicts AF recurrence after ablation, impacting mortality [25].

261 Although a growing body of evidence indicates a connection between cancer and
262 AF [26-28], the exact extent and underlying mechanisms of this association remain
263 unclear. Proposed factors such as cancer-related inflammation, anticancer treatments,
264 and other comorbidities associated with cancer are believed to influence atrial
265 remodeling, potentially heightening the susceptibility of cancer patients to AF [26].
266 Specifically, new-onset AF has been associated with higher mortality in cancer
267 patients [29, 30], and patients with solid tumors face an elevated risk of developing
268 AF compared to those without cancer. The highest risk of AF occurs within 90 days of

269 cancer diagnosis (OR 7.62, 95% CI 3.08 to 18.88), and this risk gradually decreases
270 over time [30].

271 There is considerable evidence affirming the significance of BUN as a mortality
272 predictor in patients with CVD [31-33]. Moreover, research indicates that BUN levels
273 at admission serve as a predictive marker for mortality in patients with heart failure
274 [32]. Elevated BUN levels (>13.51 mg/dL) in healthy older women have also been
275 linked to a greater incidence of heart failure [34]. The present study contributes novel
276 insights into the role of BUN in predicting mortality among AF patients through the
277 use of the CRAMB score.

278 In summary, a review of the literature on predictor variables associated with AF
279 death revealed that the variables used in constructing the CRAMB score are
280 reasonable.

281

282

283 **Comparison with similar studies**

284 Hijazi et al. proposed an innovative biomarker-based tool designed to predict
285 mortality in patients with AF [8]. This score, named ABC-death (age, biomarkers
286 (N-terminal pro B-type natriuretic peptide, troponin T, growth differentiation
287 factor-15), and clinical history of heart failure), was developed and internally
288 validated in 14,611 AF patients randomized to either apixaban or warfarin over a
289 median period of 1.9 years. External validation was conducted in a cohort of 8,548 AF

290 patients randomized to dabigatran or warfarin for 2.0 years. The study utilized a
291 two-step variable selection method, initially fitting a model with all candidate
292 predictors and then approximating a blinded, smaller model with the most predictive
293 variables through ordinary least squares (OLS) and backward elimination. However,
294 this approach has limitations, potentially excluding relevant variables and assuming
295 linearity. Due to the exclusion of more severely ill patients in these two clinical trials,
296 the scores developed from these datasets may not be applicable to other populations
297 of severely ill patients with a greater number of comorbidities.

298 In a separate study, Samaras et al. introduced the BASIC-AF risk score (biomarkers,
299 age, ultrasound, ventricular conduction delay, and clinical history), a novel prognostic
300 tool for predicting mortality in AF patients [9]. This score was developed and
301 validated using data from 1,130 patients in a single-center clinical trial; notably, the
302 dataset was not split into training and test sets due to sample size restrictions. The
303 study [9] incorporated traditional cardiac biomarkers and introduced
304 echocardiography variables, identifying indexed left atrial volume as a significant
305 predictor. This study [9] employed a random forest model for variable selection, and
306 the random forest model has challenges with missing values, as it constructs decision
307 trees based on specific features, posing difficulties in managing missing data.
308 Imputing or estimating missing values introduces potential biases, and within the
309 ensemble structure of random forests, this problem is compounded, compromising the
310 model's reliability in the case of incomplete data. XGBoost outperforms random

311 forest models in variable selection because of its advanced regularization technique,
312 gradient boosting framework for error correction, and more sophisticated
313 tree-growing strategy. This approach results in improved predictive accuracy and
314 precise estimation of variable importance.

315 Yan et al. focused on the development and validation of a prediction model for
316 in-hospital death in patients with heart failure and AF [33]. Although that study did
317 not directly align with the task of developing a score for risk of death in AF, it is
318 relevant because it addresses the intersection of heart failure and AF, which are often
319 comorbid conditions. The prediction model developed by Yan et al. may offer insights
320 into prognostic factors and risk assessment strategies that may be applicable to the
321 development of AF risk scores. The present study [33] developed a scoring model
322 using 5998 patients admitted to the ICU from the MIMIC-IV database and used
323 logistic regression to screen approximately 38 prespecified variables, ultimately
324 including 16 variables to construct the score. However, the complexity of the scoring
325 system prevented an evaluation of the importance of variables in variable screening
326 based on traditional logistic regression.

327 Compared to the ABC-death risk score and BASIC-AF risk score, the CRAMB score
328 was constructed based on the MIMIC-IV database, leading to significant differences
329 in population characteristics compared to clinical trial populations. Directly
330 comparing the AUC performance of the three scores is not appropriate. Although this
331 study identified important predictor variables, such as the cardiac biomarker troponin

332 T and comorbidities such as diabetes, their addition to the CRAMB score yielded only
333 a marginal AUC benefit of less than 0.01 in absolute terms. Conversely, replacing or
334 eliminating existing variables in the CRAMB score resulted in varying degrees of
335 AUC performance reduction. This discrepancy may be attributed to the wide
336 heterogeneity of the AF population in the MIMIC-IV database. Therefore, this study
337 addresses a gap in the development of scoring methods and screening predictor
338 variables within a broader population than did previous studies of this nature. Future
339 research on AF scores should focus on the characteristics of the population used for
340 score development, comprehensively considering the importance and applicability of
341 the variables included.

342

343 **Limitations**

344 The primary limitation of this study is its reliance on single-center data, preventing
345 external validation of the developed AF score. Importantly, the predictor variables
346 used for developing the AF risk score in this study may not be comprehensive enough.
347 Additionally, the MIMIC-IV dataset includes derived databases such as
348 MIMIC-IV-ECHO [35], which contains echocardiogram data; MIMIC-IV-ECG,
349 which contains electrocardiogram data [36]; and MIMIC-IV-Note, which contains
350 deidentified free-text clinical notes [37]. These derived datasets offer more patient
351 characteristic data, enhancing the screening of variables.

352

353 **Conclusions**

354 This study's primary contribution is the establishment of a benchmark for utilizing
355 machine learning models in the construction of a score for mortality prediction in AF
356 patients. By leveraging a large-sample population dataset and employing XGBoost
357 models for predictor screening, the CRAMB (CCI, readmission, age, metastatic solid
358 tumor, and blood urea nitrogen maximum) score was developed. The CRAMB score's
359 simplicity makes it user-friendly, allowing for coverage of a broad and heterogeneous
360 AF population. Moreover, the proposed model has superior predictive performance
361 compared to that of the clinically utilized CHA2DS2-VASc risk score for 1-year
362 mortality among AF patients. External validation of the CRAMB score in new
363 datasets holds potential value for enhancing clinical practice.

364

365 **Acknowledgments**

366 Not applicable.

367

368 **Sources of Funding**

369 This work was supported by the Real World Study Project of Hainan Boao Lecheng
370 Pilot Zone (Real World Study Base of NMPA) (No. HNLC2022RWS017).

371

372 **Disclosures**

373 None.

374 **References**

- 375 1. Schnabel RB, Yin X, Gona P, Larson MG, Beiser AS, McManus DD, et al. 50
376 year trends in atrial fibrillation prevalence, incidence, risk factors, and
377 mortality in the Framingham Heart Study: a cohort study. *Lancet*.
378 2015;386:154-62. [https://doi.org/10.1016/S0140-6736\(14\)61774-8](https://doi.org/10.1016/S0140-6736(14)61774-8)
- 379 2. Westerman S, Wenger NK. Gender Differences in Atrial Fibrillation: A
380 Review of Epidemiology, Management, and Outcomes. *Current Cardiology*
381 *Reviews*. 2019;<https://doi.org/10.2174/1573403x15666181205110624>
- 382 3. Hart RG, Benavente O, Jacobs LG. The Sin of Omission: A Systematic
383 Review of Antithrombotic Therapy to Prevent Stroke in Atrial Fibrillation.
384 *Journal of the American Geriatrics Society*.
385 2001;<https://doi.org/10.1046/j.1532-5415.2001.49016.x>
- 386 4. Lee E, Choi EK, Han KD, Lee H, Choe WS, Lee SR, et al. Mortality and
387 causes of death in patients with atrial fibrillation: A nationwide
388 population-based study. *PLoS One*. 2018;13:e0209687.
389 <https://doi.org/10.1371/journal.pone.0209687>
- 390 5. January CT, Wann LS, Calkins H, Chen LY, Cigarroa JE, Cleveland JC, Jr., et
391 al. 2019 AHA/ACC/HRS Focused Update of the 2014 AHA/ACC/HRS
392 Guideline for the Management of Patients With Atrial Fibrillation: A Report
393 of the American College of Cardiology/American Heart Association Task
394 Force on Clinical Practice Guidelines and the Heart Rhythm Society in

- 395 Collaboration With the Society of Thoracic Surgeons. *Circulation*.
396 2019;140:e125-e51. <https://doi.org/10.1161/CIR.0000000000000665>
- 397 6. Tourountzis T, Lioulios G, Ginikopoulou E, Stasini F, Skarlatou Z, Stai S, et
398 al. The prognostic value of CHA(2) DS(2) -VASc and modified-CHADS(2)
399 scores for cardiovascular events in diabetics and non-diabetics haemodialysis
400 patients. *Nephrology* (Carlton). 2023;28:387-98.
401 <https://doi.org/10.1111/nep.14165>
- 402 7. Renda G, Ricci F, Patti G, Aung N, Petersen SE, Gallina S, et al.
403 CHA(2)DS(2)VASc score and adverse outcomes in middle-aged individuals
404 without atrial fibrillation. *Eur J Prev Cardiol*. 2019;26:1987-97.
405 <https://doi.org/10.1177/2047487319868320>
- 406 8. Hijazi Z, Oldgren J, Lindback J, Alexander JH, Connolly SJ, Eikelboom JW,
407 et al. A biomarker-based risk score to predict death in patients with atrial
408 fibrillation: the ABC (age, biomarkers, clinical history) death risk score. *Eur*
409 *Heart J*. 2018;39:477-85. <https://doi.org/10.1093/eurheartj/ehx584>
- 410 9. Samaras A, Kartas A, Akrivos E, Fotos G, Dividis G, Vasdeki D, et al. A
411 novel prognostic tool to predict mortality in patients with atrial fibrillation:
412 The BASIC-AF risk score. *Hellenic J Cardiol*. 2021;62:339-48.
413 <https://doi.org/10.1016/j.hjc.2021.01.007>

- 414 10. Johnson AEW, Bulgarelli L, Shen L, Gayles A, Shammout A, Horng S, et al.
415 MIMIC-IV, a freely accessible electronic health record dataset. *Sci Data*.
416 2023;10:1. <https://doi.org/10.1038/s41597-022-01899-x>
- 417 11. Johnson A, Bulgarelli L, Pollard T, Horng S, Celi LA, Mark R. MIMIC-IV
418 (version 2.2). <https://doi.org/10.13026/6mm1-ek67>. Accessed August 26 2023.
- 419 12. Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to
420 interpreting decision curve analysis. *Diagn Progn Res*. 2019;3:18.
421 <https://doi.org/10.1186/s41512-019-0064-7>
- 422 13. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating
423 prediction models. *Med Decis Making*. 2006;26:565-74.
424 <https://doi.org/10.1177/0272989X06295361>
- 425 14. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under
426 two or more correlated receiver operating characteristic curves: a
427 nonparametric approach. *Biometrics*. 1988;44:837-45.
- 428 15. Pollard TJ, Johnson AEW, Raffa JD, Mark RG. tableone: An open source
429 Python package for producing summary statistics for research papers. *JAMIA*
430 *Open*. 2018;1:26-31. <https://doi.org/10.1093/jamiaopen/ooy012>
- 431 16. Chen T, Guestrin C: XGBoost. In: *Proceedings of the 22nd ACM SIGKDD*
432 *International Conference on Knowledge Discovery and Data Mining*.
433 Association for Computing Machinery; 2016: 785-94.

- 434 17. Middeldorp ME, Ariyaratnam JP, Kamsani SH, Albert CM, Sanders P.
435 Hypertension and atrial fibrillation. *J Hypertens.* 2022;40:2337-52.
436 <https://doi.org/10.1097/HJH.0000000000003278>
- 437 18. Saliba W, Barnett-Griness O, Rennert G. Red cell distribution width and
438 all-cause mortality in patients with atrial fibrillation: A cohort study. *J*
439 *Arrhythm.* 2017;33:56-62. <https://doi.org/10.1016/j.joa.2016.06.001>
- 440 19. Yao HM, Wang XL, Peng X, Chen SY, Wan X, Zuo W, Gan X. Increased red
441 blood cell distribution width might predict left ventricular hypertrophy in
442 patients with atrial fibrillation. *Medicine (Baltimore).* 2020;99:e22119.
443 <https://doi.org/10.1097/MD.00000000000022119>
- 444 20. Carrero JJ, Trevisan M, Sood MM, Bárány P, Xu H, Evans M, et al. Incident
445 Atrial Fibrillation and the Risk of Stroke in Adults With Chronic Kidney
446 Disease. *Clinical Journal of the American Society of Nephrology.*
447 2018;<https://doi.org/10.2215/cjn.04060318>
- 448 21. Bonde AN, Lip GY, Kamper AL, Fosbol EL, Staerk L, Carlson N, et al. Renal
449 Function and the Risk of Stroke and Bleeding in Patients With Atrial
450 Fibrillation: An Observational Cohort Study. *Stroke.* 2016;47:2707-13.
451 <https://doi.org/10.1161/STROKEAHA.116.014422>
- 452 22. Xing Y, Sun Y, Li H, Tang M, Huang W, Zhang K, et al.
453 CHA(2)DS(2)-VASc score as a predictor of long-term cardiac outcomes in

- 454 elderly patients with or without atrial fibrillation. *Clin Interv Aging*.
455 2018;13:497-504. <https://doi.org/10.2147/CIA.S147916>
- 456 23. Shih CJ, Ou SM, Chao PW, Kuo SC, Lee YJ, Yang CY, et al. Risks of Death
457 and Stroke in Patients Undergoing Hemodialysis With New-Onset Atrial
458 Fibrillation: A Competing-Risk Analysis of a Nationwide Cohort. *Circulation*.
459 2016;133:265-72. <https://doi.org/10.1161/CIRCULATIONAHA.115.018294>
- 460 24. Gulati A. Impact of Chronic Kidney Disease on Outcomes in Patients
461 Admitted to the Hospital for Paroxysmal Atrial Fibrillation: A National
462 Inpatient Sample Database Analysis. *Journal of Cardiology Research Review*
463 & Reports. 2021;[https://doi.org/10.47363/jcrrr/2021\(2\)153](https://doi.org/10.47363/jcrrr/2021(2)153)
- 464 25. Kumagai Y, Iwayama T, Arimoto T, Kutsuzawa D, Hashimoto N, Tamura H,
465 et al. Biatrial volume, estimated using magnetic resonance imaging, predicts
466 atrial fibrillation recurrence after ablation. *Pacing Clin Electrophysiol*.
467 2018;41:1635-42. <https://doi.org/10.1111/pace.13521>
- 468 26. Chu G, Versteeg HH, Verschoor AJ, Trines SA, Hemels MEW, Ay C, et al.
469 Atrial fibrillation and cancer - An unexplored field in cardiovascular oncology.
470 *Blood Rev*. 2019;35:59-67. <https://doi.org/10.1016/j.blre.2019.03.005>
- 471 27. Lateef N, Kapoor V, Ahsan MJ, Latif A, Ahmed U, Mirza M, et al. Atrial
472 fibrillation and cancer; understanding the mysterious relationship through a
473 systematic review. *J Community Hosp Intern Med Perspect*. 2020;10:127-32.
474 <https://doi.org/10.1080/20009666.2020.1726571>

- 475 28. Herrmann J. Adverse cardiac effects of cancer therapies: cardiotoxicity and
476 arrhythmia. *Nat Rev Cardiol.* 2020;17:474-502.
477 <https://doi.org/10.1038/s41569-020-0348-1>
- 478 29. Murtaza M, Baig MMA, Ahmed J, Serbanoiu LI, Busnatu SS. Higher
479 Mortality Associated With New-Onset Atrial Fibrillation in Cancer Patients: A
480 Systematic Review and Meta-Analysis. *Front Cardiovasc Med.* 2022;9:867002.
481 <https://doi.org/10.3389/fcvm.2022.867002>
- 482 30. Yuan M, Zhang Z, Tse G, Feng X, Korantzopoulos P, Letsas KP, et al.
483 Association of Cancer and the Risk of Developing Atrial Fibrillation: A
484 Systematic Review and Meta-Analysis. *Cardiol Res Pract.*
485 2019;2019:8985273. <https://doi.org/10.1155/2019/8985273>
- 486 31. Chiu CC, Wu CM, Chien TN, Kao LJ, Li C, Jiang HL. Applying an Improved
487 Stacking Ensemble Model to Predict the Mortality of ICU Patients with Heart
488 Failure. *J Clin Med.* 2022;11:<https://doi.org/10.3390/jcm11216460>
- 489 32. Khoury J, Bahouth F, Stabholz Y, Elias A, Mashiach T, Aronson D, Azzam
490 ZS. Blood urea nitrogen variation upon admission and at discharge in patients
491 with heart failure. *ESC Heart Fail.* 2019;6:809-16.
492 <https://doi.org/10.1002/ehf2.12471>
- 493 33. Yan M, Liu H, Xu Q, Yu S, Tang K, Xie Y. Development and validation of a
494 prediction model for in-hospital death in patients with heart failure and atrial

495 fibrillation. BMC Cardiovasc Disord. 2023;23:505.

496 <https://doi.org/10.1186/s12872-023-03521-3>

497 34. Lan Q, Zheng L, Zhou X, Wu H, Buys N, Liu Z, et al. The Value of Blood

498 Urea Nitrogen in the Prediction of Risks of Cardiovascular Disease in an

499 Older Population. Front Cardiovasc Med. 2021;8:614117.

500 <https://doi.org/10.3389/fcvm.2021.614117>

501 35. Gow B, Pollard T, Greenbaum N, Moody B, Johnson A, Herbst E, et al.

502 MIMIC-IV-ECHO: Echocardiogram Matched Subset (version 0.1).

503 <https://physionet.org/content/mimic-iv-echo/0.1/>. Accessed August 26 2023.

504 36. Gow B, Pollard T, Nathanson LA, Johnson A, Moody B, Fernandes C, et al.

505 MIMIC-IV-ECG: Diagnostic Electrocardiogram Matched Subset (version 1.0).

506 <https://physionet.org/content/mimic-iv-ecg/1.0/>. Accessed August 26 2023.

507 37. Johnson A, Pollard T, Horng S, Celi LA, Mark R. MIMIC-IV-Note:

508 Deidentified free-text clinical notes (version 2.2).

509 <https://physionet.org/content/mimic-iv-note/2.2/>. Accessed August 26 2023.

510

511

512 **Figure Legends**

513

514 **Figure 1** Flow chart of the study.

515 CCI: Charlson comorbidity index; AF: atrial fibrillation

516

517 **Figure 2** ROC curves for the XGBoost model in the test set and the scoring model in

518 the test set.

519 CCI: Charlson Comorbidity Index; CHA2DS2-VASc score: congestive heart failure,

520 hypertension, age, diabetes mellitus, prior stroke or TIA or thromboembolism,

521 vascular disease, age, sex category; CRAMB score: Charlson comorbidity index,

522 readmission, age, metastatic solid tumor, and blood urea nitrogen maximum.

523

524 **Figure 3** Feature importance values of the XGBoost model in the training set.

525

526 **Figure 4** Forest plot of the logistic model (CRAMB score) for predicting 1-year

527 mortality in the training set.

528

529 **Figure 5** Forest plot illustrating the ability of the Cox regression model to predict

530 1-year mortality in the training set stratified by the CRAMB score.

531

532 **Figure 6** Nomogram of the logistic model (CRAMB score) for predicting 1-year

533 mortality in the training set.

534

535 **Figure 7** Decision curve analysis of various scores in the training set.

536

537

538 **Figure 8** Calibration plots for the XGBoost model in the test set and the scoring

539 model in the test set.

540

541

Table 1 Demographic characteristics of the patients in the baseline cohort.

Item	Category	Missing	Total (59595)	Survivors (33177)	Nonsurvivors (26418)
Age, median [Q1, Q3]	-	0	77.3 [68.1,85.3]	74.2 [65.3,82.5]	81.1 [72.3,87.9]
Sex, n (%)	Female	0	26439 (44.4)	14300 (43.1)	12139 (45.9)
	Male	0	33156 (55.6)	18877 (56.9)	14279 (54.1)
In-hospital death, n (%)	No	0	56554 (94.9)	33177 (100.0)	23377 (88.5)
	Yes	0	3041 (5.1)	-	3041 (11.5)
Readmission, n (%)	No	0	19728 (33.1)	12996 (39.2)	6732 (25.5)
	Yes	0	39867 (66.9)	20181 (60.8)	19686 (74.5)
Intensive care unit admission, n (%)	No	0	41416 (69.5)	23724 (71.5)	17692 (67.0)
	Yes	0	18179 (30.5)	9453 (28.5)	8726 (33.0)
Hospital length of stay, median [Q1, Q3]	-	0	4.0 [1.0,7.0]	3.0 [1.0,6.0]	4.0 [2.0,8.0]

Item	Category	Missing	Total (59595)	Survivors (33177)	Nonsurvivors (26418)
CHA2DS2-VASc, median [Q1, Q3]	-	0	4.0 [3.0,5.0]	3.0 [2.0,5.0]	4.0 [3.0,5.0]
Charlson comorbidity index, median [Q1, Q3]	-	0	6.0 [4.0,8.0]	5.0 [3.0,7.0]	7.0 [5.0,9.0]

Table 2 Predictive performance of scores in the test set and DeLong test comparison. 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	0.659	0.563	0.738	0.720	<0.001
CHA2DS2-VASc	0.582	0.397	0.733	0.609	<0.001
CRAMB	0.690	0.594	0.768	0.756	-

CHA2DS2-VASc score: congestive heart failure, hypertension, age, diabetes mellitus, prior stroke or TIA or thromboembolism, vascular disease, age, sex category

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

Figure 1 Flow chart of the study.

CCI: Charlson comorbidity index; AF: atrial fibrillation

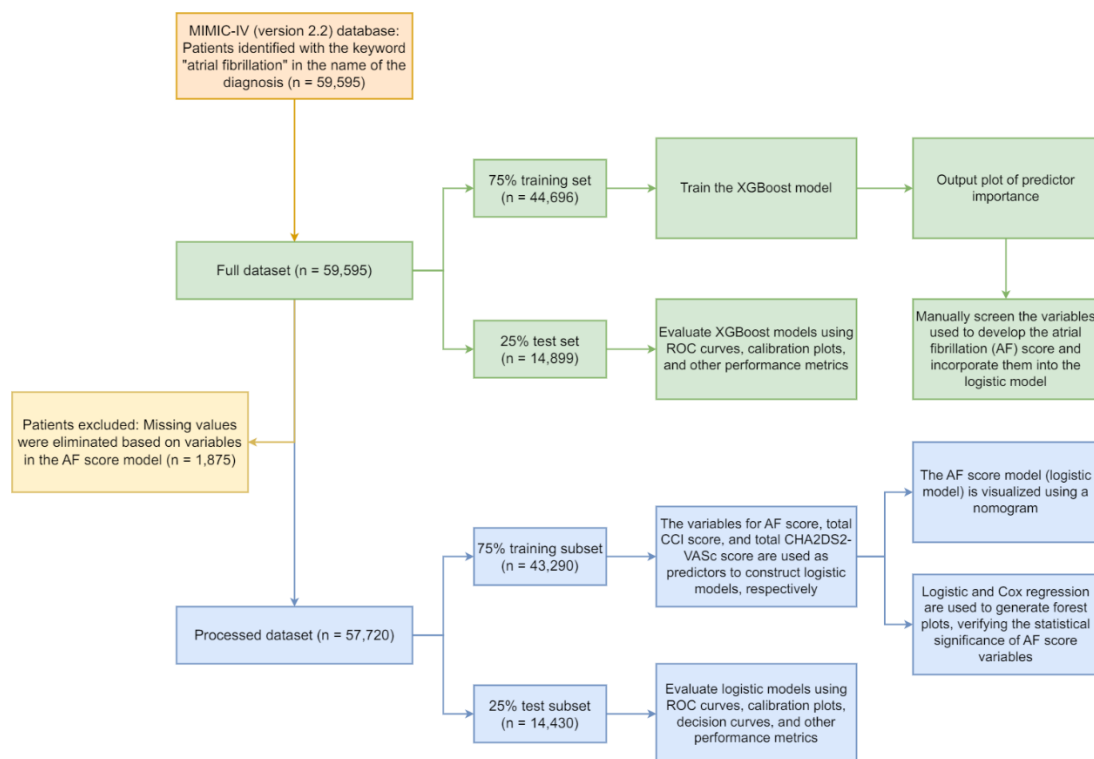


Figure 2 ROC curves for the XGBoost model in the test set and the scoring model in the test set.

CCI: Charlson Comorbidity Index; CHA2DS2-VASc score: congestive heart failure, hypertension, age, diabetes mellitus, prior stroke or TIA or thromboembolism, vascular disease, age, sex category; CRAMB score: Charlson comorbidity index, readmission, age, metastatic solid tumor, and blood urea nitrogen maximum.

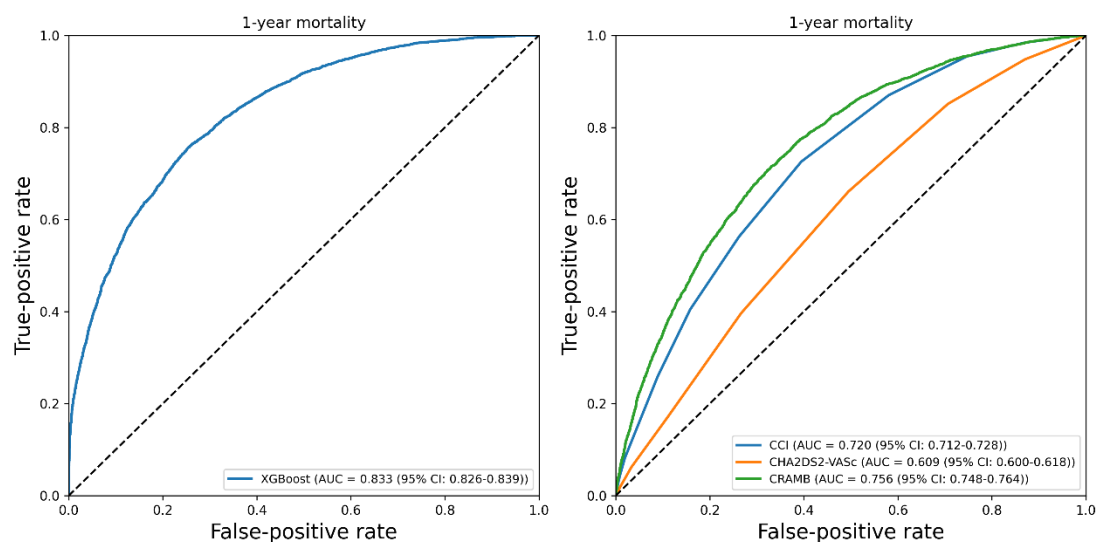


Figure 3 Feature importance values of the XGBoost model in the training set.

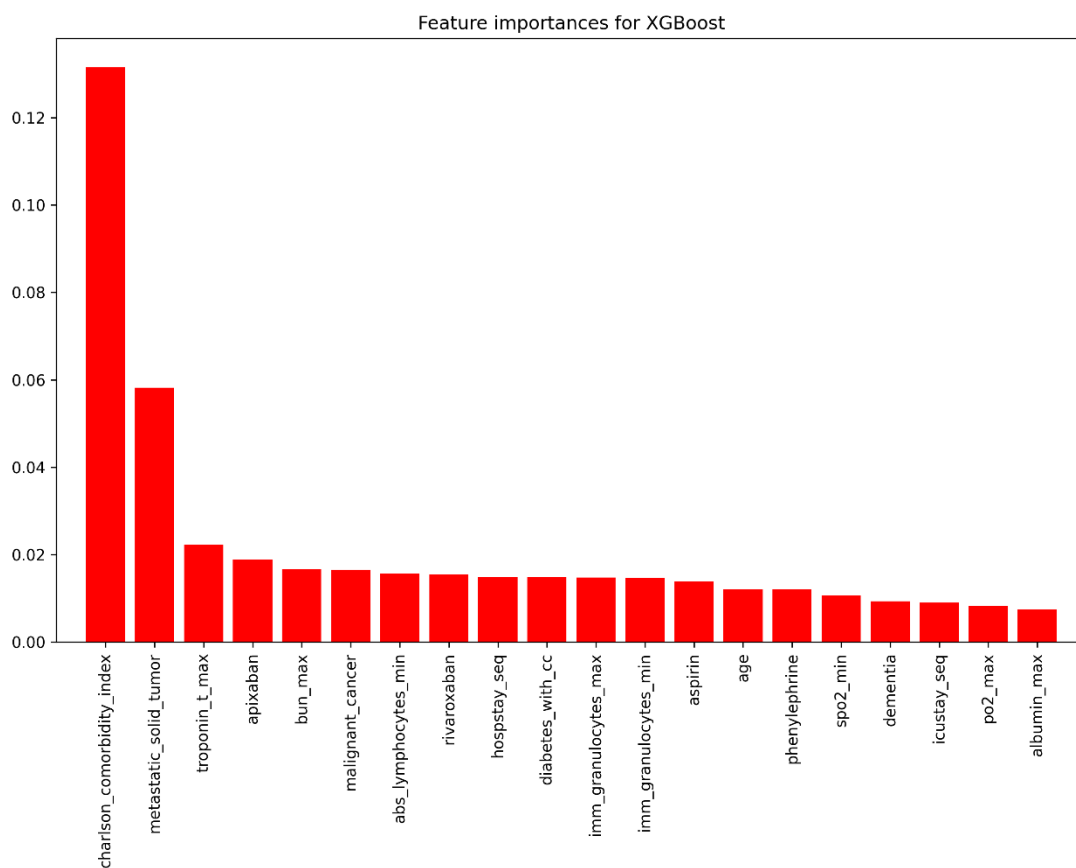


Figure 4 Forest plot of the logistic model (CRAMB score) for predicting 1-year mortality in the training set.

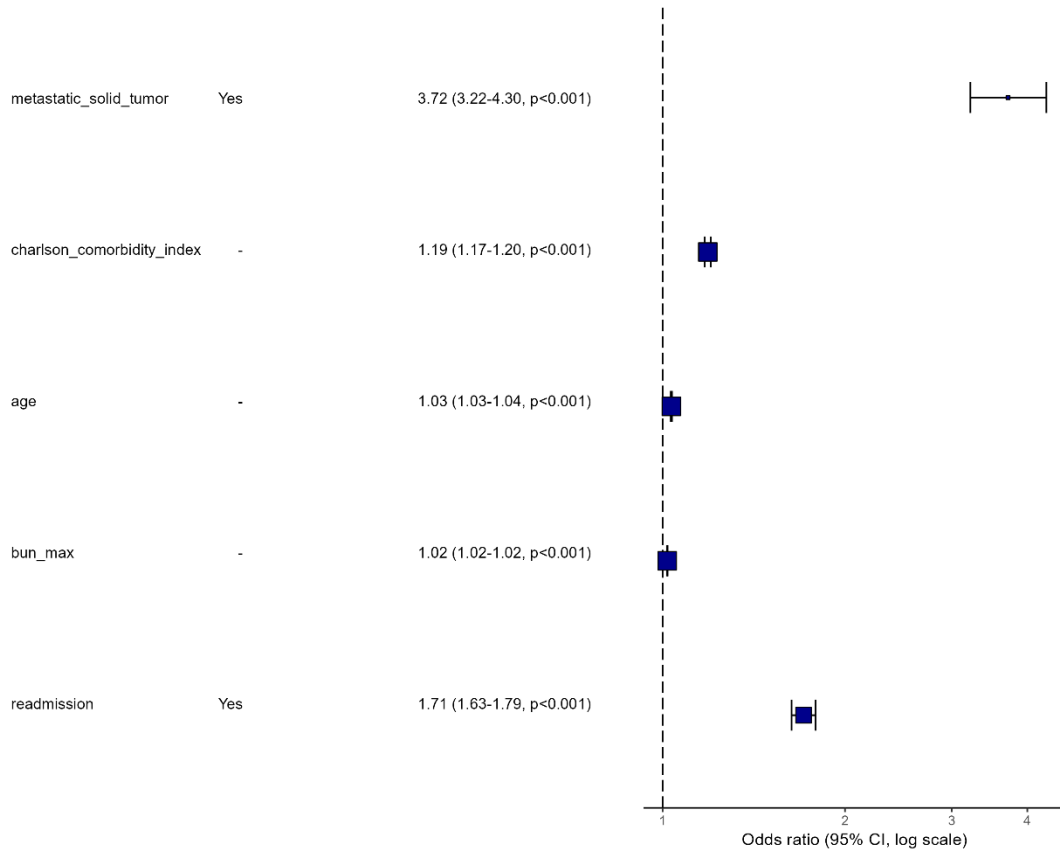


Figure 5 Forest plot illustrating the ability of the Cox regression model to predict 1-year mortality in the training set stratified by the CRAMB score.

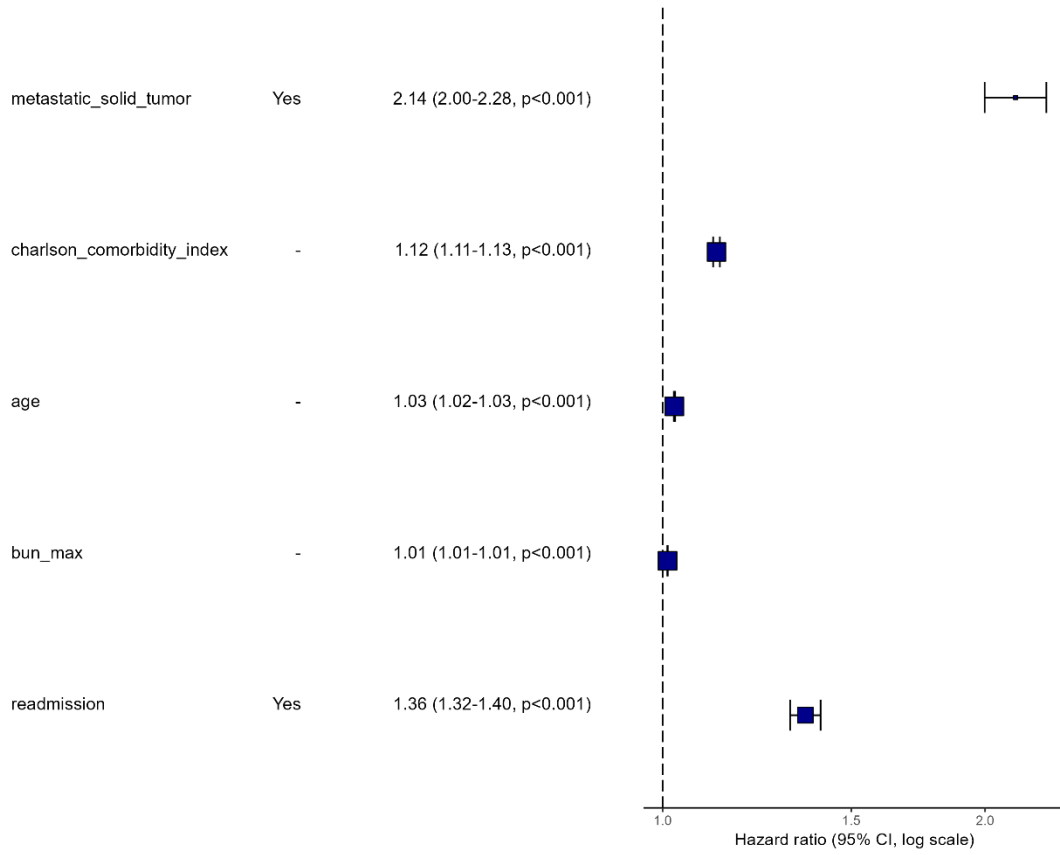
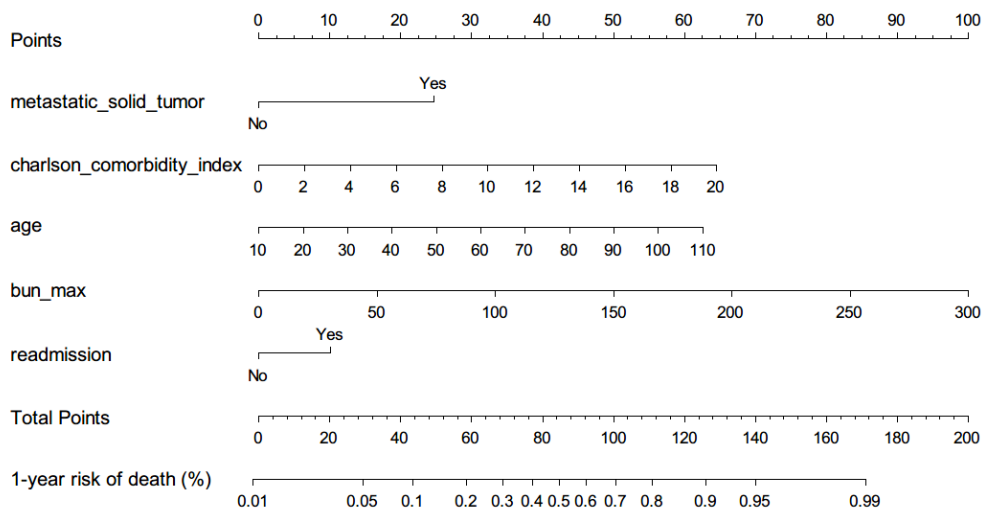


Figure 6 Nomogram of the logistic model (CRAMB score) for predicting 1-year mortality in the training set.



"bun_max": blood urea nitrogen (BUN) maximum

Figure 7 Decision curve analysis of various scores in the training set.

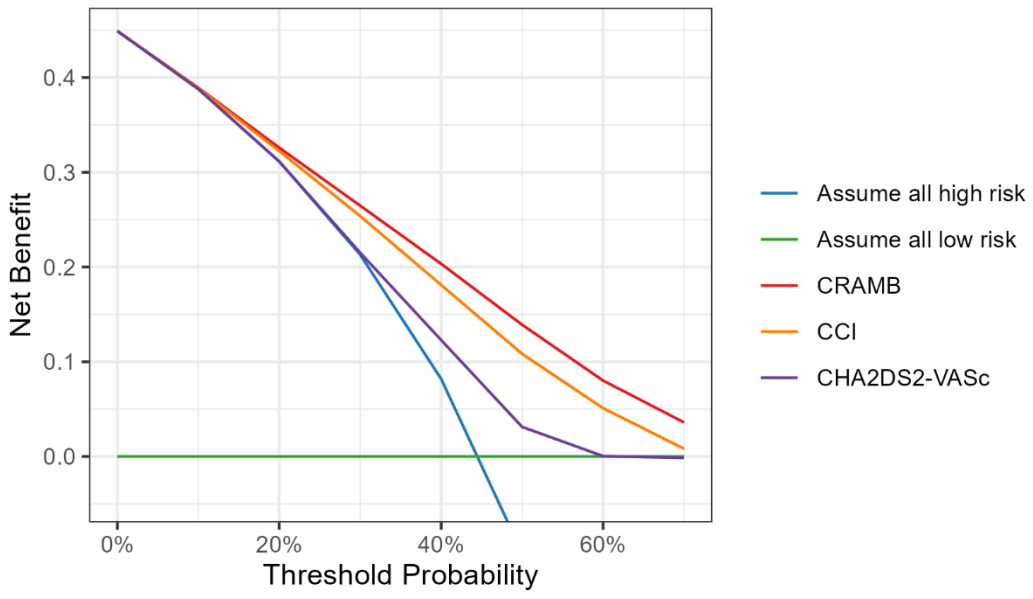


Figure 8 Calibration plots for the XGBoost model in the test set and the scoring model in the test set.

