

Machine learning–based prediction of in-hospital postoperative ischemic stroke following off-pump coronary artery bypass grafting

Zheng-Dong Kong

Plastic Surgery Hospital, Chinese Academy of Medical Sciences

Wen-Long Zou

Xuan Wu Hospital of the Capital Medical University

Guang-Pu Fan

Peking University People's Hospital

Wen-Qing Zhang

Beijing Anzhen Hospital, Capital Medical University

Chun Liu

Beijing Anzhen Hospital, Capital Medical University

Hui Xue

First Hospital of Tsinghua University

Chao-Bin Wang

Beijing Fangshan district Liangxiang Hospital

Lu-Feng Zhang

Beijing Tiantan Hospital, Capital Medical University

Dan Miao

Qingdao Hiser Hospital, Qingdao University

Xiao-Qing Li

Beijing Anzhen Hospital, Capital Medical University

Bin Cai

Plastic Surgery Hospital, Chinese Academy of Medical Sciences

Rui-Yuan Guan

School of Health Humanities, Peking University

Xiu-Bin Yang

Beijing Anzhen Hospital, Capital Medical University

Qing-Hua Cui

School of Basic Medical Sciences, Peking University, Beijing

Guang-Zhi Liu

`liugz1997@pumc.edu.cn`

Research Article

Keywords: Machine Learning, Ischemic Stroke, Off-Pump Coronary Artery Bypass Grafting, Predictive Model, Risk Factors

Posted Date: October 14th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10789974/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

This study aimed to develop and validate a machine learning (ML)-based model to assess the risk of ischemic stroke (IS) in hospitalised patients with coronary artery disease (CAD) after off-pump coronary artery bypass grafting (OPCABG).

Methods

Data were collected from the electronic health records of 25,488 hospitalised patients with CAD receiving coronary artery bypass grafting at the Beijing Anzhen Hospital of Capital Medical University (BAH-CMU cohort, training and internal validation sets) between January 1, 2019, and December 31, 2021, and at the Peking University People's Hospital (PKUPH cohort, external validation set) between May 1, 2016, and June 30, 2024. Patient and clinical characteristics available pre-, intra-, and postoperatively were used as inputs for five candidate ML models: support vector machine, random forest, logistic regression, k-nearest neighbour (KNN), and XGBoost. Model performance was compared using the area under the receiver operating characteristic curve (AUC). Shapley Additive exPlanation (SHAP) method was utilised to visualise and interpret the complex relationships between the variables and outcomes. Post-OPCABG IS Model (POISM), a web-based tool, was further designed utilising ML algorithms.

Results

A total of 11,855 patients (2452 [20.9%] female) were included in this study. The best-performing model, using perioperative data, achieved the following AUCs: internal validation (XGBoost), 0.863 (95% CI, 0.793–0.933); external validation (XGBoost), 0.786 (95% CI, 0.68–0.891); and combined validation (XGBoost), 0.827 (95% CI, 0.765–0.890). The SHAP analysis illustrated significant clinical contributors associated with risk of postoperative IS. Based on this model, we created an easy-to-use website tool, POISM, which clinicians can easily utilise for clinical decision-making in patients at risk of IS.

Conclusions

This XGBoost model showed strong predictive performance on both the training and test datasets. Moreover, the user-friendly web tool, POISM, can quickly predict the risk of postoperative IS in hospitalised patients with CAD undergoing OPCABG, thereby supporting clinical decision-making and reducing adverse outcomes.

Background

Coronary artery bypass grafting (CABG) is a major revascularisation approach for patients with coronary artery disease (CAD) and has favourable short- and long-term outcomes [1, 2]. Initially, CABG relied on on-pump techniques using saphenous vein grafts. However, off-pump coronary artery bypass grafting (OPCABG) emerged as an alternative to reduce complications associated with cardiopulmonary bypass (CPB), such as stroke, respiratory failure, and renal dysfunction [2, 3]. Among these, perioperative stroke, particularly ischemic stroke (IS), is a relatively frequent and life-threatening complication [1, 4], occurring in 1.1%–5.7% of CAD cases following CABG [5, 6]. Post-OPCABG stroke has a high negative impact on mortality and morbidity, increases the costs of both inpatient and follow-up care, and is a leading cause of rehospitalisation [7–9]. Therefore, predicting IS at admission is essential for multidisciplinary management, including medical and surgical treatments, to prevent potentially catastrophic outcomes.

To date, various risk models have been developed to screen patients with a high risk of postoperative mortality for optimal decision-making. Although current guidelines recommend them, several nonspecific scoring systems for stroke or mortality, such as the Society of Thoracic Surgeons model [10] and the European System for Cardiac Operative Risk Evaluation [11], are often ineffective in patients with CAD undergoing CABG, mainly due to their suboptimal accuracy and specificity for IS prediction [10, 12]. Moreover, most traditional models are based on logistic regression assuming independence among risk factors, whereas IS pathogenesis involves multiple interacting risk factors [12]. Hence, there is an urgent need for baseline risk stratification to develop specific prediction models for postoperative IS to improve outcomes through targeted management in high-risk patients.

Machine learning (ML) has been widely used in medicine as a powerful tool to uncover meaningful information hidden within massive amounts of complex data [13–15]. Unlike linear models that utilise logistic regression, ML algorithms use computational techniques to effectively explore parameter spaces, resulting in highly predictive models [16]. Several recent studies have attempted to develop ML-based models for predicting post-CABG complications (e.g., acute kidney injury and atrial fibrillation) [17, 18]. However, research focusing on assessing the IS risk after OPCABG remains limited. In this study, we applied ML to design an intuitive, user-friendly risk model for predicting postoperative IS in hospitalised patients with CAD undergoing OPCABG and assessed the accuracy and robustness of the model across various cohorts.

Methods

Patients and enrolment criteria

The study was approved by the Institutional Review Committee of Beijing Anzhen Hospital and Peking University People's Hospital and the requirement of informed consent was waived owing to its observational and non-invasive nature. All data were kept confidential and anonymous to ensure the protection of participants' privacy and rights.

Initially, a retrospective cohort study that included 22,558 hospitalised patients who underwent CABG consecutively was performed. These patients were admitted to Beijing Anzhen Hospital of Capital Medical University (Centre 1) between January 2018 and December 2022. Data were extracted from electronic medical records (EMRs). Data collected from patients undergoing CABG at Centre 1 were designated as the BAH-CMU cohort. Data were then randomly divided into training and internal validation sets, following an 8:2 ratio, and a random number table was used for separation. For external testing, data from 2,930 hospitalised patients undergoing CABG, who were consecutively admitted to Peking University People's Hospital (Centre 2), were used and referred to as the PKUPH cohort (**Supplementary Fig. 1**).

The following inclusion criteria were used: (1) aged ≥ 18 years and (2) undergoing OPCABG. The exclusion criteria were as follows: (1) on-pump CABG; (2) CABG combined with other surgical procedures; (3) aborted or incomplete surgery; and (4) missing data.

Definitions

All patients were adults with a diagnosis of CAD, encompassing both acute coronary syndrome (ACS) and chronic CAD. IS was defined as a sudden onset of global or focal neurological deficits lasting for > 24 h, with confirmation via computed tomography and/or magnetic resonance imaging [19]. Postoperative IS diagnosis was initially made by surgical staff, followed by independent confirmation by two neurologists. To minimise inter-rater variability in image interpretation, a standardised assessment protocol aligned with established clinical guidelines was implemented [20, 21], which included predefined radiological criteria to identify infarcted lesions and a structured reporting format.

Variable selection and model construction

Potential variables were identified from hospital admission data and various consensus statements, guidelines, and retrospective observational studies. This information was sourced from the full-admission EMRs and validated by two neurologists at each hospital. For patients with incomplete information, telephone follow-up was performed to obtain clinical CAD features. Available patient demographic data, medical history, echocardiographic findings, comorbidities, and laboratory tests were included [8, 9, 22, 23]. Laboratory test results encompassed the latest results prior to evaluation and those acquired at the time of diagnosis. Statisticians performed data cleaning and processing, and two researchers subsequently validated the extracted data to ensure reliability. Variables with $> 30\%$ missing values were excluded from the retrospective cohort analysis. The multiple imputation was used to impute missing values for variables with $< 30\%$ missing data. To ensure the model's robustness and reliability, Lasso regression was used to analyze and rank the importance of potential variables, leading to the selection of important variables for the final model.

Development, implementation, and evaluation of the models

This study used patient data from two different centres. The BAH-CMU cohort was randomly divided into training and internal test datasets. The PKUPH cohort was used as an external validation set to further evaluate the model capabilities. During the validation, hyperparameter tuning and cross-validation analyses were performed. This method provides a robust evaluation of the model before testing on different datasets and helped optimise model performance. The synthetic minority oversampling technique was employed to address data imbalance in the training and validation datasets, effectively mitigating the imbalance issue and improving model robustness.

Five established ML models, namely, support vector machine, random forest, logistic regression, k-nearest neighbours (KNN), and XGBoost, were developed to predict in-hospital postoperative IS. Several metrics were used to evaluate model performance: area under the receiver operating characteristic curve (AUC), area under the precision-recall curve (PR-AUC), calibration curve, and decision curve. The AUC and PR-AUC evaluated the overall ability of the model to differentiate between high- and low-risk individuals. The calibration curve illustrated how well the predicted outcomes matched the actual outcomes, and decision curve analysis (DCA) provided a visual representation of the net benefit for each model.

Model interpretation, development, and clinical application of a web-based tool

The Shapley Additive exPlanation (SHAP) method was employed to explore the underlying dynamics of the ML-based risk model. Each SHAP summary plot illustrated the contribution of each variable to the IS model predictions. The variables were ranked by importance, and the magnitude and direction of their effects were colour coded. Finally, POISM (Post-OPCABG IS Model), a web-based tool with a scoring function utilising ML algorithm, was designed and tested on two cases at risk of IS to evaluate its clinical utility and translational impact in perioperative care.

Statistical analysis

Continuous variables are expressed as mean $\pm$ SD or median (Q1, Q3). Student's *t*-test and the Mann–Whitney U test were conducted for normally and non-normally distributed data, respectively. Categorical variables are summarised as counts (percentages) and analysed using the chi-square test or Fisher's exact test. All statistical analyses were performed using R version 4.3.3 (R Foundation for Statistical Computing, Vienna, Austria). A *P*-value < 0.05 was deemed statistically significant.

Results

Patient characteristics

The BAH-CMU cohort comprised 10,184 adult patients who underwent OPCABG (**Supplementary Fig. 1**), and were divided into two groups: a training dataset with 8,147 patients and an internal validation dataset with 2,037 patients. In the training/internal validation dataset, the median age was 63 years

(interquartile range [IQR], 57–69 years), and 1,995 patients (19.6%) were female. The median hospitalisation duration was 13 days (IQR, 10–16 days). No differences in baseline characteristics of age, sex, or major comorbidities (e.g., diabetes and hypertension) were observed between the two groups (**Supplementary Table 1**). In this training cohort, 127 (1.58%) patients who underwent OPCABG developed IS. The basic characteristics of the patients with IS in the training dataset were detailed in Table 1. The PKUPH cohort included 1,671 patients (457 [27.3%] female), among whom 24 (1.44%) had IS (**Supplementary Table 2**).

Table 1

Baseline characteristics of patients with or without ischemic stroke (IS) in the BAH-CMU training cohort

Variables	IS patients (n = 127)	Non-IS patients (n = 8,020)	P-value
Age	65 (59, 69)	63 (57, 69)	0.003
Sex (female)	54 (42.5%)	1941 (24.2%)	< 0.001
T2DM	74 (58.3%)	3254 (40.6%)	< 0.001
Hypertension	98 (77.2%)	5230 (65.2%)	0.007
History of MI	17 (13.4%)	1311 (16.3%)	0.438
AMI	4 (3.1%)	614 (7.7%)	0.101
History of stroke	61 (48.0%)	1364 (17.0%)	< 0.001
Smoking	54 (42.5%)	3856 (48.1%)	0.248
Alcohol use	34 (26.8%)	2203 (27.5%)	0.941
History of cardiac surgery	16 (12.6%)	1295 (16.1%)	0.338
Continued Table 1. Baseline characteristics of patients in the BAH-CMU training cohort			
Carotid stenosis ($\geq 70\%$)	26 (20.5%)	451 (5.6%)	< 0.001
Emergency surgery	5 (3.9%)	257 (3.2%)	0.833
IABP	2 (1.6%)	134 (1.7%)	1.000
Preoperative SBP (mmHg)	165.00 (150.00, 182.00)	156.00 (140.00, 175.00)	0.001
Preoperative DBP (mmHg)	77.00 (70.00, 82.00)	75.00 (70.00, 82.00)	0.458
Postoperative SBP (mmHg)	126.00 (120.00, 133.00)	125.00 (120.00, 131.00)	0.044
Postoperative DBP (mmHg)	69.00 (65.00, 73.00)	69.00 (64.00, 75.00)	0.699

T2DM, type 2 diabetes mellitus; MI, myocardial infarction; AMI, acute myocardial infarction; IABP, intra-aortic balloon pump; LVEF, left ventricular ejection fraction; RDW-CV, red blood cell distribution width-coefficient of variation; MAP, mean arterial pressure; SBP, systolic blood pressure; DBP, diastolic blood pressure; POAF, postoperative atrial fibrillation; BMI, body mass index; Hs-CRP, high-sensitivity C-reactive protein; ALT, alanine aminotransferase; AST, aspartate aminotransferase; TG, triglycerides; TC, total cholesterol; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; BUN, blood urea nitrogen; UA, uric acid; FBG, fasting blood glucose; HCY, homocysteine; GA, glycated albumin; CK, creatine kinase; CK-MB, creatine kinase isoenzyme MB; LDH, lactate dehydrogenase; TP, total protein; AGR, albumin-to-globulin ratio; BNP, B-type natriuretic peptide; PT, prothrombin time; APTT, activated partial thromboplastin time; INR, international normalised ratio; FDP, fibrin degradation products; PF, plasma fibrinogen.

Variables	IS patients (n = 127)	Non-IS patients (n = 8,020)	P-value
Duration of a low intraoperative MAP (< 65 mmHg) (*5min)	1.00 (0.00, 3.00)	0.00 (0.00, 1.00)	< 0.001
Duration of surgery (h)	5.00 (4.50, 5.50)	5.00 (4.00, 5.50)	0.416
Washed erythrocyte transfusion (ml)	250.00 (0.00, 490.00)	250.00 (50.00, 420.00)	0.505
POAF	14 (11.0%)	311 (3.9%)	< 0.001
Intraoperative blood loss (ml)	700.00 (600.00, 900.00)	600.00 (600.00, 800.00)	0.376
Continued Table 1. Baseline characteristics of patients in the BAH-CMU training cohort			
LVEF (%)	60.00 (54.00, 65.00)	60.00 (55.00, 65.00)	0.429
BMI	25.71 (23.45, 27.29)	25.71 (23.74, 27.73)	0.938
Leukocyte (×10 ⁹ /L)	7.15 (5.66, 9.41)	7.51 (6.10, 9.65)	0.124
Erythrocyte (×10 ¹² /L)	4.22 (3.58, 4.59)	4.30 (3.78, 4.70)	0.042
Platelet (×10 ⁹ /L)	222.00 (180.00, 280.00)	208.00 (169.00, 251.00)	0.009
Haemoglobin (g/L)	129.00 (109.50, 140.00)	132.00 (117.00, 145.00)	0.008
RDW-CV (%)	13.00 (12.50, 13.60)	12.80 (12.40, 13.30)	0.01
Hs-CRP (mg/L)	1.59 (0.81, 4.55)	1.47 (0.66, 4.01)	0.659
ALT (U/L)	18.00 (12.50, 29.50)	21.00 (14.00, 32.00)	0.062
AST (U/L)	18.00 (15.00, 26.00)	19.00 (15.00, 26.00)	0.934
T2DM, type 2 diabetes mellitus; MI, myocardial infarction; AMI, acute myocardial infarction; IABP, intra-aortic balloon pump; LVEF, left ventricular ejection fraction; RDW-CV, red blood cell distribution width-coefficient of variation; MAP, mean arterial pressure; SBP, systolic blood pressure; DBP, diastolic blood pressure; POAF, postoperative atrial fibrillation; BMI, body mass index; Hs-CRP, high-sensitivity C-reactive protein; ALT, alanine aminotransferase; AST, aspartate aminotransferase; TG, triglycerides; TC, total cholesterol; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; BUN, blood urea nitrogen; UA, uric acid; FBG, fasting blood glucose; HCY, homocysteine; GA, glycated albumin; CK, creatine kinase; CK-MB, creatine kinase isoenzyme MB; LDH, lactate dehydrogenase; TP, total protein; AGR, albumin-to-globulin ratio; BNP, B-type natriuretic peptide; PT, prothrombin time; APTT, activated partial thromboplastin time; INR, international normalised ratio; FDP, fibrin degradation products; PF, plasma fibrinogen.			

Variables	IS patients (n = 127)	Non-IS patients (n = 8,020)	P-value
TC(mmol/L)	3.91 (3.24, 4.72)	3.84 (3.27, 4.50)	0.221
HDL-C (mmol/L)	0.97 (0.84, 1.14)	0.97 (0.84, 1.13)	0.381
TG (mmol/L)	1.42 (1.16, 2.08)	1.43 (1.08, 1.95)	0.903
Continued Table 1. Baseline characteristics of patients in the BAH-CMU training cohort			
LDL-C (mmol/L)	2.19 (1.67, 2.80)	2.15 (1.69, 2.69)	0.316
BUN (mmol/L)	5.62 (4.56, 7.59)	5.60 (4.60, 6.88)	0.14
Serum creatinine (μmol/L)	73.30 (61.95, 89.85)	73.50 (64.20, 85.00)	0.946
UA (μmol/L)	317.90 (262.25, 385.85)	320.10 (262.80, 384.80)	0.512
FBG (mmol/L)	6.78 (5.21, 9.36)	5.96 (5.02, 7.78)	0.004
GA (%)	16.84 (14.02, 19.92)	15.40 (13.73, 18.55)	0.013
HCY (μmol/L)	14.00 (11.55, 17.45)	14.10 (11.80, 18.40)	0.525
CK (U/L)	62.00 (45.50, 87.00)	67.00 (50.00, 95.00)	0.475
CK-MB (μg/mL)	2.10 (1.40, 3.15)	2.10 (1.30, 4.00)	0.567
LDH (U/L)	176.00 (158.00, 200.50)	175.00 (155.00, 200.00)	0.247
Myoglobin (μg/L)	107.00 (23.45, 295.00)	169.00 (25.30, 302.18)	0.497
Serum Na + (mmol/L)	140.10 (137.95, 142.25)	139.80 (137.80, 141.80)	0.248
Serum Cl - (mmol/L)	102.90 (100.85, 104.70)	103.30 (101.30, 105.00)	0.518

T2DM, type 2 diabetes mellitus; MI, myocardial infarction; AMI, acute myocardial infarction; IABP, intra-aortic balloon pump; LVEF, left ventricular ejection fraction; RDW-CV, red blood cell distribution width-coefficient of variation; MAP, mean arterial pressure; SBP, systolic blood pressure; DBP, diastolic blood pressure; POAF, postoperative atrial fibrillation; BMI, body mass index; Hs-CRP, high-sensitivity C-reactive protein; ALT, alanine aminotransferase; AST, aspartate aminotransferase; TG, triglycerides; TC, total cholesterol; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; BUN, blood urea nitrogen; UA, uric acid; FBG, fasting blood glucose; HCY, homocysteine; GA, glycated albumin; CK, creatine kinase; CK-MB, creatine kinase isoenzyme MB; LDH, lactate dehydrogenase; TP, total protein; AGR, albumin-to-globulin ratio; BNP, B-type natriuretic peptide; PT, prothrombin time; APTT, activated partial thromboplastin time; INR, international normalised ratio; FDP, fibrin degradation products; PF, plasma fibrinogen.

Variables	IS patients (n = 127)	Non-IS patients (n = 8,020)	P-value
Continued Table 1. Baseline characteristics of patients in the BAH-CMU training cohort			
TP (g/L)	69.30 (65.90, 74.10)	68.80 (65.10, 72.50)	0.044
AGR	1.56 (1.40, 1.73)	1.62 (1.46, 1.80)	0.007
BNP (pg/mL)	152.00 (76.00, 333.00)	122.00 (52.00, 256.00)	0.017
PT (s)	11.50 (11.00, 12.00)	11.50 (11.00, 12.20)	0.925
APTT (s)	31.20 (29.15, 33.20)	31.10 (28.90, 33.50)	0.285
INR	1.01 (0.97, 1.06)	1.02 (0.97, 1.08)	0.742
FDP (µg/mL)	0.87 (0.40, 1.60)	0.70 (0.30, 1.30)	0.045
PF (g/L)	3.42 (3.00, 4.10)	3.20 (2.80, 3.70)	< 0.001
D-dimer (µg/mL)	138.00 (83.00, 255.00)	121.00 (76.00, 199.00)	0.049
T2DM, type 2 diabetes mellitus; MI, myocardial infarction; AMI, acute myocardial infarction; IABP, intra-aortic balloon pump; LVEF, left ventricular ejection fraction; RDW-CV, red blood cell distribution width-coefficient of variation; MAP, mean arterial pressure; SBP, systolic blood pressure; DBP, diastolic blood pressure; POAF, postoperative atrial fibrillation; BMI, body mass index; Hs-CRP, high-sensitivity C-reactive protein; ALT, alanine aminotransferase; AST, aspartate aminotransferase; TG, triglycerides; TC, total cholesterol; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; BUN, blood urea nitrogen; UA, uric acid; FBG, fasting blood glucose; HCY, homocysteine; GA, glycated albumin; CK, creatine kinase; CK-MB, creatine kinase isoenzyme MB; LDH, lactate dehydrogenase; TP, total protein; AGR, albumin-to-globulin ratio; BNP, B-type natriuretic peptide; PT, prothrombin time; APTT, activated partial thromboplastin time; INR, international normalised ratio; FDP, fibrin degradation products; PF, plasma fibrinogen.			

Variable selection and model construction

For feature selection, LASSO regression was used to identify 22 features and rank their importance. The top 10 features, in descending order of importance, were fibrinogen degradation products (FDP), postoperative atrial fibrillation (POAF), female sex, severe carotid artery stenosis (stenosis degree $\geq 70\%$), red blood cell distribution width-coefficient of variation (RDW-CV), type 2 diabetes mellitus (T2DM), fasting blood glucose (FBG), plasma fibrinogen (PF), history of stroke, and prolonged intraoperative hypotension (< 65 mmHg). Subsequent modelling processes utilised these 10 predictors. The coefficient path and cross-validation curve plots of the LASSO regression are shown in **Supplementary Fig. 2A and 2B**.

Predictive performance of ML models

The performance of various ML models in predicting IS was analysed (Fig. 1). Receiver operating characteristic (ROC) curves for all five models were plotted. These models exhibited varying abilities to predict IS. Among the five classification algorithms, XGBoost achieved the highest AUC of 0.863 (95% CI, 0.793–0.933). Furthermore, the XGBoost model was tested externally, with an AUC of 0.786 (95% CI, 0.68–0.891) and 0.827 (95% CI, 0.765–0.890) for the PKUPH test dataset and combined BAH-CMU/PKUPH test cohorts, respectively, reflecting its effective predictive capabilities (Fig. 2 and **Supplementary Table 3**). Moreover, the calibration plots confirmed that the risks of IS evaluated using the XGBoost algorithm aligned well with the actual observed outcomes. The XGBoost model exhibited high consistency between predicted risks and actual IS observations in patients undergoing OPCABG (Fig. 3). To investigate the net benefit of the XGBoost model, a DCA was conducted. The analysis revealed that the model exhibited favourable clinical utility (Fig. 4), suggesting that our prediction model had significant clinical relevance.

Model interpretation and clinical practicality of POISM

The computed feature importance using the SHAP method is shown in **Supplementary Fig. 3**. Based on this analysis, 10 key features were selected, including prolonged intraoperative hypotension (intraoperative_low_MAP_duration), history of stroke (IS_history), PF (PF), FBG (FBG), T2DM (diabetes_mellitus), RDW-CV (RDW_CV), severe carotid artery stenosis (carotid_stenosis), female sex (female), POAF (POAF), and FDP (FDP).

Practical applications of POISM (accessible at https://qfsky.shinyapps.io/ais_opcabg_prediction/) are illustrated in **Supplementary Table 3**, with an example depicting its execution process. In Case 1, a 63-year-old man with prolonged intraoperative hypotension and severe carotid stenosis had a predicted probability of 2.36% for developing IS. Based on this risk evaluation, the medical team adopted a proactive approach, such as blood pressure monitoring and maintenance of sufficient blood volume and cerebral perfusion, to prevent IS occurrence. In Case 2, a 66-year-old woman with pre-existing T2DM exhibited significantly elevated preoperative FBG levels, with a probability (2.31%) of developing IS. Accordingly, the clinical team prioritised perioperative glycemic control by adjusting the hypoglycemic regimen, consequently preventing IS.

Discussion

In this multicentre, retrospective cohort study, a new ML-based tool to predict IS risk following OPCABG was developed. This solution is effective, exemplified by multiple instances in which ML-based methods outperformed conventional risk stratification approaches [24–27]. We applied various advanced ML algorithms and assessed each model to detect patients at high IS risk. Consequently, the XGBoost model demonstrated a superior discriminative capability, with the highest sensitivity and specificity. Based on this model, POISM, an easy-to-use website tool, was created and displayed favourable clinical utility.

Among all the variables, female sex, T2DM, history of stroke, RDW-CV, FBG, PF, FDP, carotid stenosis $\geq 70\%$, prolonged intraoperative hypotension, and POAF were identified as the key predictors. Prolonged intraoperative hypotension, the greatest contributor to model output, is associated with worsened survival and increased incidence of stroke [28, 29]. Although its pathophysiological role in postoperative stroke remains unclear, multiple mechanisms, including catecholamine increase, mechanical trauma to red blood cells triggering inflammatory conditions, and reduced vasomotor reactivity after CPB in the presence of hypotension, may aggravate end-organ damage, leading to adverse outcomes such as stroke [29]. Hence, continuous invasive blood pressure monitoring is essential for optimising fluid management and maintaining hemodynamic stability throughout the perioperative period.

Patients' baseline characteristics, such as previous IS and severe carotid stenosis ($\geq 70\%$), two conventional risk factors, are directly linked to IS occurrence [8, 22, 30–33]. The most commonly used preoperative and intraoperative interventions for carotid stenosis include pharmacological therapy and revascularisation options, such as carotid artery stenting and carotid endarterectomy. However, the efficacy of these revascularisation procedures in reducing postoperative stroke among patients with severe carotid stenosis remains elusive [32, 34]. A better understanding of the effectiveness of these interventions will improve clinical practice.

Consistent with previous research [8, 9, 30, 35–37], our study demonstrated a significant impact of sex on perioperative stroke occurrence, with female sex showing a remarkably higher IS risk than males. Notably, a study by Keeling et al. revealed that female sex was associated with an increased 30-day mortality, attributed to anatomical and physiological factors [36]. Hence, sufficient postoperative attention should be paid to female patients to alleviate these risks.

In the present study, a significant association between T2DM and postoperative IS was observed, consistent with previous studies [9]. Elevated glycemic levels are correlated with both stroke incidence and poor short-term prognosis following CABG [38–43]. Glycemic levels directly influence the occurrence of post-CABG stroke. Acute hyperglycemia decreases nitric oxide bioavailability, impairs endothelial function, increases platelet aggregation, and stimulates coagulation [44], suggesting that it increases IS risk. Therefore, controlling preoperative blood glycemic levels is crucial to reduce the incidence of postoperative IS. Based on these findings, it was recommended that perioperative blood glucose levels should be controlled below 9.99 mmol/L [45].

Hypercoagulability is associated with a greater risk for MI, stroke, and mortality in the early period after CABG. We found that PF and FDP levels were associated with IS. Corroborating this finding, a recent study revealed that elevated fibrinogen levels could be a causative risk factor for IS, with the risk doubling for each 1 g/L increase. However, the underlying pathogenesis of CAD complicated by IS after CABG remains unclear. Therefore, we hypothesise that increased PF production directly activates multiple mechanisms, accelerates atherosclerosis and thrombogenesis [46], and consequently results in IS. However, this issue needs to be clarified in future studies.

As an independent risk factor, red blood cell distribution width (RDW) has been used for predicting post-CABG IS and transient ischemic attack [23, 47, 48]. We validated this finding by revealing a clear link between the RDW-CV and IS risk. Although the pathogenic role of increased RDW in IS remains largely unknown, it has been speculated to be associated with inflammatory status and oxidative stress [48, 49]. Nevertheless, future studies with larger sample sizes are needed to further confirm its role and clinical utility in predicting IS across diverse populations.

POAF is an important and potentially modifiable risk factor [50, 51]. POAF is associated with increased risk of stroke and in-hospital mortality [52] and it also contributes to additional treatment, prolonged hospitalisation, and increased costs [53–55]. The aetiology likely involves multiple factors, such as pericardial inflammation, autonomic stimulation, oxidative stress, and electrolyte disturbances [56]. Additionally, a recent study reported that short-term IS risk in patients with POAF after CABG was associated with prior IS, older age, heart failure, and MI [57]. Therefore, preventive measures including β -blockers and careful periprocedural anticoagulation for POAF may help reduce postoperative stroke [58, 59].

Among many algorithm-based models, XGBoost is a widely preferred model used for clinical prediction of several human diseases such as brain injury, and COVID-19 [60–62]. Because of the presence of numerous intricate variables in CAD admissions with unclear relationships, the XGBoost model was used to enhance data analysis and prediction accuracy in this study. Although the XGBoost model is liable to overfit, this issue was addressed by persistently refining this model with a large sample size and validating it independently using multicentre cohorts in our study [63]. Furthermore, the SHAP method was utilised to elucidate the XGBoost model, in order to better understand how variables affect outcomes. More importantly, integration of the SHAP method with XGBoost model resulted in POISM, a website tool that can be easily used by clinicians to guide decision-making in treating patients at risk of IS.

Conclusion

This study demonstrated for the first time that the ML model, particularly XGBoost, performs well in predicting in-hospital postoperative IS in patients undergoing OPCABG. POISM was created based on this model, and it can be conveniently used to evaluate IS risk effectively and serves as a reliable tool for targeted management following OPCABG. Hence, this study provides a solid foundation for early identification and therapy for patients at risk of IS, leading to improved outcomes and reduced economic burden.

Limitations

Our study has several limitations. First, the incidence of post-CABG IS was low. Although more than 10,000 cases were enrolled, the limited number of positive cases may compromise overall statistical power and hinder robust statistical analysis, affecting the reliability of the findings. Second, as a

retrospective study, the results may have been affected by incomplete data recording and missing values. Third, owing to limitations in longitudinal follow-up data from the hospital's electronic records, we defined in-hospital postoperative IS as the primary outcome rather than 30-day IS after OPCABG. Finally, all data were sourced from the Beijing region, which may introduce selection bias and constrain the generalisability of the results. Hence, future multicentre prospective studies on a larger scale are warranted to validate our findings.

Future work

Future studies should confirm the associations identified in our ML-based analysis using prospective data. Conducting larger multicentre studies across diverse regions and performing external validation will improve the generalisability of the prediction model to other populations and clinical scenarios. Additionally, a comprehensive comparison of the XGBoost model with traditional models (e.g., the STS model) is needed to assess their predictive performance. Investigating the efficacy of targeted interventions (e.g., intraoperative hypotension management) and identifying other potential risk factors (e.g., family history, environmental influences, and lifestyle choices) may provide deeper insight into the determinants contributing to in-hospital IS risk after OPCABG.

Abbreviations

AUC

Area under the receiver operating characteristic curve

CABG

Coronary artery bypass grafting

CAD

Coronary artery disease

FBG

Fasting blood glucose

FDP

Fibrinogen degradation products

IS

Ischemic stroke

ML

Machine learning

OPCABG

Off-pump coronary artery bypass grafting

POAF

Postoperative atrial fibrillation

PF

Plasma fibrinogen

RDW-CV

Red blood cell distribution width-coefficient of variation
SHAP
Shapley Additive exPlanation
T2DM
Type 2 diabetes mellitus

Declarations

Ethics approval and consent to participate

Our research was conducted in accordance with the Declaration of Helsinki, and was approved by the ethical review board of Beijing Anzhen Hospital and Peking University People's Hospital.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Funding

The present work was supported by China National Natural Science Foundation (82071342).

Author Contribution

Z.D.K., W.L.Z., G.Z.L., and W.Q.Z. conceived and designed the study. C.L., H.X., C.B.W., and L.F.Z. performed the data collection. D.M., X.Q.L., and B.C. carried out the data analysis and interpretation. R.Y.G. and X.B.Y. developed the methodology and provided software support. Q.H.C. and G.P.F. prepared the figures and tables. Z.D.K. and W.L.Z. drafted the main manuscript text. All authors reviewed the manuscript and approved the final version.

Acknowledgements

We express our sincere gratitude to all the participants, research assistants, and outcome assessors involved in this study. Our thanks also go to Editage (www.editage.com) for English language editing.

Data Availability

Data availability statement The clinical data that support the findings of this study were obtained from [Hospital/Institution Name] and are subject to ethical and privacy restrictions. These data are not publicly available due to patient confidentiality and institutional review board (IRB)/ethics committee regulations. De-identified data may be available from the corresponding author upon reasonable request, subject to institutional approval and a signed data access agreement. Requests for data access should be directed to the corresponding author at email: liugz1997@pumc.edu.cn

References

1. Head SJ, Börgermann J, Osnabrugge RLJ, Kieser TM, Falk V, Taggart DP, et al. Coronary artery bypass grafting: Part 2—Optimizing outcomes and future prospects. *Eur Heart J*. 2013;34:2873–86. 10.1093/eurheartj/ehs284.
2. Yoo K-J. The past, present, and future of off-pump coronary artery bypass grafting. *J Chest Surg*. 2025;58:121–33. 10.5090/jcs.24.122.
3. Ramsingh R, Bakaeen FG. Coronary artery bypass grafting: Practice trends and projections. *Cleve Clin J Med*. 2025;92:181–91. 10.3949/ccjm.92a.23071.
4. Bakaeen FG, Gaudino M, Whitman G, Doenst T, Ruel M, Taggart DP et al. 2021: The American Association for Thoracic Surgery Expert Consensus Document: Coronary artery bypass grafting in patients with ischemic cardiomyopathy and heart failure. *J Thorac Cardiovasc Surg*. 2021;162:829–850.e1. 10.1016/j.jtcvs.2021.04.052
5. Lorusso R, Moscarelli M, Di Franco A, Grazioli V, Nicolini F, Gherli T, et al. Association between coronary artery bypass surgical techniques and postoperative stroke. *J Am Heart Assoc*. 2019;8:e013650. 10.1161/JAHA.119.013650.
6. Mohamed MO, Hirji S, Mohamed W, Percy E, Bradley P, Chung J, et al. Incidence and predictors of postoperative ischemic stroke after coronary artery bypass grafting. *Int J Clin Pract*. 2021;75:e14067. 10.1111/ijcp.14067.
7. Redžek A, Mironicki M, Gvozdenović A, Petrović M, Čemerlić-Adić N, Ilić A, et al. Predictors for hospital readmission after cardiac surgery. *J Card Surg*. 2015;30:1–6. 10.1111/jocs.12441.
8. Jonsson K, Barbu M, Nielsen SJ, Hafsteinsdottir B, Gudbjartsson T, Jensen EM, et al. Perioperative stroke and survival in coronary artery bypass grafting patients: A SWEDEHEART study. *Eur J Cardiothorac Surg*. 2022;62:ezac025. 10.1093/ejcts/ezac025.
9. Sultan I, Bianco V, Kilic A, Jovin T, Jadhav A, Jankowitz B, et al. Predictors and outcomes of ischemic stroke after cardiac surgery. *Ann Thorac Surg*. 2020;110:448–56. 10.1016/j.athoracsur.2020.02.025.
10. Shahian DM, Jacobs JP, Badhwar V, Kurlansky PA, Furnary AP, Cleveland JC Jr, et al. The Society of Thoracic Surgeons 2018 adult cardiac surgery risk models: part 1—background, design considerations, and model development. *Ann Thorac Surg*. 2018;105:1411–8. 10.1016/j.athoracsur.2018.03.002.
11. Nashef SAM, Roques F, Sharples LD, Nilsson J, Smith C, Goldstone AR, et al. EuroSCORE II. *Eur J Cardiothorac Surg*. 2012;41:734–44. 10.1093/ejcts/ezs043. discussion 744.

12. Choi JW, Hwang HY, Kim KB. The performance of the current risk prediction scoring systems in patients undergoing anaortic off-pump coronary artery bypass grafting. *J Korean Med Sci.* 2021;36:e163. 10.3346/jkms.2021.36.e163.
13. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* 2015;521:436–44. 10.1038/nature14539.
14. Delen D, Walker G, Kadam A. Predicting breast cancer survivability: A comparison of three data mining methods. *Artif Intell Med.* 2005;34:113–27. 10.1016/j.artmed.2004.07.002.
15. Shi J-Y, Wang X, Ding G-Y, Dong Z, Han J, Guan Z, et al. Exploring prognostic indicators in the pathological images of hepatocellular carcinoma based on deep learning. *Gut.* 2021;70:951–61. 10.1136/gutjnl-2020-320930.
16. Deo RC. Machine learning in medicine. *Circulation.* 2015;132:1920–30. 10.1161/CIRCULATIONAHA.115.001593.
17. Song Y, Zhai W, Ma S, Wu Y, Ren M, Van den Eynde J, et al. Machine learning-based prediction of off-pump coronary artery bypass grafting-associated acute kidney injury. *J Thorac Dis.* 2024;16:4535–42. 10.21037/jtd-24-711.
18. Parise O, Parise G, Vaidyanathan A, Occhipinti M, Gharaviri A, Tetta C, et al. Machine learning to identify patients at risk of developing new-onset atrial fibrillation after coronary artery bypass. *J Cardiovasc Dev Dis.* 2023;10:82. 10.3390/jcdd10020082.
19. Lee S, Song I-U, Na S-H, Jeong D-S, Chung S-W. Association between long-term functional outcome and change in hs-CRP level in patients with acute ischemic stroke. *Neurologist.* 2020;25:122–5. 10.1097/NRL.0000000000000278.
20. Latchaw RE, Alberts MJ, Lev MH, Connors JJ, Harbaugh RE, Higashida RT, et al. Recommendations for imaging of acute ischemic stroke: A scientific statement from the American Heart Association. *Stroke.* 2009;40:3646–78. 10.1161/STROKEAHA.108.192616.
21. Wintermark M, Sanelli PC, Albers GW, Bello J, Derdeyn C, Hetts SW, et al. Imaging recommendations for acute stroke and transient ischemic attack patients: A joint statement by the American Society of Neuroradiology, the American College of Radiology, and the Society of Neurointerventional Surgery. *AJNR Am J Neuroradiol.* 2013;34:E117–27. 10.3174/ajnr.A3690.
22. Mao Z, Zhong X, Yin J, Zhao Z, Hu X, Hackett ML. Predictors associated with stroke after coronary artery bypass grafting: A systematic review. *J Neurol Sci.* 2015;357:1–7. 10.1016/j.jns.2015.07.006.
23. Tatlisuluoglu D, Tezcan B, Mungan İ, Çakirli YA, Tümer NB, Taşoğlu İ. Predicting postoperative ischemic stroke problems in patients following coronary bypass surgery using neutrophil-lymphocyte ratio, platelet-lymphocyte ratio, and red blood cell distribution width values. *Kardiochir Torakochirurgia Pol.* 2022;19:90–5. 10.5114/kitp.2022.117499.
24. Brinkhaus HO, Rajan K, Schaub J, Zielesny A, Steinbeck C. Open data and algorithms for open science in AI-driven molecular informatics. *Curr Opin Struct Biol.* 2023;79:102542. 10.1016/j.sbi.2023.102542.
25. Hindocha S, Charlton TG, Linton-Reid K, Hunter B, Chan C, Ahmed M, et al. A comparison of machine learning methods for predicting recurrence and death after curative-intent radiotherapy for non-

- small cell lung cancer: Development and validation of multivariable clinical prediction models. *EBiomedicine*. 2022;77:103911. 10.1016/j.ebiom.2022.103911.
26. Salikhanov I, Roth V, Gahl B, Reid G, Kolb R, Dimanski D, et al. Machine learning-based prediction of short-term mortality after coronary artery bypass grafting: A retrospective cohort study. *Biomedicines*. 2025;13:2023. 10.3390/biomedicines13082023.
 27. Zou W, Zhao H, Ren M, Cui C, Yuan G, Yuan B, et al. Risk factors and prediction model for acute ischemic stroke after off-pump coronary artery bypass grafting based on Bayesian network. *BMC Med Inf Decis Mak*. 2024;24:349. 10.1186/s12911-024-02762-2.
 28. Sun LY, Chung AM, Farkouh ME, van Diepen S, Weinberger J, Bourke M, et al. Defining an intraoperative hypotension threshold in association with stroke in cardiac surgery. *Anesthesiology*. 2018;129:440–7. 10.1097/ALN.0000000000002298.
 29. de La Hoz MA, Rangasamy V, Bastos AB, Xu X, Novack V, Saugel B, et al. Intraoperative hypotension and acute kidney injury, stroke, and mortality during and outside cardiopulmonary bypass: A retrospective observational cohort study. *Anesthesiology*. 2022;136:927–39. 10.1097/ALN.0000000000004175.
 30. Mehta A, Choxi R, Gleason T, Wechsler L, Jovin T, Thirumala PD. Carotid artery disease as a predictor of in-hospital postoperative stroke after coronary artery bypass grafting from 1999 to 2011. *J Cardiothorac Vasc Anesth*. 2018;32:1587–96. 10.1053/j.jvca.2017.10.010.
 31. Wang S, Ran Y, Cheng S, Lyu Y, Liu J. Determinants and clinical outcomes of stroke following revascularization among patients with reduced ejection fraction. *Brain Behav*. 2023;13:e2927. 10.1002/brb3.2927.
 32. Chen S, Mi C, Zhang S, Li Y, Yun Y, Zhang X, et al. The role of carotid artery stenosis in predicting stroke after coronary artery bypass grafting in a Chinese cohort study. *Sci Rep*. 2023;13:21536. 10.1038/s41598-023-47640-5.
 33. Masabni K, Raza S, Blackstone EH, Gornik HL, Sabik JF. Does preoperative carotid stenosis screening reduce perioperative stroke in patients undergoing coronary artery bypass grafting? *J Thorac Cardiovasc Surg*. 2015;149:1253–60. 10.1016/j.jtcvs.2015.02.003.
 34. Klarin D, Patel VI, Zhang S, Xian Y, Kosinski A, Yerokun B, et al. Concomitant carotid endarterectomy and cardiac surgery does not decrease postoperative stroke rates. *J Vasc Surg*. 2020;72:589–e5963. 10.1016/j.jvs.2019.10.072.
 35. Rubens FD, Femes SE, Grubic N, Fergusson D, Taljaard M, van Walraven C. Outcomes following coronary artery bypass grafting with multiple arterial grafting by pump status in men and women. *J Thorac Cardiovasc Surg*. 2024;167:1796–e180715. 10.1016/j.jtcvs.2023.02.008.
 36. Qu WW, Wei JW, Binongo JN, Keeling WB. Sex differences in failure-to-rescue after coronary artery bypass grafting. *Ann Thorac Surg*. 2022;114:1596–602. 10.1016/j.athoracsur.2021.09.070.
 37. Bryce Robinson NB, Naik A, Rahouma M, Morsi M, Wright D, Hameed I, et al. Sex differences in outcomes following coronary artery bypass grafting: A meta-analysis. *Interact Cardiovasc Thorac Surg*. 2021;33:841–7. 10.1093/icvts/ivab191.

38. Jones KW, Cain AS, Mitchell JH, Millar RC, Rimmasch HL, French TK, et al. Hyperglycemia predicts mortality after CABG: Postoperative hyperglycemia predicts dramatic increases in mortality after coronary artery bypass graft surgery. *J Diabetes Complications*. 2008;22:365–70. 10.1016/j.jdiacomp.2007.05.006.
39. Alavi-Moghaddam M, Parsa-Mahjoob M, Ghodssi-Ghassemabadi R, Bitazar B. Association of admission blood glucose level with major adverse cardiac events in acute coronary syndrome; a cohort study. *Arch Acad Emerg Med*. 2019;7:e26.
40. Thiele RH, Hucklenbruch C, Ma JZ, Colquhoun D, Zuo Z, Nemergut EC, et al. Admission hyperglycemia is associated with poor outcome after emergent coronary bypass grafting surgery. *J Crit Care*. 2015;30:1210–6. 10.1016/j.jcrc.2015.09.004.
41. Li X, Hou X, Zhang H, Qian X, Feng X, Shi N, et al. Association between stress hyperglycaemia and in-hospital cardiac events after coronary artery bypass grafting in patients without diabetes: A retrospective observational study of 5450 patients. *Diabetes Obes Metab*. 2023;25(Suppl 1):34–42. 10.1111/dom.15013.
42. McAlister FA, Man J, Bistritz L, Amad H, Tandon P. Diabetes and coronary artery bypass surgery. An examination of perioperative glycemic control and outcomes. *Diabetes Care*. 2003;26:1518–24.
43. Imran SA, Ransom TPP, Buth KJ, Clayton D, Al-Shehri B, Ur E, et al. Impact of admission serum glucose level on in-hospital outcomes following coronary artery bypass grafting surgery. *Can J Cardiol*. 2010;26:151–4. 10.1016/S0828-282X(10)70357-3.
44. Worthley MI, Holmes AS, Willoughby SR, Kucia AM, Heresztyn T, Stewart S, et al. The deleterious effects of hyperglycemia on platelet function in diabetic patients with acute coronary syndromes mediation by superoxide production, resolution with intensive insulin administration. *J Am Coll Cardiol*. 2007;49:304–10. 10.1016/j.jacc.2006.08.053.
45. Lawton JS, Tamis-Holland JE, Bangalore S, Bates ER, Beckie TM, Bischoff JM, et al. 2021 ACC/AHA/SCAI guideline for coronary artery revascularization: Executive summary: A report of the American College of Cardiology/American Heart Association joint committee on clinical practice guidelines. *Circulation*. 2021;145:e4–17. 10.1161/CIR.0000000000001039.
46. Surma S, Banach M. Fibrinogen and atherosclerotic cardiovascular diseases—review of the literature and clinical studies. *Int J Mol Sci*. 2021;23:193. 10.3390/ijms23010193.
47. Feng G-H, Li H-P, Li Q-L, Fu Y, Huang R-B. Red blood cell distribution width and ischaemic stroke. *Stroke Vasc Neurol*. 2017;2:172–5. 10.1136/svn-2017-000071.
48. Xie K-H, Liu L-L, Liang Y-R, Su C-Y, Li H, Liu R-N, et al. Red cell distribution width: A novel predictive biomarker for stroke risk after transient ischaemic attack. *Ann Med*. 2022;54:1167–77. 10.1080/07853890.2022.2059558.
49. Xue J, Zhang D, Zhang X-G, Zhu X-Q, Xu X-S, Yue Y-H. Red cell distribution width is associated with stroke severity and unfavorable functional outcomes in ischemic stroke. *Front Neurol*. 2022;13:938515. 10.3389/fneur.2022.938515.

50. Dk M, Lr R. Stroke and atrial fibrillation following cardiac surgery. *WMJ Off Publ State Med Soc Wis.* 2003;102:26–30.
51. Lahtinen J, Biancari F, Salmela E, Mosorin M, Satta J, Rainio P, et al. Postoperative atrial fibrillation is a major cause of stroke after on-pump coronary artery bypass surgery. *Ann Thorac Surg.* 2004;77:1241–4. 10.1016/j.athoracsur.2003.09.077.
52. Kosmidou I, Chen S, Kappetein AP, Serruys PW, Gersh BJ, Puskas JD, et al. New-onset atrial fibrillation after PCI or CABG for left main disease: The EXCEL trial. *J Am Coll Cardiol.* 2018;71:739–48. 10.1016/j.jacc.2017.12.012.
53. Almassi GH, Schowalter T, Nicolosi AC, Aggarwal A, Moritz TE, Henderson WG, et al. Atrial fibrillation after cardiac surgery: A major morbid event? *Ann Surg.* 1997;226:501–11. 10.1097/00000658-199710000-00011. discussion 511.
54. Creswell LL, Schuessler RB, Rosenbloom M, Cox JL. Hazards of postoperative atrial arrhythmias. *Ann Thorac Surg.* 1993;56:539–49. 10.1016/0003-4975(93)90894-N.
55. Hakala T, Pitkänen O, Hippeläinen M. Feasibility of predicting the risk of atrial fibrillation after coronary artery bypass surgery with logistic regression model. *Scand J Surg.* 2002;91:339–44. 10.1177/145749690209100406.
56. Halonen J, Kärkkäinen J, Jäntti H, Martikainen T, Valtola A, Ellam S, et al. Prevention of atrial fibrillation after cardiac surgery: A review of literature and comparison of different treatment modalities. *Cardiol Rev.* 2024;32:248–56. 10.1097/CRD.0000000000000499.
57. Taha A, Martinsson A, Nielsen SJ, Rezk M, Pivodic A, Gudbjartsson T, et al. New-onset atrial fibrillation after coronary surgery and stroke risk: A nationwide cohort study. *Heart.* 2024;111:18–26. 10.1136/heartjnl-2024-324573.
58. Neumann F-J, Sousa-Uva M, Ahlsson A, Alfonso F, Banning AP, Benedetto U, et al. Linee guida ESC/EACTS 2018 sulla rivascularizzazione miocardica. Task force sulla Rivascularizzazione Miocardica della Società Europea di Cardiologia (ESC) e dell'Associazione Europea di Chirurgia Cardiotoracica (EACTS) [ESC/EACTS Guidelines on myocardial revascularization]. *G Ital Cardiol (Rome).* 2019;20:S1–61. 10.1714/3203.31801.
59. ., Authors/Task Force Members, Jeppsson A, (Co-Chairperson) (Sweden), Rocca B, (Co-Chairperson) (Italy), Hansson EC et al. 2024 EACTS Guidelines on perioperative medication in adult cardiac surgery. *Eur J Cardiothorac Surg.* 2024;67:ezae355. 10.1093/ejcts/ezae355
60. Wang R, Wang L, Zhang J, He M, Xu J. XGBoost machine learning algorism performed better than regression models in predicting mortality of moderate-to-severe traumatic brain injury. *World Neurosurg.* 2022;163:e617–22. 10.1016/j.wneu.2022.04.044.
61. Fang ZG, Yang SQ, Lv CX, An SY, Wu W. Application of a data-driven XGBoost model for the prediction of COVID-19 in the USA: A time-series study. *BMJ Open.* 2022;12:e056685. 10.1136/bmjopen-2021-056685.
62. Rahman MS, Chowdhury AH. A data-driven eXtreme gradient boosting machine learning model to predict COVID-19 transmission with meteorological drivers. *PLoS ONE.* 2022;17:e0273319.

63. Shah NH, Milstein A, Bagley PhD SC. Making machine learning models clinically useful. JAMA. 2019;322:1351–2. 10.1001/jama.2019.10306.

Figures

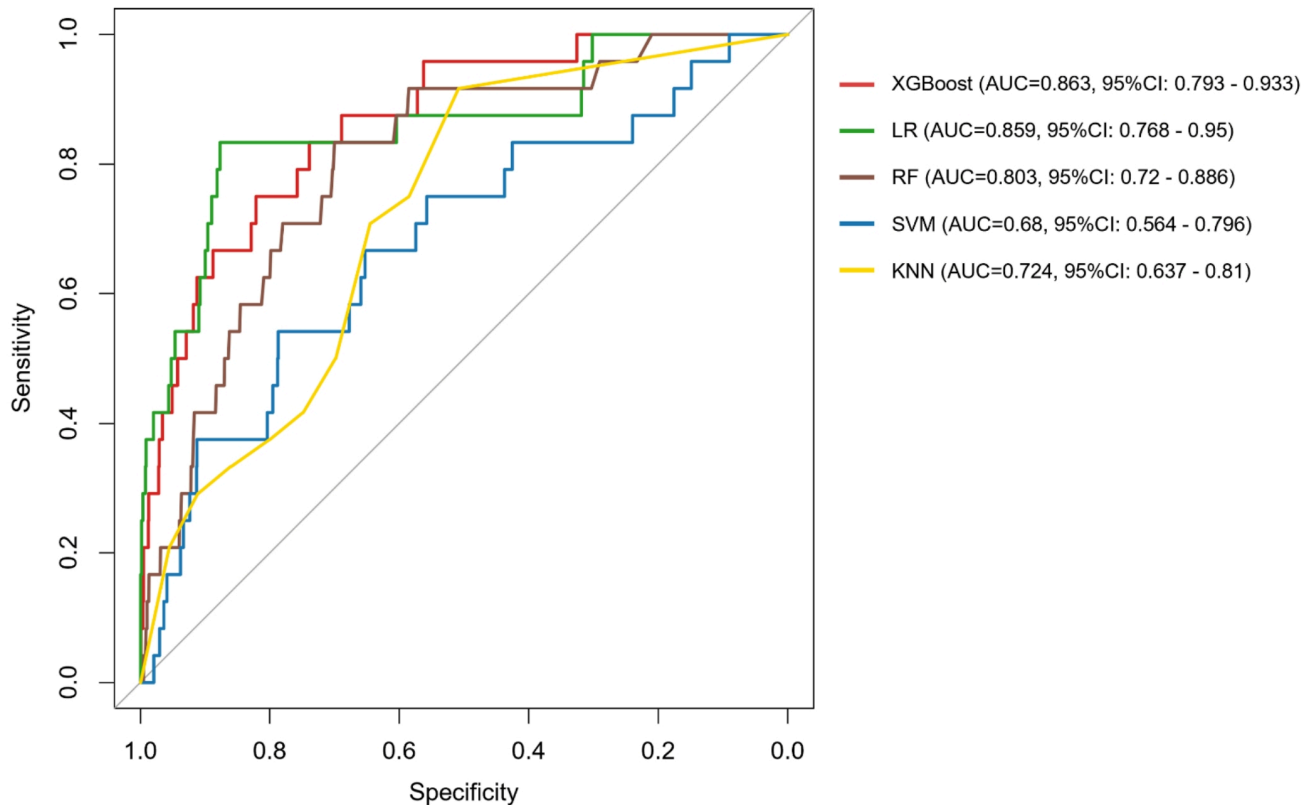


Figure 1

Plot of the area under the receiver operating characteristic (ROC) curve of different machine learning (ML) models in the BAH-CMU test dataset. BAH-CMU, Beijing Anzhen Hospital of Capital Medical University. XGBoost, eXtreme Gradient Boosting; LR, Logistic Regression; RF, Random Forest; SVM, Support Vector Machine; KNN, k-Nearest Neighbours; AUC, Area Under the ROC Curve; CI, Confidence Interval.

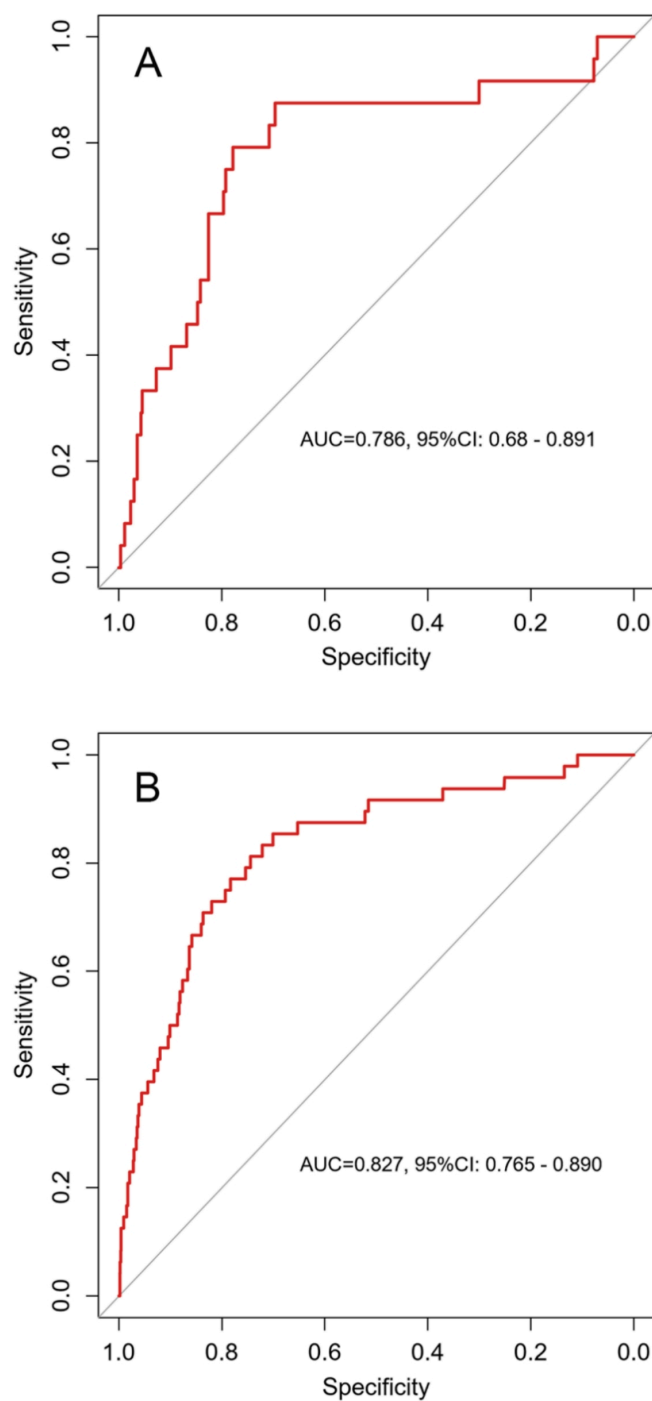


Figure 2

Plot of the area under the receiver operating characteristic (ROC) curve of the XGBoost model in (A) the PKUPH-test cohort, and (B) the combined test (BAH-CMU-test and PKUPH) cohorts. BAH-CMU, Beijing Anzhen Hospital of Capital Medical University; PKUPH, Peking University People's Hospital.

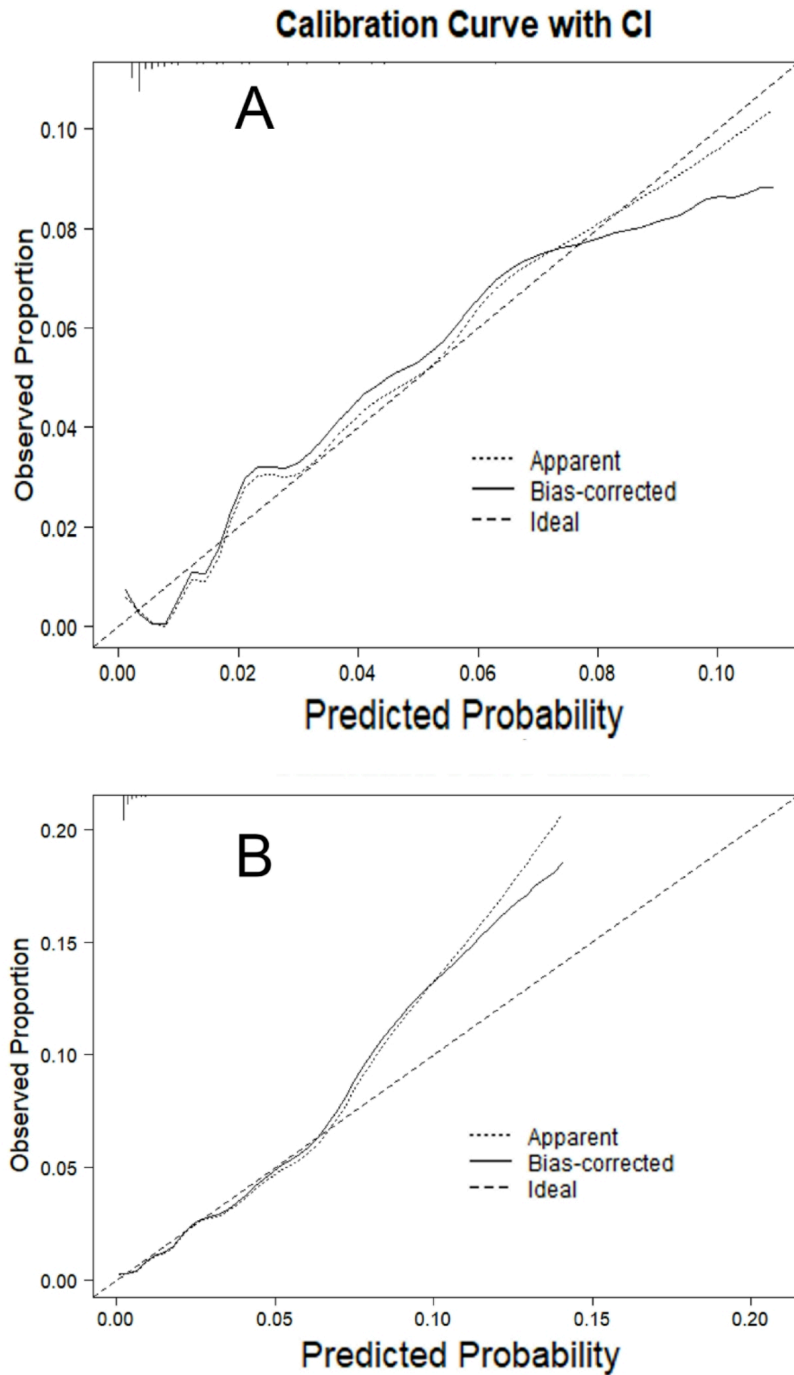


Figure 3

Calibration plots for the XGBoost model in (A) the PKUPH-test cohort, and (B) the combined test (BAH-CMU-test and PKUPH) cohorts. BAH-CMU, Beijing Anzhen Hospital of Capital Medical University; PKUPH, Peking University People's Hospital.

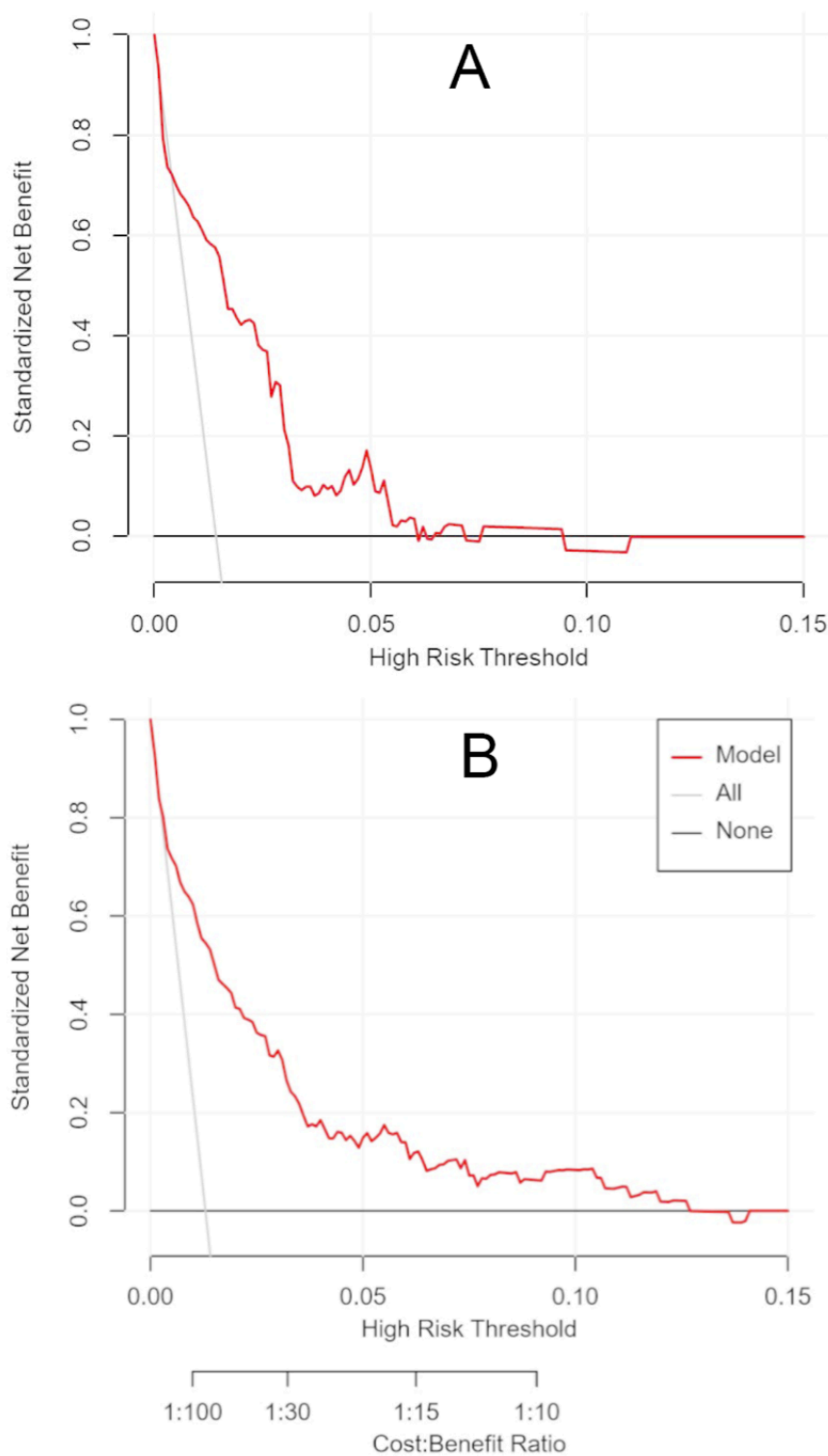


Figure 4

Decision curve analysis (DCA) plots for the XGBoost model in (A) the PKUPH-test cohort, and (B) the combined test (BAH-CMU-test and PKUPH) cohorts. BAH-CMU, Beijing Anzhen Hospital of Capital Medical University; PKUPH, Peking University People's Hospital.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [MLCABGAdditionalFilev28.docx](#)