

Machine learning model-based risk prediction of severe complications after off-pump coronary artery bypass grafting

Yang Zhang^{1,A–D,F}, Lin Li^{1,A,B,D–F}, Ye Li^{2,A,F}, Zhihe Zeng^{1,B,C,F}

¹ Department of Anesthesiology, General Hospital of Northern Theater Command, Shenyang, China

² Microsoft (China) Co., Ltd, Beijing, China

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation; D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online)

Adv Clin Exp Med. 2023;32(2):185–194

Address for correspondence

Lin Li
E-mail: lilinlashofmine@qq.com

Funding sources

This work was supported by the Key R&D Guiding Program of Liaoning Province 2019010209 (grant No. 2019JH8/10300083).

Conflict of interest

None declared

Received on April 14, 2022

Reviewed on June 10, 2022

Accepted on August 23, 2022

Published online on October 13, 2022

Cite as

Zhang Y, Li L, Li Y, Zeng Z. Machine learning model-based risk prediction of severe complications after off-pump coronary artery bypass grafting. *Adv Clin Exp Med.* 2023;32(2):185–194. doi:10.17219/acem/152895

DOI

10.17219/acem/152895

Copyright

Copyright by Author(s)

This is an article distributed under the terms of the Creative Commons Attribution 3.0 Unported (CC BY 3.0) (<https://creativecommons.org/licenses/by/3.0/>)

Abstract

Background. Compared with coronary artery bypass grafting (CABG) under cardiopulmonary bypass, off-pump coronary artery bypass (OPCAB) is minimally invasive and reduces the risk of intraoperative blood transfusion and acute kidney injury. Nonetheless, OPCAB-related complications still pose a threat. Machine learning technology can analyze a large number of clinical data, establish risk prediction models and help clinicians make early and correct clinical decisions.

Objectives. Risk prediction models are available for mortality and morbidity after cardiac surgery, but they are not specific to OPCAB. This study aimed to develop a predictive model of severe complications after OPCAB, based on machine learning.

Materials and methods. Anesthesia records of OPCAB from the General Hospital of the Northern Theater Command (Shenyang, China) collected between January 1, 2019, and June 15, 2020, were analyzed. The endpoint of the study was the occurrence of serious complications after OPCAB (postoperative unplanned intra-aortic balloon pump, secondary surgery and death). The features entered into the models were as follows: intraoperative ventricular fibrillation, number of saphenous vein grafts, nerve block (NeB), venous oxygen saturation (SvO₂), skin incision-bypass time, and hypertension. A total of 8 machine learning algorithms were tested: logistic regression analysis (LRA), k-nearest neighbor (KNN), naïve Bayes (NB), support vector machine (SVM), random forest (RF), extreme gradient boosting (XGBoost), light gradient boosting machine (LightGBM), and categorical features gradient boosting (CatBoost).

Results. Among the 506 patients found in the records, 27 met the endpoint. The highest area under the curve (AUC) value was achieved with the XGBoost model (AUC = 0.94), and the lowest with the SVM model (AUC = 0.75). The highest and lowest accuracy were observed with the XGBoost and NB models, respectively, while the highest and lowest precision were achieved using the SVM and NB models, respectively. Based on the receiver operating characteristic (ROC) curves, the XGBoost model was selected as the most useful in this study.

Conclusions. This study suggests using the XGBoost model to predict the risk of complications after OPCAB.

Key words: complications, machine learning, off-pump coronary artery bypass grafting, prediction model

Background

Revascularization is paramount to the management of acute coronary syndrome (ACS). It aims to improve blood flow to the myocardium¹ and is performed using percutaneous coronary intervention or coronary artery bypass graft (CABG).^{1–3} The latter can be performed either off-pump (i.e., without the assistance of a heart-lung machine) or on-pump. On-pump CABG is associated with more severe surgical trauma, while off-pump coronary artery bypass (OPCAB) can reduce perioperative bleeding and allogeneic blood transfusions, as well as reduce the risk of acute kidney injury (AKI) in patients with kidney dysfunction.⁴ The OPCAB does not appear to increase 30-day mortality compared with on-pump CABG, but an extensive systematic review of observational studies suggested that OPCAB might reduce short-term mortality.^{5,6} Therefore, OPCAB is probably a good option for selected patients.^{7,8}

There are still some risks related to the use of OPCAB,^{4,9,10} including perioperative complications such as mortality, stroke, kidney failure, respiratory failure, and blood loss.^{11–13} Furthermore, OPCAB appears to be associated with higher 10-year rates of incomplete revascularization, repeat revascularization and mortality, compared with on-pump CABG.¹⁰ Additionally, OPCAB is associated with increased adverse events at 1 year and mortality at 5 years.^{14,15} Although OPCAB has a similar risk of myocardial infarction compared to on-pump CABG, the data are inconsistent for the risk of stroke.^{5,6,16} A decreased left ventricular ejection fraction (LVEF) is observed in about 22% of the patients after OPCAB and can compromise their short- and long-term outcomes.^{17–19}

Some tools are available for estimating the risk of mortality and morbidity after CABG. The Society of Thoracic Surgeons (STS) score can be used to determine the risk of mortality and morbidity after cardiac surgery, but it is not specific to OPCAB.²⁰ The European System for Cardiac Operative Risk Evaluation (EuroSCORE) can overestimate the risk of complications in the highest-risk and lowest-risk patients undergoing CABG,²¹ as well as in patients undergoing OPCAB.²² Other risk models are available but they are not specific to OPCAB.^{23–25}

Machine learning algorithms can be used to analyze data and establish risk models more accurately than traditional statistical models.^{26,27} Indeed, machine learning has been used to create models that predict mortality after cardiac surgery,^{28–31} as well as estimate the length of hospital stay after CABG.³² Regardless, these models are still not specific to OPCAB.

Objectives

Using machine learning, this study aimed to build a predictive model for the detection of early postoperative serious complications after OPCAB. The results could

provide reference data for optimizing the clinical pathway for OPCAB, and anesthesia strategy to maintain vital signs, regulate the circulation, balance myocardial oxygen supply and demand, and reduce complications.

Materials and methods

Study design

All data were taken from the Do-care anesthesia record system (v. 5.0, MEDICAL SYSTEM Co., Ltd., Suzhou, China). All records of OPCABs performed from January 1, 2019, to June 15, 2020, at the 2nd Ward of the Department of Anesthesiology of the General Hospital of the Northern Theater Command (Shenyang, China) were included. At that hospital, OPCAB has been carried out for 20 years. In this study, 3 teams comprised of 15 surgeons were involved, who all had the qualification of chief surgeon, with an annual operation volume of 300–450 cases per surgeon.

This study was approved by the Ethics Committee of the hospital (Approval No. k(2020)01). The requirement for individual informed consent was waived by the Committee due to the retrospective nature of this study.

Inclusion and exclusion criteria

All patients who underwent CABG were screened. The exclusion criteria were: 1) CABG under cardiopulmonary bypass; 2) CABG combined with other surgical procedures; 3) cancellation of the operation; or 4) intraoperative change of the initial surgical plan (e.g., intraoperative decision for valve replacement).

Data collection and definitions

Demographic data (sex, age and body mass index (BMI)), data on comorbidities and intraoperative parameters (heart rate (HR), mean arterial pressure (MAP), respiratory rate, and mixed venous oxygen saturation (SvO₂)) were collected retrospectively. The SvO₂ of the patients was continuously measured and recorded using a Swan–Ganz catheter (Edward Company, Irvine, USA) placed through the internal jugular vein.

The endpoint was the occurrence of serious complications after OPCAB, defined as postoperative unplanned intra-aortic balloon pump (IABP) assistance, secondary surgery (e.g., thoracotomy and repeat revascularization), intraoperative emergency conversion to on-pump CABG, and death. Revascularization was defined as revascularization for acute graft failure during the same hospital stay and emergency revascularization for bleeding (i.e., hemorrhagic shock caused by bleeding from the anastomotic site of the transplanted blood vessel during postoperative hospitalization, with repeat revascularization after 2 emergency operations). A patient in whom any of the above

events occurred after OPCAB and before discharge from the hospital was considered to have met the endpoint.

Feature selection and model evaluation

The core principle of model feature screening was based on the feature importance of the machine learning model, combined with Pearson's correlation analysis and statistical analysis of the difference. The specific implementation was as follows:

1. Four algorithms with characteristically important parameters were selected: logical regression analysis (LRA), support vector machine (SVM), random forest (RF), and extreme gradient boosting (XGBoost).
2. The feature importance of the standard features was calculated and ranked based on the above 4 models.
3. The top 10 features of each model were selected as the feature groups (a total of 4 feature groups).
4. Pearson's correlation analysis was carried out on the standard features, and the features with the top 10 correlation coefficients were selected as the feature group.
5. The χ^2 test was performed on the standard features, and the features with statistically significant differences ($p < 0.05$) were selected as the feature group.
6. All 6 feature groups were compared, and the features that appeared 4 times or more were selected as the main features.
7. Finally, Pearson's correlation coefficient was used to distinguish the variables that might affect the endpoint.

After removing meaningless features (Supplementary Table 1), interpolating missing values (Supplementary Table 2), discretizing the numerical variables, and selecting the features, the remaining 6 features were entered into the model. The features were intraoperative ventricular fibrillation (VF), number of saphenous vein grafts (SVG), nerve block (NeB), mixed venous oxygen saturation, skin incision-bypass time (T1), and hypertension (HBP). Results of the tests used for feature selection are shown in Supplementary File 1 and Supplementary Figure 3. Results of verifying the assumptions for the application of the preferred tests are shown in Supplementary Tables 3–10.

"Simpleimputer" in the "sklearn" module was used for the interpolation of the missing data. Mean interpolation was used for numerical variables and mode interpolation for binary and hierarchical variables. Meaningless features were first deleted, and the remaining features were interpolated one by one according to the characteristics and distribution of each feature, rather than based on the 54 features.

All numerical features were discretized by segmentation of continuous numerical data into discrete intervals. The segmentation principle was based on equal frequency, equal distance or optimization methods. Data discretization is also required by many algorithms, since discretization can speed up model training and enhance the robustness of the model by converting the continuous variables

into category variables through discretization. In order to unify the characteristic segmentation of different dimensions, this study used the mean \pm standard deviation ($M \pm SD$) as the segmentation principle. Specifically, all values of the characteristic column were divided into 4 segments according to the nodes of $M-1 \times SD$ or M or $M+1 \times SD$, and each segment was marked as 0, 1, 2, or 3. All features were defined as standard features after meaningless feature removal and numerical feature discretization.

Eight machine learning algorithms were tested in this study: LRA, k-nearest neighbor (KNN), naïve Bayes (NB), SVM, RF, XGBoost, light gradient boosting machine (LightGBM), and categorical features gradient boosting (Catboost).

Statistical analyses

Statistical analysis was performed using the SciPy v. 1.4.1 scientific computing module within the Python 3.8 environment (<https://pypi.org/project/scipy/1.4.1/>). Data were assessed for normality using the Shapiro–Wilk test. Continuous data conforming to a normal distribution were presented as $M \pm SD$ and analyzed using the independent samples t test. Those not conforming to a normal distribution were presented as median (range) and analyzed using the Mann–Whitney U test. Categorical data were presented as n (%) and analyzed using the χ^2 test. Correlation analyses were performed using the Pearson's analysis. Training and validation sets were divided using k-fold cross-validation. The k-fold module in sklearn (https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.KFold.html) was used to randomly divide the database into 5 equal and non-overlapping groups, and the proportion of negative and positive samples in each group was the same. Each time, 4 groups were used as the training set, and 1 group was used as the validation set for model training verification. Precision, recall, F1-score (combining precision and recall into one metric by calculating the harmonic mean between those two³³), and the area under the curve (AUC) were calculated. The above process was performed 5 times to ensure that each group was used as the validation set. Each time, the model was retrained and validated to avoid overfitting, and the average score of the 5 cross-validations was used as the final performance score of the model. The value of $p < 0.05$ was considered statistically significant.

Results

Patient selection

Figure 1 presents the patient selection process. Among the 11,495 patients included in the database, 1238 underwent CABG, and 506 from them were selected based on the eligibility criteria. They were then divided into the training set ($n = 405$) and the validation set ($n = 101$). Table 1 presents the characteristics of the patients. Among the 506 patients

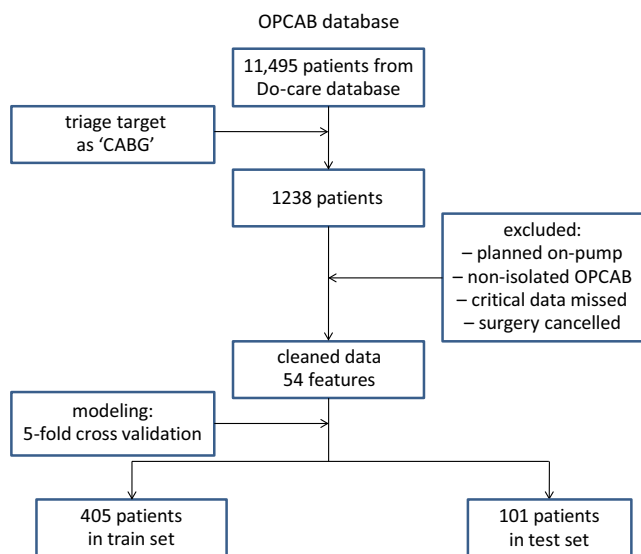


Fig. 1. Screening process of the patients. The patients were selected from the off-pump coronary artery bypass (OPCAB) graft database of the General Hospital of the Northern Theater Command (Shenyang, China)

CABG – coronary artery bypass grafting.

chosen, 27 met the endpoint (positive group), including postoperative emergency IABP assistance ($n = 10$), secondary surgery ($n = 2$), death without other outcomes ($n = 3$), postoperative emergency IABP assistance with secondary surgery ($n = 2$), death after postoperative emergency IABP ($n = 6$), death after secondary surgery ($n = 2$), and death after postoperative emergency IABP assistance and secondary surgery ($n = 2$). The in-hospital mortality rate was 2.6% (13/506). Compared with the controls, the patients who met the endpoint had a lower LVEF ($52 \pm 7\%$ compared to $55 \pm 6\%$, $p = 0.027$), lower fractional shortening ($26 \pm 5\%$ compared to $28 \pm 4\%$, $p = 0.013$), worse New York Heart Association (NYHA) classification score ($p = 0.041$), lower frequency of preoperative diabetes mellitus (DM; 22% compared to 46%, $p = 0.016$), lower intraoperative urine output

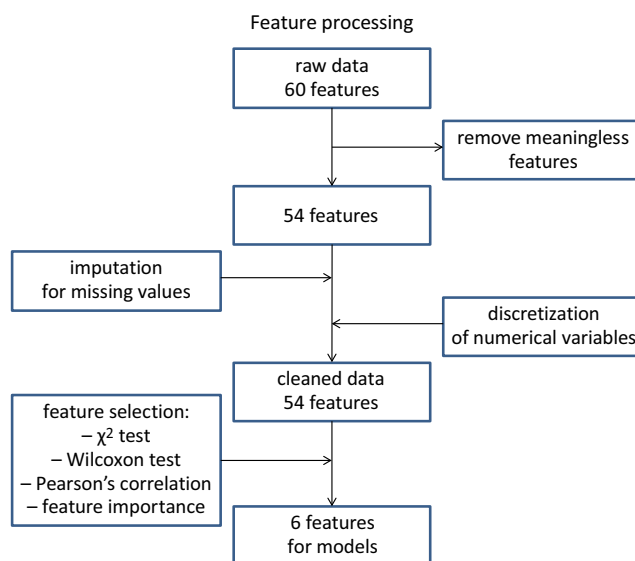


Fig. 2. Feature selection process

(639 ± 2512 mL compared to 771 ± 426 mL, $p = 0.018$), shorter T1 (62.0 ± 17.8 min compared to 69.3 ± 18.7 min, $p = 0.047$), higher SvO₂ values (74% compared to 53%, $p = 0.036$), and a smaller numbers of grafts ($p < 0.001$). Unilateral (left) internal mammary artery to the anterior descending artery anastomosis was performed in all patients. Radial arteries or other arteries were not used as graft vessels.

Feature selection

Figure 2 presents the feature selection process. From the initial 60 features, 6 were removed due to meaninglessness. After missing data imputation and discretization of the continuous variables, 54 clean features were tested, and 6 were retained (intraoperative VF, number of SVG, NeB, SvO₂, T1, and HBP). The results of the correlation analysis of the 6 features are presented in Fig. 3.

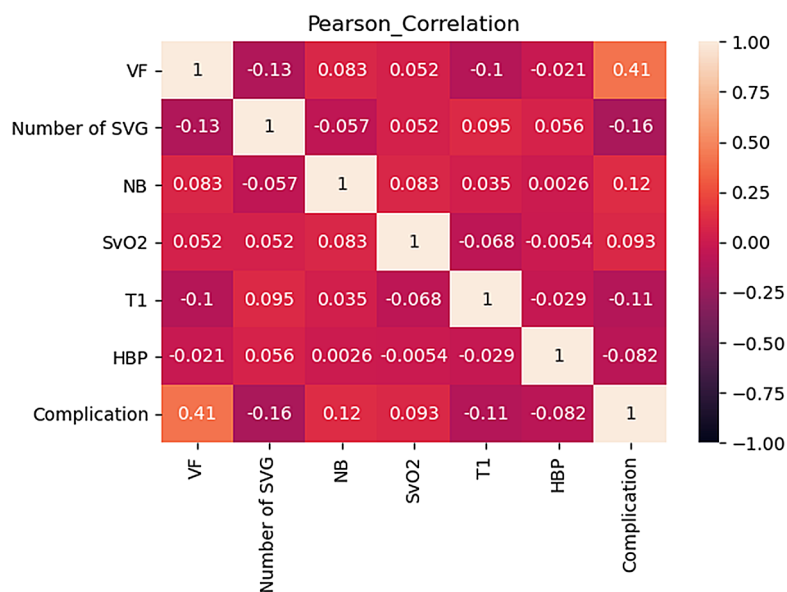


Fig. 3. Correlation analysis among the characteristic variables. Pearson's correlation coefficient was used to distinguish the variables that might affect the endpoint. Finally, these 6 variables were selected to be included in the final model after reordering

VF – ventricular fibrillation; NeB – nerve block; SvO₂ – mixed venous oxygen saturation; T1 – skin incision-bypass time; HBP – high blood pressure; SVG – saphenous vein grafts.

Table 1. Characteristics of the patients with and without serious complications after off-pump coronary artery bypass (OPCAB)

Characteristic	Patients with complications (n = 27)	Patients without complications (n = 479)	Test	Test result	p-value
Age [years], M ±SD	63.3 ±8.6	64.1 ±7.3	Student's t	0.534	0.594
Males, n (%)	22 (81)	362 (76)	χ ²	0.440	0.507
BMI [kg/m ²], M ±SD	23.9 ±2.6	24.7 ±3.0	Student's t	1.295	0.196
LVD [mm], M ±SD	49.6 ±11.6	46.0 ±4.4	Student's t	−1.553	0.132
LVEF (%), M ±SD	52 ±7	55 ±6	Student's t	2.337	0.027
FS (%), M ±SD	26 ±5	28 ±4	Student's t	2.658	0.013
Emergency surgery, n (%)	0 (0)	22 (5)	Fisher's	0.427	0.513
NYHA			Wilcoxon rank sum	5961.0	0.041
I	0 (0)	5 (1)			
II	24 (89)	454 (95)			
III	3 (11)	18 (4)			
IV	0 (0)	2 (0)			
Preoperative comorbidities/risk factors, n (%)					
HBP	14 (52)	330 (69)	χ ²	3.410	0.065
DM	6 (22)	220 (46)	χ ²	5.812	0.016
CVD	4 (15)	98 (20)	χ ²	0.216	0.642
AF	0 (0)	7 (1)	Fisher's	0.046	0.830
AMI	5 (19)	81 (17)	χ ²	0.047	0.829
PCI	6 (22)	80 (17)	χ ²	0.552	0.457
LM	9 (33)	114 (24)	χ ²	1.263	0.261
RCA	24 (89)	435 (91)	χ ²	0.000	0.996
Intraoperative parameters					
HR	62 ±14	61 ±13	Student's t	−0.473	0.636
MAP [mm Hg], M ±SD	76 ±12	75 ±12	Student's t	−0.571	0.568
mPAP [mm Hg]	20 ±5	19 ±5	Student's t	−1.460	0.160
TP [°C], M ±SD	36.2 ±0.4	36.2 ±0.6	Student's t	−0.079	0.937
CI [L/min × m ²], M ±SD	1.90 ±0.40	1.95 ±0.51	Student's t	1.495	0.606
ACCT [min], M ±SD	13.9 ±5.8	15.9 ±6.3	Student's t	1.649	0.100
Intraoperative transfusion [mL], M ±SD	1791 ±798	1652 ±671	Student's t	−1.037	0.300
Urine [mL], M ±SD	639 ±2512	771 ±426	Student's t	2.473	0.018
T1 [min]	62.0 ±17.8	69.3 ±18.7	Student's t	1.749	0.047
SvO ₂ < 75%, n (%)	20 (74)	256 (53)	χ ²	−0.079	0.036
SVG, n (%)			Wilcoxon rank sum	4301.0	<0.001
1	1 (4)	1 (0)			
2	11 (41)	76 (16)			
3	13 (48)	309 (64)			
4	2 (7)	90 (19)			
5	0 (0)	3 (1)			

BMI – body mass index; LVD – left ventricular diameter; LVEF – left ventricular ejection fraction; FS – fractional shortening; NYHA – New York Heart Association; HBP – high blood pressure; DM – diabetes mellitus; CVD – cardiovascular disease; AF – atrial fibrillation; AMI – acute myocardial infarction; PCI – percutaneous coronary intervention; LM – left main artery; RCA – right coronary artery; HR – heart rate; MAP – mean arterial pressure; mPAP – mean pulmonary arterial pressure; TP – temperature; CI – cardiac index; ACCT – aortic cross-clamp time; SvO₂ – mixed venous oxygen saturation; T1 – skin incision-bypass time; SVG – saphenous vein grafts; M \pm SD – mean \pm standard deviation.

Algorithms

Prediction values for the 8 machine learning models are presented in Table 2. The highest AUC was achieved with

the XGBoost model (AUC = 0.94) and the lowest AUC with the SVM model (AUC = 0.75). The highest and lowest accuracy were observed with the XGBoost and NB models, respectively, while the highest and lowest precision were

Table 2. Cross-validation comparison between models

Model	Accuracy (95% CI)	Precision (95% CI)	Recall (95% CI)	F1 (95% CI)	AUC (95% CI)
LRA	0.81 (0.76–0.85)	0.89 (0.86–0.93)	0.70 (0.64–0.65)	0.76 (0.71–0.81)	0.80 (0.72–0.89)
KNN	0.81 (0.76–0.85)	0.86 (0.81–0.90)	0.74 (0.69–0.80)	0.79 (0.75–0.84)	0.80 (0.71–0.89)
NB	0.64 (0.58–0.70)	0.68 (0.62–0.74)	0.63 (0.57–0.68)	0.59 (0.53–0.65)	0.81 (0.72–0.89)
SVM	0.83 (0.79–0.88)	0.94 (0.91–0.97)	0.71 (0.65–0.76)	0.80 (0.75–0.85)	0.75 (0.65–0.85)
RF	0.84 (0.79–0.88)	0.92 (0.89–0.95)	0.75 (0.69–0.80)	0.82 (0.78–0.87)	0.92 (0.86–0.98)
XGBoost	0.84 (0.80–0.89)	0.92 (0.89–0.95)	0.75 (0.70–0.80)	0.83 (0.78–0.87)	0.94 (0.89–0.99)
LightGBM	0.82 (0.78–0.87)	0.87 (0.83–0.91)	0.77 (0.72–0.82)	0.81 (0.76–0.86)	0.92 (0.86–0.98)
CatBoost	0.84 (0.80–0.88)	0.92 (0.89–0.96)	0.75 (0.70–0.80)	0.83 (0.78–0.87)	0.89 (0.82–0.86)

AUC – area under the curve; 95% CI – 95% confidence interval; LRA – logistic regression analysis; KNN – k-nearest neighbor; NB – naïve Bayes; SVM – support vector machine; RF – random forest; XGBoost – extreme gradient boosting; LightGBM – light gradient boosting machine; CatBoost – categorical features gradient boosting. The F1-score combines precision and recall into one metric by calculating the harmonic mean between those two.³⁵

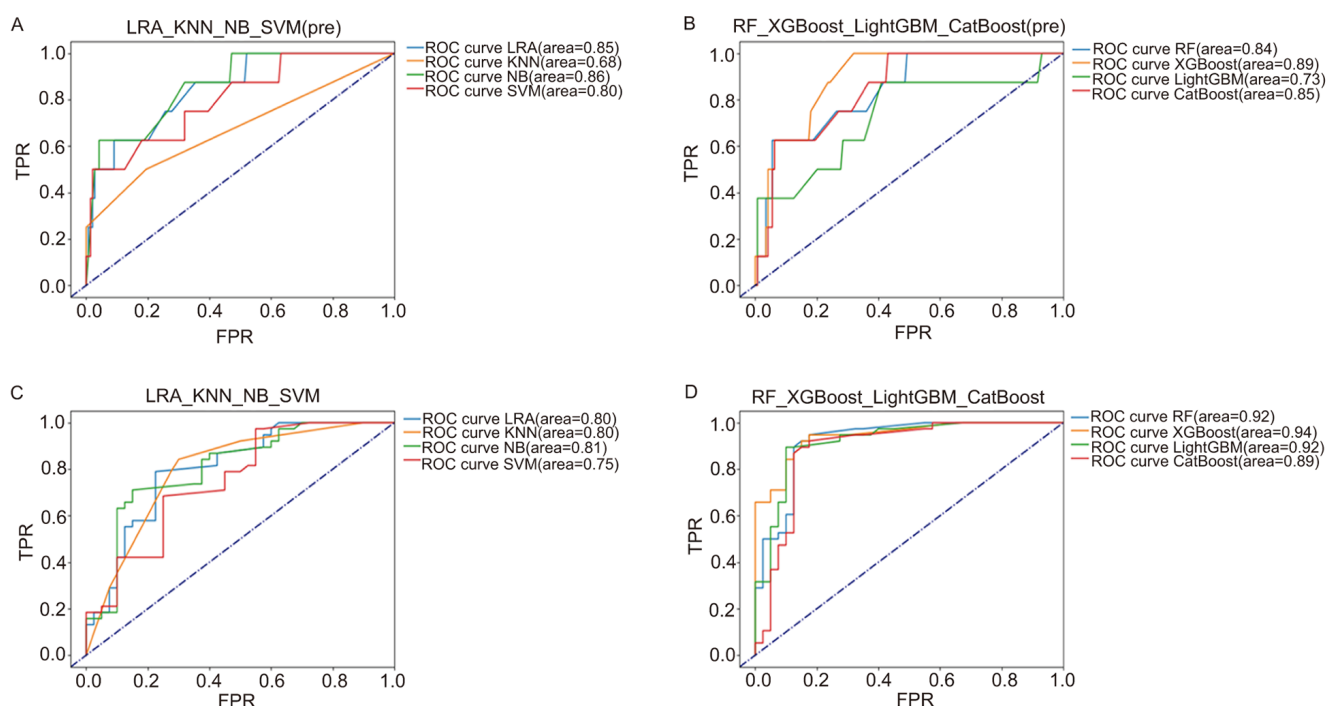


Fig. 4. Receiver operating characteristic (ROC) curve of 2 types of machine learning models before and after data processing algorithm. A. LRA_KNN_NB_SVM before data processing; B. RF_XGBoost_LightGBM_CatBoost before data processing; C. LRA_KNN_NB_SVM after data processing; D. RF_XGBoost_LightGBM_CatBoost after data processing. The XGBoost was confirmed as the final model for this study

TPR – true positive rates; FPR – false positive rates; LRA – logistic regression analysis; KNN – k-nearest neighbor; NB – naïve Bayes; SVM – support vector machine; RF – random forest; XGBoost – extreme gradient boosting; LightGBM – light gradient boosting machine; CatBoost – categorical features gradient boosting.

achieved using the SVM and NB models, respectively. Based on receiver operating characteristics (ROC) curve analysis, the XGBoost model was selected as the final model for the study (Fig. 4). Figure 5 shows the importance of the different variables when analyzed by the different models. Table 3 displays all of the variables evaluated in this study.

Discussion

Results suggest that it is possible to use machine learning algorithms to predict the risk of complications after OPCAB. The highest predictive value was achieved using

the XGBoost model, based on VE, SVG, NeB, SvO₂, T1, and HBP, as revealed by the AUC, which can be used as the main metric to determine the optimal classifier.³⁴

Previous studies used machine learning to predict the mortality and morbidity of cardiac surgery. In a study by Kartal, mortality risk was predicted using the EuroSCORE and the C4.5 algorithm: both the EuroSCORE and the C4.5 algorithm included age, serum creatinine, LVEF, and mean pulmonary hypertension (mPAP).²⁸ They used their algorithm to develop a web application for risk prediction after cardiac surgery. Castela Forte et al. used machine learning to evaluate 88 perioperative variables in order to predict 5-year mortality after cardiac surgery;

Table 3. Features taken into account in this study

No.	Name	Abbreviation	No.	Name	Abbreviation
1.	gender	/	29.*	baseline mean arterial pressure	MAP0
2.*	age	/	30.	skin incision-bypass mean arterial pressure variables	MAP1
3.*	body mass index	BMI	31.	bypass-end mean arterial pressure variables	MAP2
4.	high blood pressure	HBP	32.*	baseline mean pulmonary arterial pressure	mPAP0
5.	diabetes mellitus	DM	33.	skin incision-bypass mean pulmonary arterial pressure variables	mPAP1
6.	cerebrovascular disease	CVD	34.	bypass-end mean pulmonary arterial pressure variables	mPAP2
7.	chronic renal failure	CRF	35.*	temperature	TP
8.	bronchial asthma	BA	36.	mixed venous oxygen saturation	SvO ₂
9.	hyperlipidemia	HL	37.*	cardiac index	CI
10.	chronic obstructive pulmonary disease	COPD	38.	mean pulmonary capillary wedge pressure	mPCWP
11.	atrial fibrillation	AF	39.	American Society of Anesthesiologists score	ASA
12.	acute myocardial infarction	AMI	40.	morphine	/
13.	percutaneous coronary intervention	PCI	41.	etomidate	Etomi
14.	New York Heart Association classification score	NYHA	42.	nerve block	NeB
15.	left main	LM	43.	skin-bypass single push vasoactive agents	VD1
16.	root cause analysis	RCA	44.	bypass-end single push vasoactive agents	VD2
17.*	left ventricle diameter	LVD	45.	internal mammary artery	IMA
18.*	left ventricular ejection fraction	LVEF	46.	radial artery	RA
19.*	fraction shortening	FS	47.	saphenous vein	SV
20.	thickened ventricular septum	TVS	48.*	aortic cross-clamp time	ACCT
21.	valvular heart disease	VHD	49.*	fluid intake	FI
22.	complete left bundle branch block	CLBBB	50.*	urine	/
23.	emergency surgery	/	51.	number of saphenous vein grafts	SVG
24.*	skin incision-bypass time	T1	52.	potassium supplement (intraoperative)	PS
25.*	operation time	T2	53.	calcium supplement (intraoperative)	CS
26.*	baseline heart rate	HR0	54.	on-pump coronary artery bypass	ONCAB
27.	skin incision-bypass heart rate variability	HR1	55.	ventricular fibrillation	VF
28.	bypass-end heart rate variability	HR2	56.	preventive intra-aortic balloon pump	IABP_pre

* continuous variable.

they observed that postoperative urea concentration, age and creatinine concentration, achieved the best predictive values across different cardiac surgery types.²⁹ Kim et al. examined deep neural network, GBM and a generalized linear model to predict major adverse cardiovascular events 1, 6 and 12 months after cardiac surgery, and achieved accuracies >95%.³⁰ Zhong et al. used deep learning to predict the risk of septic shock, thrombocytopenia and liver dysfunction after open-heart surgery.³¹ They examined the performance of XGBoost, RF, KNN and logistic regression, and showed that the XGBoost model achieved the best predictive value for complications. Alshakhs et al. used machine learning to determine the length of hospital stay after CABG, which might be considered a surrogate for the occurrence of postoperative complications.³² They also showed that an RF model including age, height, EuroSCORE II, and the use of IABP achieved the best predictive value.

In the present study, the Pearson's correlation analysis was used to consider the importance of extracting features from different directions (machine learning direction and statistical direction) to make the screened features more convincing. The data indicated that VF, SVG, NeB, SvO₂, T1, and HBP in the XGBoost model achieved the best predictive value. The XGBoost is an advanced complex implementation of gradient boosting algorithms.^{35,36} It can handle both regularization and over/underfitting issues.^{35,36} The parameters selected by the user (i.e., the hyperparameters) usually have a strong effect on the performance of a machine learning algorithm.^{37,38} Still, XGBoost can adapt to the selected hyperparameters to achieve the best fitting,^{35,36,39} which explains its good performance in the present study.

Comparisons among studies are difficult. Indeed, various studies have examined different machine learning

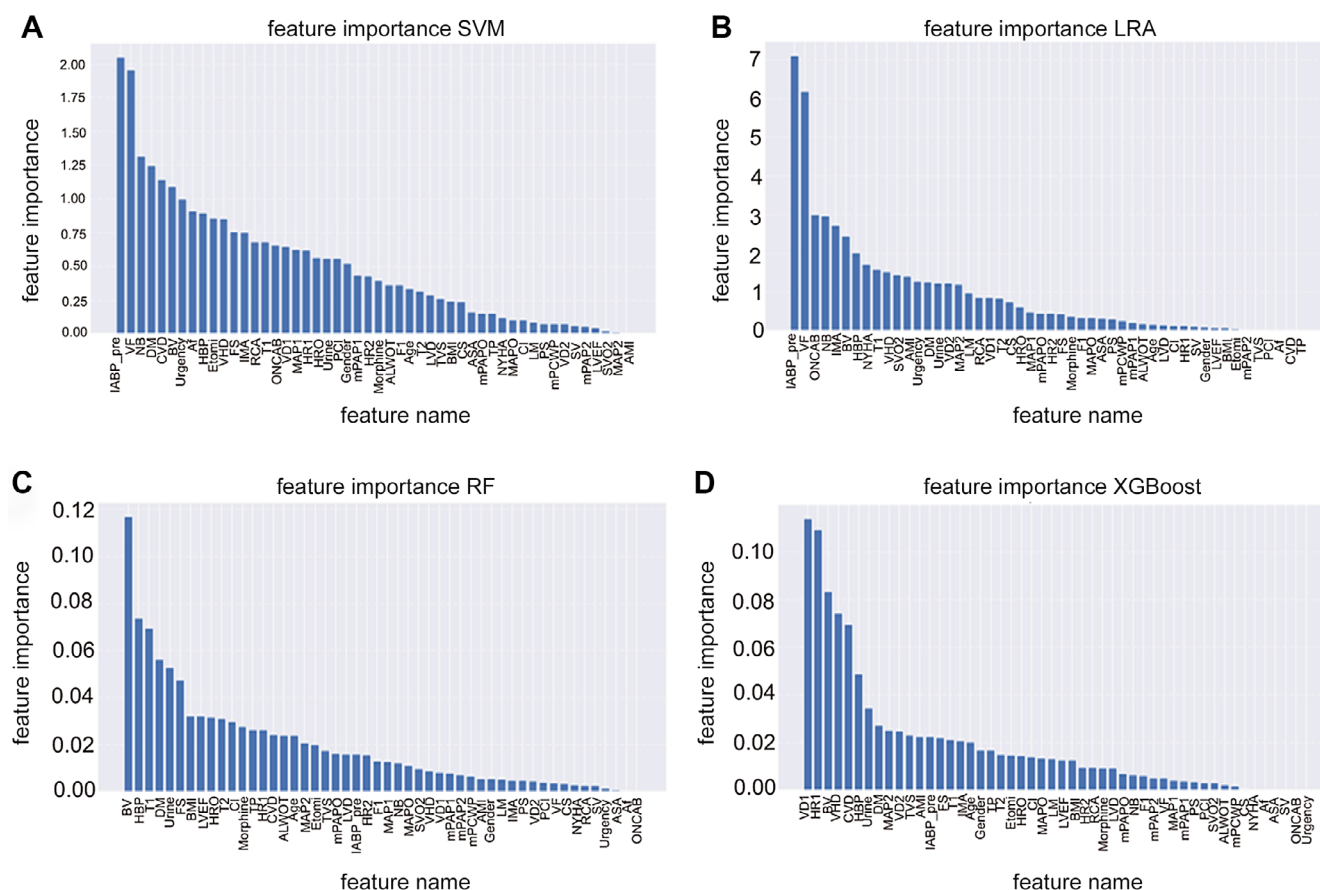


Fig. 5. The feature importance of different algorithms. A. Support vector machine (SVM); B. Logistic regression analysis (LRA); C. Random forest (RF); D. Extreme gradient boosting (XGBoost)

models based on a wide variety of different variables. In addition, the endpoints and the definitions of complications vary, and the study populations have included various types of surgery. In the present study, only patients who underwent OPCAB were included, and the endpoint was the occurrence of IABP assistance, secondary surgery and death. Nonetheless, various studies have shown that the XGBoost model achieved good predictive value. Indeed, similar to above, Zhong et al. used the XGBoost model for predicting complications after open-heart surgery.³¹ Kilic et al. used the XGBoost model to predict the occurrence of operative mortality (AUC = 0.771), renal failure (AUC = 0.776), prolonged ventilation (AUC = 0.739), reoperation (AUC = 0.637), stroke (AUC = 0.684), and deep wound infection (AUC = 0.599) after aortic valve replacement.⁴⁰ Additionally, Lee et al. showed that the XGBoost model had the highest predictive power for AKI after cardiac surgery.⁴¹

Apart from comparison with other deep learning models, the model established here should be compared with well-known and recognized models. Indeed, the EuroSCORE II and the original EuroSCORE have been used for decades to predict mortality risk after cardiac surgery and help improve patient outcomes.^{42,43} The STS score can also be used to determine the risk of CABG.²⁰ However, both scores are not specific to OPCAB. Furthermore,

the data used in the present study were taken directly from the anesthesia monitor system, and some components were not included in the EuroSCORE II and STS scores. Future studies should be set up to allow such direct comparisons using the same set of patients.

Patients who met the study endpoint had a low frequency of DM and high SvO₂. Diabetes mellitus is associated with poor outcomes after CABG or cardiac surgery.^{44–47} On the other hand, poor outcomes after CABG have been associated with either high SvO₂⁴⁸ or low SvO₂.⁴⁹ Considering the small number of patients who met the endpoint in the present study, no conclusion can be drawn on these points.

Limitations

This study has a number of limitations. The data were unbalanced, with the proportion of patients who met the complication endpoint being small. Although the category imbalance was corrected at the data and algorithm levels, it inevitably affected the fitting degree of the model. Follow-up studies are required to optimize the algorithm based on category imbalance characteristics, in order to reduce the impact of category imbalance on the model performance. Although the predictive factors selected in this study related to the endpoint as much as possible, some

predictive factors that had not been discovered or confirmed might have been omitted. In the future, more predictive factors could be added through an in-depth study of OPCAB-related risk factors to improve the performance of the model.

The sample size of this study was small, and it was a single-center retrospective study. The data were from a single center or a single physician team, which limited generalizability and probably introduced some bias caused by varying experience of the surgeons and anesthesiologists. Future studies should be extended to multiple centers. As a retrospective study, this investigation collected the data of all patients who met the criteria in our center during the study period. Relevant data from this period are relatively complete, and data quality cannot be guaranteed in earlier cases. After June 2020, the number of operations decreased due to the coronavirus pandemic, which might have led to bias. An independent validation dataset was also lacking. Therefore, the final model might have poor generalizability. Continuous iterations of the model, through large multicenter samples and prospective validation studies, should increase the generalizability of the model. Since this study only predicted specific, not all complications of OPCAB surgery, its purpose was not to compare the performance of the final model with the EuroSCORE. Data were insufficient to allow separate analyses of patients undergoing total artery bypass grafting.

Conclusions

This study verified the effectiveness of different machine learning models and provided suggestions for the best mathematical model for predicting the risk of complications after OPCAB. This knowledge could be used to continuously optimize the model and introduce it into the clinical medical electronic system, which would allow clinicians to use optimizing treatment strategies in real-time.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Supplementary files

The Supplementary Files are available at <https://doi.org/10.5281/zenodo.7063461>. The package contains the following files:

Supplementary Fig. 1. Results of the Pearson's correlation analysis.

Supplementary File 2. Features selection results of Wilcoxon and χ^2 tests).

Supplementary Table 1. The list of meaningless features.

Supplementary Table 2. The list of missing data.

Supplementary Table 3. Verification results of LRA.

Supplementary Table 4. Verification results of KNN.

Supplementary Table 5. Verification results of naive Bayes (NB).

Supplementary Table 6. Verification results of SVM.

Supplementary Table 7. Verification results of RF.


Supplementary Table 8. Verification results of XGBoost.


Supplementary Table 9. Verification results of LightGBM.


Supplementary Table 10. Verification results of CatBoost.

ORCID iDs

Yang Zhang  <https://orcid.org/0000-0002-3246-1552>

Lin Li  <https://orcid.org/0000-0001-7262-5788>

Ye Li  <https://orcid.org/0000-0002-0154-2464>

Zhihe Zeng  <https://orcid.org/0000-0003-2335-2300>

References

- Diodato M, Chedrawy EG. Coronary artery bypass graft surgery: The past, present and future of myocardial revascularisation. *Surg Res Pract*. 2014;2014:726158. doi:10.1155/2014/726158
- O'Gara PT, Kushner FG, Ascheim DD, et al. 2013 ACCF/AHA Guideline for the Management of ST-Elevation Myocardial Infarction: A report of the American College of Cardiology Foundation/American Heart Association Task Force on practice guidelines. *Circulation*. 2013;127(4):e362–e425. doi:10.1161/CIR.0b013e3182742cf6
- Ibanez B, James S, Agewall S, et al. 2017 ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation. *Eur Heart J*. 2018;39(2):119–177. doi:10.1093/eurheartj/ehx393
- Writing Committee Members, Hillis LD, Smith PK, et al. 2011 ACCF/AHA Guideline for Coronary Artery Bypass Graft Surgery: A report of the American College of Cardiology Foundation/American Heart Association Task Force on practice guidelines. *Circulation*. 2011;124(23):2610–2642. doi:10.1161/CIR.0b013e31823c074e
- Møller CH, Penninga L, Wetterslev J, Steinbrüchel DA, Gluud C. Off-pump versus on-pump coronary artery bypass grafting for ischaemic heart disease. *Cochrane Database Syst Rev*. 2012;14(3):CD007224. doi:10.1002/14651858.CD007224.pub2
- Kuss O, von Salviati B, Börgermann J. Off-pump versus on-pump coronary artery bypass grafting: A systematic review and meta-analysis of propensity score analyses. *J Thorac Cardiovasc Surg*. 2010;140(4):829.e13–835.e13. doi:10.1016/j.jtcvs.2009.12.022
- Matkovic M, Tutus V, Bilbija I, et al. Long term outcomes of the off-pump and on-pump coronary artery bypass grafting in a high-volume center. *Sci Rep*. 2019;9(1):8567. doi:10.1038/s41598-019-45093-3
- Diegeler A, Börgermann J, Kappert U, et al. Five-year outcome after off-pump or on-pump coronary artery bypass grafting in elderly patients. *Circulation*. 2019;139(16):1865–1871. doi:10.1161/CIRCULATIONAHA.118.035857
- Sousa-Uva M, Neumann FJ, Ahlsson A, et al. 2018 ESC/EACTS Guidelines on Myocardial Revascularization. *Eur J Cardiothorac Surg*. 2019;55(1):4–90. doi:10.1093/ejcts/ezy289
- Chikwe J, Lee T, Itagaki S, Adams DH, Egorova NN. Long-term outcomes after off-pump versus on-pump coronary artery bypass grafting by experienced surgeons. *J Am Coll Cardiol*. 2018;72(13):1478–1486. doi:10.1016/j.jacc.2018.07.029
- Daviewala PM. Current outcomes of off-pump coronary artery bypass grafting: Evidence from real world practice. *J Thorac Dis*. 2016;8(S10):S772–S786. doi:10.21037/jtd.2016.10.102
- Wunsch H, Linde-Zwirble WT, C. Angus D. Methods to adjust for bias and confounding in critical care health services research involving observational data. *J Crit Care*. 2006;21(1):1–7. doi:10.1016/j.jcrc.2006.01.004
- Mack MJ, Pfister A, Bachand D, et al. Comparison of coronary bypass surgery with and without cardiopulmonary bypass in patients with multivessel disease. *J Thorac Cardiovasc Surg*. 2004;127(1):167–173. doi:10.1016/j.jtcvs.2003.08.032

14. Shroyer AL, Grover FL, Hattler B, et al. On-pump versus off-pump coronary-artery bypass surgery. *N Engl J Med*. 2009;361(19):1827–1837. doi:10.1056/NEJMoa0902905
15. Shroyer AL, Hattler B, Wagner TH, et al. Five-year outcomes after on-pump and off-pump coronary-artery bypass. *N Engl J Med*. 2017;377(7):623–632. doi:10.1056/NEJMoa1614341
16. Afilalo J, Rasti M, Ohayon SM, Shimony A, Eisenberg MJ. Off-pump vs. on-pump coronary artery bypass surgery: An updated meta-analysis and meta-regression of randomized trials. *Eur Heart J*. 2012;33(10):1257–1267. doi:10.1093/eurheartj/ehs307
17. Caputti GM, Palma JH, Gaia DF, Buffolo E. Off-pump coronary artery bypass surgery in selected patients is superior to the conventional approach for patients with severely depressed left ventricular function. *Clinics (Sao Paulo)*. 2011;66(12):2049–2053. doi:10.1590/S1807-59322011001200009
18. Maltais S, Ladouceur M, Cartier R. The influence of a low ejection fraction on long-term survival in systematic off-pump coronary artery bypass surgery. *Eur J Cardiothorac Surg*. 2011;39(5):e122–e127. doi:10.1016/j.ejcts.2010.12.022
19. Xia L, Ji Q, Song K, et al. Early clinical outcomes of on-pump beating-heart versus off-pump technique for surgical revascularization in patients with severe left ventricular dysfunction: The experience of a single center. *J Cardiothorac Surg*. 2017;12(1):11. doi:10.1186/s13019-017-0572-x
20. Ragosta M. Left main coronary artery disease: Importance, diagnosis, assessment, and management. *Curr Probl Cardiol*. 2015;40(3):93–126. doi:10.1016/j.cpcardiol.2014.11.003
21. Grant SW, Hickey GL, Dimarakis I, et al. How does EuroSCORE II perform in UK cardiac surgery: An analysis of 23 740 patients from the Society for Cardiothoracic Surgery in Great Britain and Ireland National Database. *Heart*. 2012;98(21):1568–1572. doi:10.1136/heartjnl-2012-302483
22. Parolari A, Pesce LL, Trezzi M, et al. Performance of EuroSCORE in CABG and off-pump coronary artery bypass grafting: Single institution experience and meta-analysis. *Eur Heart J*. 2008;30(3):297–304. doi:10.1093/eurheartj/ehn581
23. Hannan EL, Wu C, Bennett EV, et al. Risk stratification of in-hospital mortality for coronary artery bypass graft surgery. *J Am Coll Cardiol*. 2006;47(3):661–668. doi:10.1016/j.jacc.2005.10.057
24. Singh M, Gersh BJ, Li S, et al. Mayo Clinic Risk Score for percutaneous coronary intervention predicts in-hospital mortality in patients undergoing coronary artery bypass graft surgery. *Circulation*. 2008;117(3):356–362. doi:10.1161/CIRCULATIONAHA.107.711523
25. Birim Ö, van Gameren M, Bogers AJJC, Serruys PW, Mohr FW, Kappelein AP. Complexity of coronary vasculature predicts outcome of surgery for left main disease. *Ann Thorac Surg*. 2009;87(4):1097–1105. doi:10.1016/j.athoracsur.2008.11.079
26. Weller GB, Lovely J, Larson DW, Earnshaw BA, Huebner M. Leveraging electronic health records for predictive modeling of post-surgical complications. *Stat Methods Med Res*. 2018;27(11):3271–3285. doi:10.1177/0962280217696115
27. Soguero-Ruiz C, Hindberg K, Mora-Jiménez I, et al. Predicting colorectal surgical complications using heterogeneous clinical data and kernel methods. *J Biomed Inform*. 2016;61:87–96. doi:10.1016/j.jbi.2016.03.008
28. Kartal E. Machine learning techniques in cardiac risk assessment. *Turk Gogus Kalp Dama*. 2018;26(3):394–401. doi:10.5606/tgkdc.dergisi.2018.15559
29. Castela Forte J, Munroop HE, de Geus F, et al. Ensemble machine learning prediction and variable importance analysis of 5-year mortality after cardiac valve and CABG operations. *Sci Rep*. 2021;11(1):3467. doi:10.1038/s41598-021-82403-0
30. Kim YJ, Saqlan M, Lee JY. Deep learning-based prediction model of occurrences of major adverse cardiac events during 1-year follow-up after hospital discharge in patients with AMI using knowledge mining. *Pers Ubiquit Comput*. 2022;26(2):259–267. doi:10.1007/s00779-019-01248-7
31. Zhong Z, Yuan X, Liu S, Yang Y, Liu F. Machine learning prediction models for prognosis of critically ill patients after open-heart surgery. *Sci Rep*. 2021;11(1):3384. doi:10.1038/s41598-021-83020-7
32. Alshakhs F, Alharthi H, Aslam N, Khan IU, Elasheri M. Predicting post-operative length of stay for isolated coronary artery bypass graft patients using machine learning. *Int J Gen Med*. 2020;13:751–762. doi:10.2147/IJGM.S250334
33. Orozco-Arias S, Piña JS, Tabares-Soto R, Castillo-Ossa LF, Guyot R, Isaza G. Measuring performance metrics of machine learning algorithms for detecting and classifying transposable elements. *Processes*. 2020;8(6):638. doi:10.3390/pr8060638
34. Provost F, Fawcett T. Robust classification for imprecise environments. *arXiv*. 2000;2000:arXiv:cs/0009007. doi:10.48550/ARXIV.CS/0009007
35. Xia Y, Liu C, Li Y, Liu N. A boosted decision tree approach using Bayesian hyper-parameter optimization for credit scoring. *Expert Syst Appl*. 2017;78:225–241. doi:10.1016/j.eswa.2017.02.017
36. Zięba M, Tomczak SK, Tomczak JM. Ensemble boosted trees with synthetic features generation in application to bankruptcy prediction. *Expert Syst Appl*. 2016;58:93–101. doi:10.1016/j.eswa.2016.04.001
37. Brown SD, de Juan A. ICRM-2011 international chemometrics research meeting. *Chemom Intell Lab Syst*. 2012;111(1):66. doi:10.1016/j.chemolab.2011.12.002
38. Mantovani RG, Rossi ALD, Vanschoren J, Bischl B, de Carvalho ACPLF. Effectiveness of random search in SVM hyper-parameter tuning. In: *2015 International Joint Conference on Neural Networks (IJCNN)*. Killarney, Ireland: IEEE; 2015:1–8. doi:10.1109/IJCNN.2015.7280664
39. Budholiya K, Shrivastava SK, Sharma V. An optimized XGBoost based diagnostic system for effective prediction of heart disease. *J King Saud Univ Comput Inf Sci*. 2022;34(7):4514–4523. doi:10.1016/j.jksuci.2020.10.013
40. Kilic A, Goyal A, Miller JK, Gleason TG, Dubrawski A. Performance of a machine learning algorithm in predicting outcomes of aortic valve replacement. *Ann Thorac Surg*. 2021;111(2):503–510. doi:10.1016/j.athoracsur.2020.05.107
41. Lee HC, Yoon HK, Nam K, et al. Derivation and validation of machine learning approaches to predict acute kidney injury after cardiac surgery. *J Clin Med*. 2018;7(10):322. doi:10.3390/jcm7100322
42. Nashef SAM, Roques F, Sharples LD, et al. EuroSCORE II. *Eur J Cardiothorac Surg*. 2012;41(4):734–745. doi:10.1093/ejcts/ezs043
43. Biancari F, Vasques F, Mikkola R, Martin M, Lahtinen J, Heikkinen J. Validation of EuroSCORE II in patients undergoing coronary artery bypass surgery. *Ann Thorac Surg*. 2012;93(6):1930–1935. doi:10.1016/j.athoracsur.2012.02.064
44. van Straten AHM, Soliman Hamad MA, van Zundert AAJ, et al. Diabetes and survival after coronary artery bypass grafting: Comparison with an age- and sex-matched population. *Eur J Cardiothorac Surg*. 2010;37(5):1068–1074. doi:10.1016/j.ejcts.2009.11.042
45. Kogan A, Ram E, Levin S, et al. Impact of type 2 diabetes mellitus on short- and long-term mortality after coronary artery bypass surgery. *Cardiovasc Diabetol*. 2018;17(1):151. doi:10.1186/s12933-018-0796-7
46. Ram E, Sternik L, Klempfner R, et al. Type 2 diabetes mellitus increases the mortality risk after acute coronary syndrome treated with coronary artery bypass surgery. *Cardiovasc Diabetol*. 2020;19(1):86. doi:10.1186/s12933-020-01069-6
47. Järvinen O, Hokkanen M, Huhtala H. Diabetics have inferior long-term survival and quality of life after CABG. *Int J Angiol*. 2019;28(01):50–56. doi:10.1055/s-0038-1676791
48. Balzer F, Sander M, Simon M, et al. High central venous saturation after cardiac surgery is associated with increased organ failure and long-term mortality: An observational cross-sectional study. *Crit Care*. 2015;19(1):168. doi:10.1186/s13054-015-0889-6
49. Holm J, Håkanson E, Vánky F, Svedjeholm R. Mixed venous oxygen saturation predicts short- and long-term outcome after coronary artery bypass grafting surgery: A retrospective cohort analysis. *Br J Anesth*. 2011;107(3):344–350. doi:10.1093/bja/aer166