Journal of Cardiovascular Translational Research (2025) 18:185–197
https://doi.org/10.1007/s12265-024-10565-z

ORIGINAL ARTICLE

Machine Learning Model for Predicting Risk Factors of Prolonged Length of Hospital Stay in Patients with Aortic Dissection: a Retrospective Clinical Study

Luo Li[1] · Yihuan Chen[1] · Hui Xie[1] · Peng Zheng[2] · Gaohang Mu[1] · Qian Li[1] · Haoyue Huang[1] · Zhenya Shen[1]

Received: 1 February 2024 / Accepted: 4 September 2024 / Published online: 10 October 2024

Abstract

The length of hospital stay (LOS) is crucial for assessing medical service quality. This study aimed to develop machine learning models for predicting risk factors of prolonged LOS in patients with aortic dissection (AD). The data of 516 AD patients were obtained from the hospital's medical system, with 111 patients in the prolonged LOS (> 30 days) group based on three quarters of the LOS in the entire cohort. Given the screened variables and prediction models, the XGBoost model demonstrated superior predictive performance in identifying prolonged LOS, due to the highest area under the receiver operating characteristic curve, sensitivity, and F1-score in both subsets. The SHapley Additive exPlanation analysis indicated that high density lipoprotein cholesterol, alanine transaminase, systolic blood pressure, percentage of lymphocyte, and operation time were the top five risk factors associated with prolonged LOS. These findings have a guiding value for the clinical management of patients with AD.

Keywords Machine learning · Aortic dissection · Length of hospital stay · Risk factors · Prediction

Abbreviations

AD	Aortic dissection
LOS	Length of hospital stay
LICU	Length of ICU stay
AI	Artificial intelligence
EMR	Electronic medical record
KNN	K-nearest neighbors
GNB	Gaussian naive bayes
CNB	Complement naive bayes
MLP	Multi-layer perceptron neural network
SVM	Support vector machine
ROC	Receiver operating characteristic
AUC	Area under the receiver operating characteristic curve
ACC	Accuracy
DCA	Decision curve analysis
SHAP	SHapley Additive exPlanation

Luo Li and Yihuan Chen equally contributed to this work.

Associate Editor Sandra Loerakker oversaw the review of this article.

✉ Haoyue Huang
hhysz@suda.edu.cn

✉ Zhenya Shen
uuzyshen@aliyun.com

[1] Department of Cardiovascular Surgery of the First Affiliated Hospital & Institute for Cardiovascular Science, Soochow University, Suzhou Medical College, Soochow University, 899 Pinghai Road, Jiangsu 215123, Suzhou, China

[2] Department of Cardiology, School of Medicine, Zhongda Hospital, Southeast University, 87 Dingjiaqiao, Jiangsu 210009, Nanjing, China

Introduction

Aortic dissection (AD) typically refers to the occurrence of interlaminar bleeding resulting from tears in the intima and media layers of the vessel wall [1]. The Stanford classification categorizes AD involving the ascending aorta as type A and that affecting only the descending aorta as type B [2]. The prevalence of AD is approximately 0.003%, and mortality rate in untreated AD patients increases by 1–2% per hour after symptom onset [3]. However, when the ascending aorta of patients is affected, mortality rates are further elevated [4]. Surgical intervention plays a crucial role in treating AD patients in clinical practice, and recent advancements in

the techniques have led to impressive clinical outcomes, including reduced complications and mortality rates, as well as shorter hospital stays [5].

Length of hospital stay (LOS) is commonly used as a patient-centered indicator, reflecting the quality of care and prognosis of hospitalized individuals [6]. A recent study has revealed the importance of reducing postoperative LOS in patients with cardiovascular disease [7]. Moreover, shortening LOS can alleviate the burden on hospital management and prevent unnecessary waste of medical resources [8]. Indeed, previous studies have extensively investigated the risk factors for LOS in patients undergoing heart transplantation or coronary artery bypass grafting [9, 10]. However, there is a dearth of research investigating the risk factors of prolonged LOS in AD.

The recent progress in data mining technology has resulted in the widespread adoption of artificial intelligence (AI)-based machine learning algorithms for data analysis and the development of robust predictive models [11]. The automated analysis of large volumes of data enables AI to accurately model intricate relationships within the extensive datasets at an accelerated pace. In the medical field, machine learning algorithm-based models are frequently used to enhance the ability in the diagnosis and prognosis of disease, thereby improving overall healthcare [12]. The application of machine learning models in clinical practice, such as heart failure, is gradually gaining traction to ensure their practicality [13–15]. This study aimed to use clinical data and develop machine learning models to predict prolonged LOS risk in patients with AD during hospitalization.

Methods

Patient Population

The protocol of this study was approved by the Ethical Review Committee of the First Affiliated Hospital of Soochow University (Ethics number: No. 2021–424). The study protocol was in accordance with the Declaration of Helsinki (revised in 2013). Considering the retrospective nature of this study, informed consent from the patients was waived by the Ethical Review Committee of the First Affiliated Hospital of Soochow University.

This study enrolled 586 patients diagnosed with AD at the Department of Cardiovascular Surgery of the First Affiliated Hospital of Soochow University from January 2020 to May 2024. The inclusion criteria for patients [16] were as follows: (1) age over 18 years old; (2) received surgical treatment during hospitalization; (3) pathologic anatomy consistent with AD disease based on imaging examinations or surgical observation; and (4) clinical symptom onset within 14 days prior to surgical treatment. The exclusion criteria for patients included: (1) abandonment of surgery or preoperative dissection rupture; (2) rheumatic and autoimmune disorders; (3) hematologic conditions; (4) malignant tumors; (5) severe coronary and heart disease; (6) history of anticoagulant medication before hospitalization; (6) LOS less than 3 days; and (7) data missing more than 30%. A total of 70 patients were eventually excluded from this study. The AD patients were classified based on the third quartile of hospitalization duration in the entire study cohort (Q3 = 30 days). Among 516 patients diagnosed with AD, a total of 111 individuals experienced a prolonged LOS (exceeding 30 days).

Data Collection

The collected variables are displayed in Table S1. Patients' clinical data were collected using electronic medical record (EMR) system of the hospital. These data mainly included demographics (gender, age, height, weight, body surface area [BSA], body mass index [BMI], systolic blood pressure [SBP], diastolic blood pressure [DBP] upon admission, length of ICU stay [LICU], and LOS). The comorbidities were recorded from EMR system, such as hypertension (HTN), coronary heart disease (CHD), diabetes mellitus (DM), cerebral infarction history (CIH), arrhythmia, hepatic insufficiency (HepIS), and acute renal dysfunction (ARD). Additionally recorded indicators were echocardiography parameters (left ventricular ejection fraction [LVEF] and diameter of aorta [AA]), cardiac valve diseases, operation time (OT), and hospital costs (HOS). Some laboratory parameters, including routine blood tests, ratio of neutrophil to lymphocyte (NLR), ratio of platelet to lymphocyte (PLT/Lym), glucose (GLU), electrolytes, high-sensitivity C-reactive protein (hs-CRP), serum creatinine (CrS), Urea, albumin (ALB), globulin (GLB), ratio of Urea to GLB (Urea/GLB), and bilirubin levels, were documented in this study. These indicators were assessed after a 12-h fasting period upon admission of the patients. Missing data (Table S2 and Table S3) were supplemented using multiple imputation method by the R programming language, specifically utilizing the "mice" package (version 3.12.0). In this process, 10 imputed datasets were generated, and the average data were selected as the final data to be included in the model training and evaluation.

Statistic Analysis

Categorical variables were presented as frequencies and percentages, and the comparisons among groups were conducted using the chi-square test. Continuous variables with a normalized distribution were expressed as mean ± standard deviations (SD), and the comparisons among groups were analyzed using Student's t-test. The non-normally distributed continuous variables were presented as median ($P_{25\%}$,

$P_{75\%}$), and the comparisons between two groups were analyzed using the Mann–Whitney U test. The data analysis was conducted using SPSS 28.0, and all p values were two-sided tests, with $p < 0.05$ considered a significant difference.

Feature Selection

The Boruta algorithm is a feature selection method that utilizes the random forest technique. It is employed to iteratively assess the importance of each variable. By incorporating shadow features, the significance of both the original variable and its corresponding shadow variable are compared during each round of iteration. An original variable is deemed significant if it exhibits a significantly higher level of importance compared to the shadow variable. Conversely, if the original variable demonstrates much greater importance than its shadow counterpart, it is considered insignificant. In this study, the Boruta algorithm was employed to identify the key features associated with prolonged LOS in patients and subsequently these crucial variables were utilized for model development. This process effectively mitigates overfitting and optimizes parameterization.

Machine Learning Models Development

Data collected from a single center were randomly divided into two cohorts (70% of the data set was randomly selected as the training subset, and the remaining 30% was used as the validation subset for internal validation). In this process, the resampling method was used to find more model parameters. In this study, XGBoost, AdaBoost, k-nearest neighbors (KNN), logistic regression, lightGBM, gaussian naive bayes (GNB), multi-layer perceptron neural network (MLP), complement naive bayes (CNB), and support vector machine (SVM) algorithms were used to develop machine learning models to predict prolonged LOS risk in AD patients. The model with the highest prediction performance was used as the final prediction model. The performance of the model was evaluated by the area under the receiver operating characteristic curve (AUC), sensitivity, specificity, accuracy (ACC), and F1-score. Generally, the AUC value is recognized as the most crucial indicator for evaluating the model's performance. Furthermore, when considering the AUCs of the machine learning models, other parameters (sensitivity, specificity, ACC, and F1-score, etc.) are also used as assessing indicators. The significance of AUCs among models were assessed using the Delong test, and $p < 0.05$ was considered statistically significant. Meanwhile, calibration curve and decision curve analysis (DCA) were obtained in the validation set to evaluate the disparity between clinical practice and predictive outcomes. After the index evaluation was completed, the model with the highest AUC, sensitivity, and other indexes in the training and validation subsets were selected as the final prediction model. The flowchart of this study is presented in Fig. 1.

Model Interpretation

The SHapley Additive exPlanation (SHAP) method was employed in this study to characterize the optimal model, visualize the significance of each feature in the model, and exhibit further the influence of individual features on classification outcomes. In this study, SHAP was utilized to demonstrate a reasonably balanced distribution across feature values. These features played a crucial role in ensuring confidence in the risk factors associated with prolonged LOS in AD patients. The SHAP value data for each feature were derived by calculating the sample data of the patients using the entire cohort, thereby determining each individual feature's contribution to the prediction results.

Results

Baseline Characteristics of Patients

A total of 516 AD patients were included in this study, with 111 cases in the prolonged LOS group. The median age of the participants was recorded as 52 years old, with 445 male cases (86.24%). There were significant differences between the two groups in terms of degree of aortic regurgitation (AR, $p = 0.031$), ARD ($p = 0.01$), SBP ($p = 0.016$), LICU ($p < 0.001$), OT ($p < 0.001$), HOS ($p < 0.001$), degree of tricuspid regurgitation (TR, $p < 0.001$), NLR ($p = 0.030$), lymphocyte count (Lym, $p = 0.046$), percent of lymphocyte (LymP, $p = 0.034$), mean corpuscular volume (MCV, $p = 0.021$), mean corpuscular hemoglobin concentration (MCHC, $p = 0.018$), cystatin-c (CysC, $p = 0.012$), serum K^+ ($p = 0.013$), low-density lipoprotein cholesterol (LDL-C, $p = 0.027$), high-density lipoprotein cholesterol (HDL-C, $p < 0.001$), total cholesterol (TC, $p = 0.013$), Urea ($p = 0.010$), ALB ($p = 0.004$), Urea/GLB ($p = 0.017$), total protein (TP, $p = 0.039$), N-Terminal pro-brain natriuretic peptide (NT-proBNP, $p = 0.032$), high-sensitivity troponin T (hs-cTnT, $p = 0.006$), prothrombin time (PT, $p = 0.002$), international normalized ratio (INR, $p = 0.001$), AA ($p < 0.001$), and AD type ($p < 0.001$). The detail content is shown in Table 1.

Variable Selection and Machine Learning Model Development

The Boruta algorithm was employed in this study to select the key variables associated with prolonged LOS in patients with AD. Ultimately, a total of 22 variables were identified and utilized for the development of machine learning

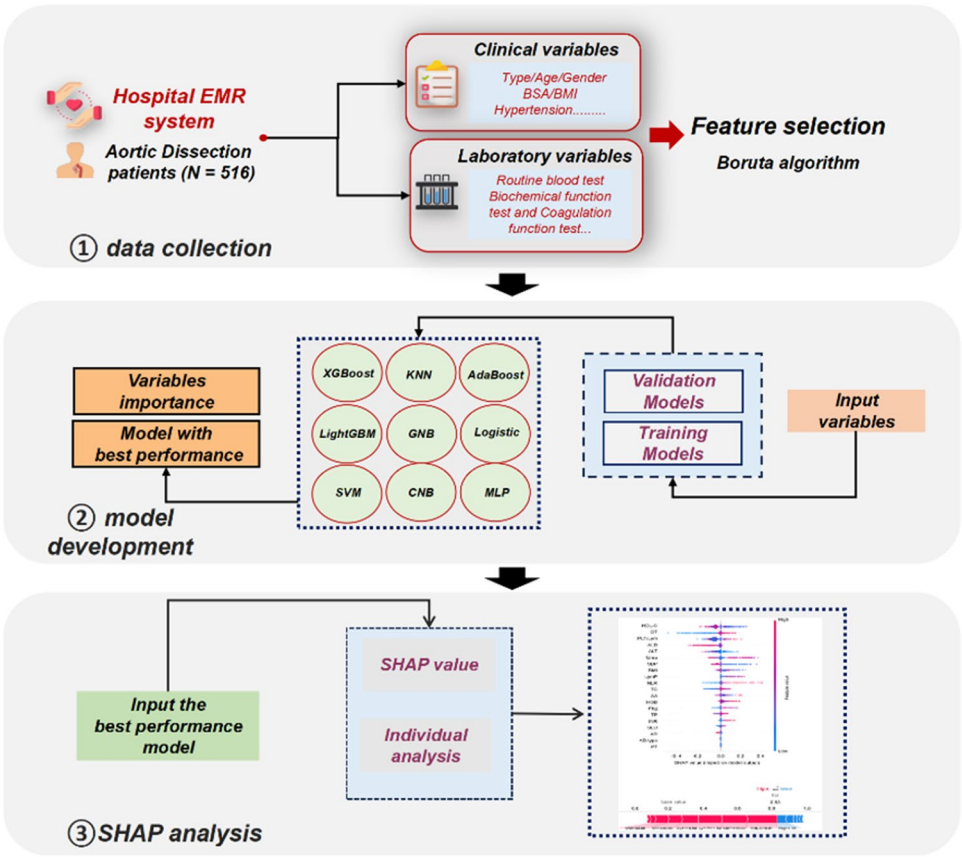

Fig. 1 The flowchart of this study. The flowchart of this study consists of data collection and model development. EMR, electronic medical record; KNN, k-nearest neighbors; GNB, gaussian naive bayes; SVM, support vector machine; CNB, complement naive bayes; MLP, multi-layer perceptron neural network

models (Fig. 2). The selected variables included AD type, SBP, OT, AA, NLR, PLT/Lym, HDL-C, ALB, BMI, DM, AR, TR, Lym, hemoglobin (HGB), TC, GLU, Urea, TP, alanine transaminase (ALT), PT, fibrinogen (Fbg), and INR.

The receiver operating characteristic (ROC) curves of the prediction models were plotted in this study, and the corresponding AUC values were obtained. Among nine developed models, the XGBoost model showed a good performance in predicting prolonged LOS risk in AD patients, with the highest AUC value (0.96 in the training subset and 0.71 in the validation subset) (Fig. 3). Additionally, the sensitivity, specificity, accuracy, and F1-score of the XGBoost model in the training subset were 0.92, 0.90, 0.90 and 0.82, respectively. When the prediction models were used for internal validation in the validation subset, the XGBoost model still showed the best performance (Table 2 and Table 3). This was further confirmed by using Delong test, and the p values between XGBoost and other models in the Delong test were all less than 0.05 (Table S4 and Table S5). The XGBoost model had an excellent calibration curve revealing its superior fit by closely matching predicted probabilities to actual probabilities (Fig. 4). The DCA showed that the XGBoost model outperformed both the all-line and none-line, especially at threshold probabilities of 16% to 83%. The results from the DCA curve indicated a significant net benefit of patient intervention in the prediction of the XGBoost model, thereby demonstrating strong clinical validity within this range (Fig. 5). Therefore, the XGBoost model was considered the final predictive model in this study.

Model Application

The SHAP model was used to evaluate the significance of characteristics in the XGBoost model with optimal performance. HDL-C, SBP, LymP, ALT, and OT were identified as the top five indicators influencing the XGBoost model in SHAP analysis. The SHAP summary plots effectively demonstrated the impact of different variables on distinguishing LOS in patients with AD (Fig. 6). The XGBoost model constructed in this study predicted the probability of prolonged LOS in individuals with AD to be 83.00%. The results exhibited that HDL-C of 0.81 mmol/L, SBP of 116.00 mmHg, LymP of 11.5%, ALT of 16.50 U/L, and OT of 440 min were the top five variables promoting prolonged LOS (Fig. 7).

Discussion

Machine learning is broadly defined as a system's ability to autonomously acquire knowledge by identifying patterns within large data sets [17]. Machine learning-based

Table 1 Baseline characteristics of patients in the two groups

Indicator		Total patients (N = 516)	Normal LOS group (N = 405)	Prolonged LOS group (N = 111)	P-value
Age (year)		52.0 (41.0, 60.0)	52.0 (42.0, 60.0)	50.0 (39.0, 62.0)	0.195
Gender (n, %)	**Male**	455 (86.24)	349 (86.17)	96 (86.49)	0.932
	Female	71 (13.76)	56 (13.83)	15 (13.51)	
AD type (n, %)	**A**	303 (58.72)	212 (52.35)	91 (81.98)	**< 0.001**
	B	213 (41.28)	193 (47.65)	20 (18.02)	
BSA (m^2)		1.92 (1.79, 2.06)	1.92 (1.78, 2.05)	1.92 (1.80, 2.08)	0.488
BMI (kg/m^2)		25.95 (23.44, 28.76)	25.95 (23.44, 28.65)	25.47 (23.66, 29.41)	0.566
SBP (mmHg)		147.00 (130.00, 161.00)	148.00 (132.00, 161.00)	136.00 (122.00, 161.00)	**0.016**
DBP (mmHg)		78.00 (67.00, 90.00)	78.00 (68.00, 91.00)	77.00 (64.00, 88.00)	0.110
HepIS (n, %)	**No**	497 (96.32)	391 (96.54)	106 (95.50)	0.604
	Yes	19 (3.68)	14 (3.46)	5 (4.50)	
CHD (n, %)	**No**	451 (87.40)	353 (87.16)	98 (88.29)	0.751
	Yes	65 (12.60)	52 (12.84)	13 (11.71)	
Arrhythmia (n, %)	**No**	504 (97.67)	397 (98.02)	107 (96.40)	0.313
	Yes	12 (2.33)	8 (1.98)	4 (3.60)	
HTN (n, %)	**No**	167 (32.36)	133 (32.84)	34 (30.63)	0.659
	Yes	349 (67.64)	272 (67.16)	77 (69.37)	
DM (n, %)	**No**	495 (95.93)	392 (96.79)	103 (92.79)	0.059
	Yes	21 (4.07)	13 (3.21)	8 (7.21)	
ARD (n, %)	**No**	476 (92.25)	380 (93.83)	96 (86.49)	**0.010**
	Yes	40 (7.75)	25 (6.17)	15 (13.51)	
CIH (n, %)	**No**	484 (93.80)	378 (93.33)	106 (95.50)	**0.403**
	Yes	32 (6.20)	27 (6.67)	5 (4.50)	
hsCRP (n, %)	**0.00–0.56**	6 (1.16)	6 (1.48)	0 (0.00)	Na
	0.56–3.00	34 (6.59)	27 (6.67)	7 (6.31)	
	3.00–15.36	247 (47.87)	185 (45.68)	62 (55.86)	
	> 15.36	229 (44.38)	187 (46.17)	42 (37.84)	
WBC (10^9/L)		11.46 (8.88, 14.07)	11.30 (8.86, 13.89)	12.17 (9.56, 14.63)	0.112
Lym (10^9/L)		0.89 (0.60, 1.25)	0.91 (0.62, 1.26)	0.85 (0.57, 1.16)	**0.046**
Mon (10^9/L)		0.55 (0.36, 0.80)	0.55 (0.35, 0.77)	0.57 (0.39, 0.86)	0.191
Neu (10^9/L)		10.02 (7.35, 12.57)	9.80 (7.33, 12.27)	10.47 (7.58, 13.36)	0.091
Eos (10^9/L)		0.00 (0.00, 0.03)	0.00 (0.00, 0.03)	0.01 (0.00, 0.03)	0.693
Bas (10^9/L)		0.02 (0.01, 0.03)	0.02 (0.01, 0.03)	0.02 (0.01, 0.03)	0.425
LymP (%)		7.50 (5.00, 12.10)	7.70 (5.20, 12.50)	7.10 (4.20, 11.00)	**0.034**
MonP (%)		5.10 (3.40, 6.90)	5.10 (3.30, 6.90)	5.10 (3.50, 6.90)	0.767
NeuP (%)		86.80 (80.80, 91.20)	86.40 (80.40, 91.10)	87.90 (81.80, 91.70)	0.208
EosP (%)		0.00 (0.00, 0.30)	0.00 (0.00, 0.30)	0.10 (0.00, 0.20)	0.735
BasP (%)		0.20 (0.10, 0.20)	0.20 (0.10, 0.20)	0.20 (0.10, 0.20)	0.456
NLR		11.51 (6.76, 18.42)	11.36 (6.50, 17.66)	12.60 (7.77, 22.00)	**0.030**
RBC (10^{12}/L)		4.42 (4.01, 4.71)	4.43 (4.04, 4.70)	4.38 (3.92, 4.81)	0.635
HGB (g/L)		132.00 (120.00, 141.00)	132.00 (121.00, 141.00)	131.00 (118.00, 141.00)	0.539
HCT (L/L)		0.40 (0.37, 0.43)	0.40 (0.37, 0.43)	0.39 (0.36, 0.43)	0.218
MCV (fl)		90.70 (88.00, 93.20)	90.90 (88.50, 93.30)	89.60 (87.00, 92.50)	**0.021**
MCH (pg)		30.10 (29.20, 31.00)	30.10 (29.20, 31.00)	30.10 (29.10, 30.90)	0.551
MCHC (g/L)		332.00 (326.00, 337.00)	331.00 (326.00, 336.00)	334.00 (328.00, 338.00)	**0.018**
RDW (%)		13.00 (12.60, 13.40)	13.00 (12.60, 13.40)	12.90 (12.60, 13.40)	0.865
PLT (10^9/L)		176.00 (139.00, 211.00)	176.00 (139.00, 210.00)	177.00 (133.00, 218.00)	0.862
PLT/Lym		196.34 (137.82, 285.11)	195.00 (135.63, 278.00)	201.12 (153.85, 309.72)	0.103
PCT (%)		0.17 (0.14, 0.21)	0.17 (0.14, 0.21)	0.17 (0.14, 0.21)	0.977
MPV (fl)		9.90 (9.20, 10.70)	9.90 (9.20, 10.70)	9.90 (9.10, 11.00)	0.787

Table 1 (continued)

Indicator		Total patients (N = 516)	Normal LOS group (N = 405)	Prolonged LOS group (N = 111)	P-value
PDW (%)		16.30 (16.00, 16.50)	16.30 (16.00, 16.50)	16.30 (16.10, 16.50)	0.619
CysC (mg/L)		1.01 (0.86, 1.26)	1.00 (0.85, 1.22)	1.09 (0.92, 1.44)	**0.012**
PAB (mg/L)		230.50 (185.80, 269.30)	233.00 (188.50, 270.70)	224.20 (177.70, 260.80)	0.211
HBDH (U/L)		163.60 (137.30, 208.50)	162.10 (136.60, 206.90)	170.70 (140.20, 220.70)	0.162
CK (U/L)		94.70 (61.10, 184.90)	96.90 (61.20, 184.90)	92.90 (61.10, 186.40)	0.897
LDH (U/L)		221.70 (186.60, 283.00)	218.30 (184.50, 277.00)	233.70 (194.20, 318.20)	0.076
Serum P^+ (mmol/L)		1.13 (0.94, 1.33)	1.13 (0.93, 1.32)	1.17 (0.99, 1.34)	0.233
Serum Ca^{2+} (mmol/L)		2.16 (2.07, 2.22)	2.16 (2.08, 2.22)	2.15 (2.06, 2.20)	0.089
Serum Cl^- (mmol/L)		105.62 ± 3.39	105.61 ± 3.43	105.64 ± 3.25	0.943
Serum Na^+ (mmol/L)		139.90 (138.10, 141.50)	139.90 (138.10, 141.50)	139.80 (138.00, 141.70)	0.750
Serum K^+ (mmol/L)		3.83 (3.53, 4.06)	3.81 (3.51, 4.04)	3.92 (3.63, 4.24)	**0.013**
LDL-C (mmol/L)		2.42 (1.93, 3.07)	2.45 (2.01, 3.11)	2.17 (1.80, 2.83)	**0.027**
HDL-C (mmol/L)		0.96 (0.79, 1.19)	0.97 (0.82, 1.21)	0.84 (0.72, 1.04)	**< 0.001**
TG (mmol/L)		1.13 (0.84, 1.60)	1.11 (0.82, 1.56)	1.14 (0.88, 1.69)	0.133
TC (mmol/L)		4.14 (3.50, 4.79)	4.19 (3.54, 4.83)	3.94 (3.27, 4.51)	**0.013**
GLU (mmol/L)		6.90 (5.85,7.94)	6.92 (5.84,7.96)	6.78 (5.90, 7.81)	0.898
UA (μmol/L)		390.80 (307.10, 474.50)	387.00 (307.10, 462.00)	407.40 (309.40, 504.30)	0.146
CrS (μmol/L)		80.90 (65.00, 110.00)	79.20 (64.80, 107.40)	87.20 (67.60, 123.50)	0.059
Urea (mmol/L)		7.20 (5.70, 9.50)	7.00 (5.60, 9.10)	7.90 (6.30, 10.80)	**0.010**
A/G		1.70 (1.40, 1.90)	1.70 (1.50, 1.90)	1.60 (1.40, 1.80)	0.074
GLB (g/L)		23.50 (20.70, 25.90)	23.40 (20.70, 26.10)	23.50 (20.30, 25.40)	0.663
ALB (g/L)		38.70 (35.60, 41.10)	39.00 (35.70, 41.40)	37.60 (35.00, 39.60)	**0.004**
Urea/GLB		0.31 (0.23, 0.43)	0.31 (0.23, 0.41)	0.34 (0.26, 0.48)	**0.017**
TP (g/L)		62.00 (57.80, 65.90)	62.20 (58.30, 66.20)	60.80 (56.00, 64.20)	**0.039**
ALP (U/L)		68.80 (55.00, 83.10)	69.00 (55.20, 83.50)	68.20 (54.40, 78.10)	0.430
GGT (U/L)		27.80 (17.60, 51.20)	27.80 (17.60, 50.90)	29.60 (17.40, 51.70)	0.806
AST (U/L)		20.30 (15.30, 31.50)	19.70 (15.20, 30.50)	22.00 (16.80, 33.40)	0.293
ALT (U/L)		20.90 (13.20, 36.60)	21.00 (12.60, 37.20)	20.70 (16.50, 34.60)	0.337
IBIL (μmol/L)		10.40 (8.00, 14.90)	10.40 (8.00, 14.80)	10.50 (8.00, 15.40)	0.826
DBIL (μmol/L)		5.80 (4.30, 8.30)	5.60 (4.30, 8.10)	6.40 (4.50, 9.00)	0.103
TBIL (μmol/L)		16.40 (12.40, 23.20)	16.00 (12.30, 23.00)	17.70 (13.20, 23.90)	0.387
PT (sec)		14.00 (13.50, 14.70)	14.00 (13.40, 14.60)	14.30 (13.80, 15.00)	**0.002**
APTT (sec)		36.50 (33.30, 40.40)	36.40 (33.30, 40.40)	36.50 (33.60, 40.10)	0.734
TT (sec)		16.70 (15.80, 17.60)	16.70 (15.70, 17.60)	16.70 (16.00, 17.60)	0.409
Fbg (g/L)		2.81 (2.16, 3.72)	2.82 (2.18, 3.67)	2.69 (2.07, 3.80)	0.765
INR		1.10 (1.05, 1.17)	1.09 (1.04, 1.16)	1.12 (1.07, 1.21)	**0.001**
AT-IIIA (%)		92.00 (85.00, 101.00)	92.00 (85.00, 101.00)	93.00 (86.00, 101.00)	0.543
DD (n, %)	**0.0–0.5**	40 (7.75)	34 (8.40)	6 (5.41)	Na
	0.5–20.0	363 (70.35)	283 (69.88)	80 (72.07)	
	> 20.0	113 (21.90)	88 (21.73)	25 (22.52)	
CK-MB (ng/mL)		1.93 (1.09, 3.62)	1.87 (1.08, 3.47)	2.18 (1.18, 4.16)	0.230
Myo (ng/mL)		46.88 (28.79, 96.39)	45.86 (26.71, 96.39)	52.71 (33.81, 88.93)	0.172
NT-proBNP (pg/mL)		178.30 (81.82, 421.20)	167.00 (78.23, 384.80)	219.70 (108.20, 532.00)	**0.032**
hs-cTnT (pg/mL)		12.67 (8.16, 30.04)	11.86 (7.79, 27.21)	19.35 (8.88, 36.53)	**0.006**
FDP (mg/L)		15.95 (5.88, 45.86)	15.84 (5.46, 46.45)	16.17 (6.52, 42.68)	0.819
MR (n, %)	**Normal**	271 (52.52)	230 (56.79)	41 (36.94)	Na
	Mild	236 (45.74)	168 (41.48)	68 (61.26)	
	Moderate	8 (1.55)	6 (1.48)	2 (1.80)	
	Heavy	1 (0.19)	1 (0.25)	0 (0.00)	
AR (n, %)	**Normal**	187 (36.24)	157 (38.77)	30 (27.03)	**0.031**

Table 1 (continued)

Indicator		Total patients (N = 516)	Normal LOS group (N = 405)	Prolonged LOS group (N = 111)	P-value
	Mild	235 (45.54)	179 (44.20)	56 (50.45)	
	Moderate	78 (15.12)	60 (14.81)	18 (16.22)	
	Heavy	16 (3.10)	9 (2.22)	7 (6.31)	
TR (n, %)	**Normal**	286 (55.43)	242 (59.75)	44 (39.64)	**< 0.001**
	Mild	226 (43.80)	161 (39.75)	65 (58.56)	
	Moderate	4 (0.78)	2 (0.49)	2 (1.80)	
AA (mm)		41.00 (37.60, 45.00)	40.90 (37.00, 44.70)	43.00 (39.00, 47.00)	**< 0.001**
LVEF (%)		62.00 (59.00, 64.00)	62.00 (59.00, 64.00)	61.00 (58.00, 64.00)	0.225
LICU (d)		7.00 (5.00, 11.00)	7.00 (4.00, 10.00)	11.00 (7.00, 19.00)	**< 0.001**
LOS (d)		25.00 (19.00, 30.00)	23.00 (18.00, 27.00)	38.00 (33.00, 43.00)	**< 0.001**
OT (min)		345.00 (141.00, 465.00)	310.00 (130.00, 450.00)	420.00 (345.00, 490.00)	**< 0.001**
HOS (¥)		31.30 (19.90,39.50)	27.80 (18.70,36.20)	42.00 (34.40,51.50)	**< 0.001**

Continuous variables were presented as mean ± standard deviations (SD) or median (interquartile spacing). Categorical variables were presented as numerical values and proportions. Normal LOS group indicates patients with length of hospital stay of less than 30 days; Prolonged LOS group indicates patients with length of hospital stay greater than 30 days

AD aortic dissection, *BSA* body surface area, *BMI* body mass index, *SBP* systolic blood pressure, *DBP* diastolic blood pressure, *HepIS* hepatic insufficiency, *CHD* coronary heart disease, *HTN* hypertension, *DM* diabetes mellitus, *ARD* acute renal dysfunction, *CIH* cerebral infarction history, *hsCRP* hypersensitivity-c-reactive protein, *WBC* white blood cell, *Lym* lymphocyte count, *Mon* monocyte count, *Neu*, neutrophil count, *NLR* ratio of neutrophil to lymphocyte, *Eos* eosinophil count, *Bas* basophil count, *LymP* percentage of lymphocyte, *MonP* percentage of monocyte, *NeuP* percentage of neutrophil, *EosP* percentage of eosinophil, *BasP* percentage of basophil, *RBC* red blood cell, *HGB* hemoglobin, *HTC* hematocrit, *MCV* mean corpuscular volume, *MCH* mean corpuscular hemoglobin, *MCHC* mean corpuscular hemoglobin concentration, *RDW* red blood cell distribution width, *PLT* platelet, *PLT/Lym* ratio of platelet to lymphocyte, *PCT* platelet ratio, *MPV* mean platelet volume, *PDW* platelet volume distribution width, *CysC* cystatin-c, *PAB* prealbumin, *HBDH* α-hydroxybutyrate dehydrogenase, *CK* creatine kinase, *LDH* lactate dehydrogenase, P^{+} phosphorus, Ca^{2+} calcium, Cl^{-} chlorine, Na^{+} sodium, K^{+} Potassium, *LDL-C* low density lipoprotein cholesterol, *HDL-C* high density lipoprotein cholesterol, *TG* triglyceride, *TC* total cholesterol, *GLU* glucose, *UA* uric acid, *CrS* serum creatinine, *ALB* albumin, *GLB* globulin, *A/G* ratio of albumin to globulin, *Urea/GLB* ratio of Urea to globulin, *TP* total protein, *ALP* alkaline phosphatase, *GGT* L-glutamine, *AST* aspartate transaminase, *ALT* alanine transaminase, *IBIL* indirect bilirubin, *DBIL* direct bilirubin, *TBIL* total bilirubin, *PT* prothrombin time, *APTT* activated partial thromboplastin time, *TT* thrombin time, *Fbg* fibrinogen, *INR* international Normalized ratio, *AT-IIIA* antithrombin-III, *DD* D-dimer, *CK-MB* creatine kinase isoenzymes, *Myo* myoglobin, *NT-proBNP* N-Terminal pro-brain natriuretic peptide, *hs-cTnT* high-sensitivity troponin T, *FDP* fibrinogen degradation product, *MR* mitral regurgitation, *AR* aortic regurgitation, *TR* tricuspid regurgitation, *AA* diameter of aorta, *LVEF* left ventricular eject fraction, *LICU* length of ICU stay, *LOS* length of hospital stay, *OT* operational time, *HOS* hospital cost, *Na* not applicable

models perform continuous basic operations, and the complex and unpredictable nature of human physiology has been shown to be better captured by machine learning algorithms [18, 19]. This capability allows for the exploration of intricate correlation features, contributing to the increasing adoption of machine learning in the field of medicine [20, 21]. Significant advancements in machine learning have demonstrated substantial potential in disease diagnosis, complication monitoring, and prognosis prediction [22]. In this study, AD patients were divided into either the normal or prolonged LOS group according to binary classification (a threshold value of 30 days of LOS). The variables associated with prolonged LOS were screened and selected by the Boruta algorithm. Subsequently, these variables were used to develop machine learning-based models, including XGBoost, AdaBoost, KNN, logistic regression, LightGBM, GNB, MLP, CNB, and SVM. Our findings revealed that the XGBoost model had outstanding predictive performance in prolonged LOS risk in AD patients, evidenced by its highest AUC values in both the training and validation subsets. Additionally, other parameters of the model further supported this finding. Therefore, the XGBoost model was considered the final prediction model for prolonged LOS in AD patients.

The SHAP analysis can offer a novel approach to interpreting machine learning models, proving effective for both local and global interpretability. Furthermore, the SHAP provides a more pragmatic elucidation of the effectiveness of machine learning algorithms in delivering precise predictions for specific patient cohorts. In our study, machine learning algorithms-based models and the SHAP method was combined to improve the accuracy in predicting risk factors for prolonged LOS in AD patients. Moreover, the SHAP analysis showed intuitive explanations to assist clinicians in comprehensive understanding of the decision-making process for assessing disease severity and optimizing opportunities for early intervention [23]. In this study, the SHAP analysis for the XGBoost identified HDL-C, SBP, LymP, ALT, and OT as the top five variables affecting prolonged LOS in AD patients.

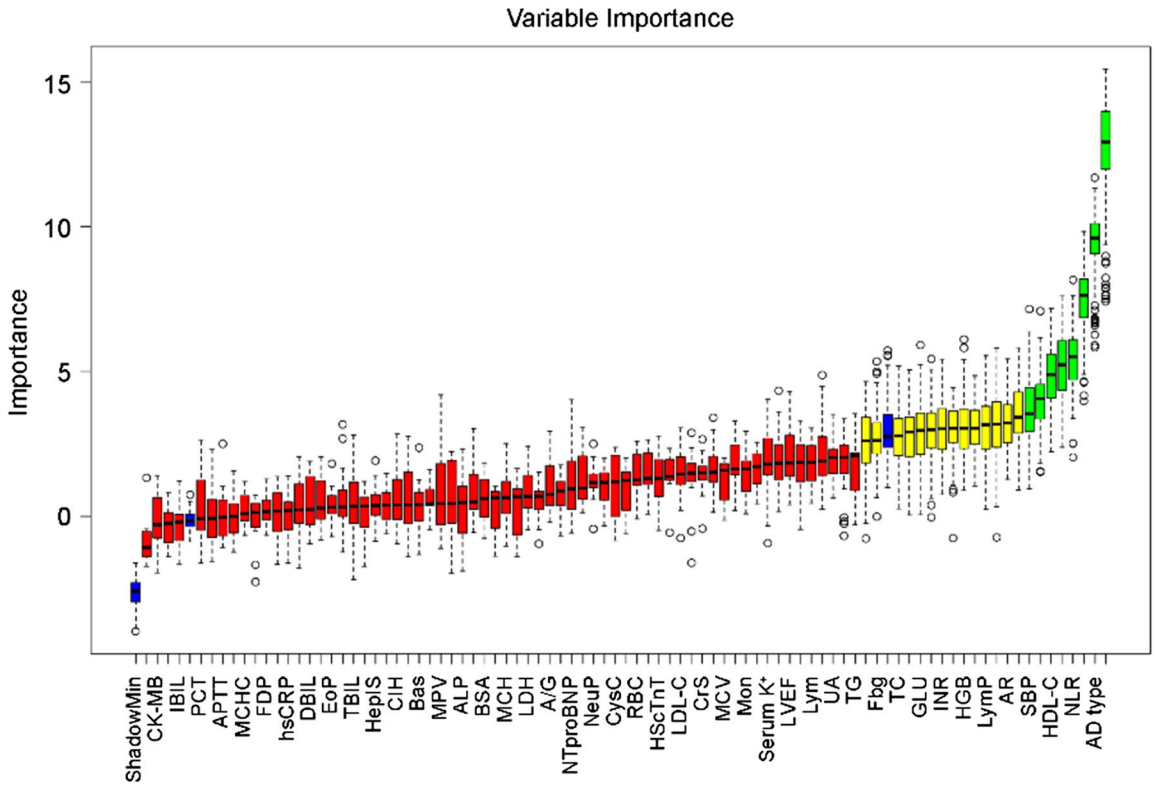

Fig. 2 The Boruta algorithm for feature selection. The features were selected for developing models when the importance of features exceeds the ShadowMax value

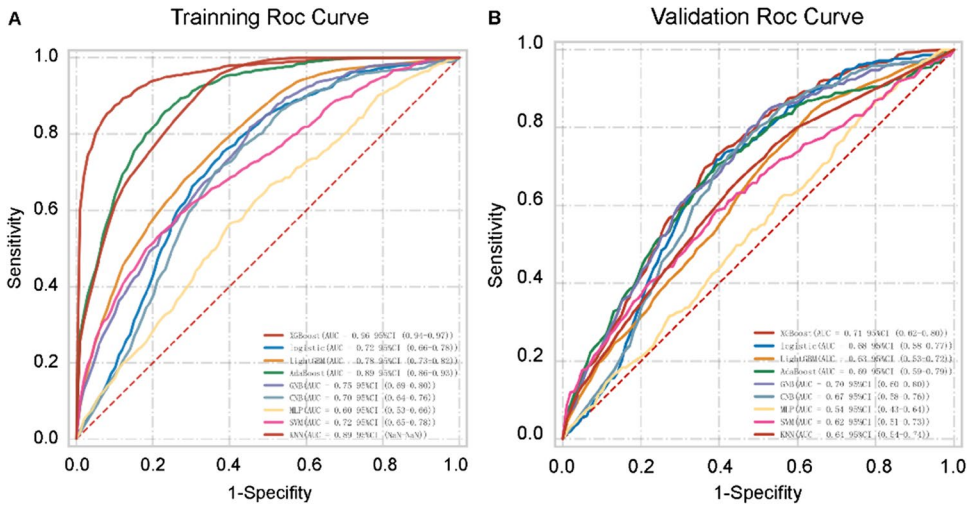

Fig. 3 The training and validation ROC curves of machine learning-based models. (A) The comparison of machine learning-based models in the area under the receiver operating characteristic curve in the training set. **(B)** The comparison of machine learning-based models in the area under the receiver operating characteristic curve in the validation set. KNN, k-nearest neighbors; GNB, gaussian naive bayes; SVM, support vector machine; CNB, complement naive bayes; MLP, multi-layer perceptron neural network

HDL-C can efficiently transport and remove excess cholesterol from peripheral tissues, thereby facilitating its return to the liver [24]. The study has reported that low levels of HDL-C increase the susceptibility to cardiovascular and cerebrovascular events in patients following cardiac intervention therapy [25]. A prospective cohort study involving 3 million participants revealed a 'J-shaped' association between HDL-C levels and cardiovascular disease mortality, where both high and low HDL-C levels were related with increased risk [26]. In this study, HDL-C levels in the prolonged LOS group were significantly lower than those in the normal LOS group. Although we did not observe a 'J-shaped' relationship between HDL-C levels and LOS in AD patients, our study identified that reduced HDL-C levels might contribute to prolonged LOS in these patients. Additionally, elevated HDL-C levels have been shown to reduce

Table 2 Models' performance in the training subset

Models	AUC	Sensitivity	Specificity	ACC	PPV	NPV	F1-Score
XGBoost	0.96 (0.94–0.97)	0.92 (0.87–0.96)	0.90 (0.85–0.95)	0.90 (0.86–0.95)	0.75 (0.64–0.86)	0.97 (0.96–0.99)	0.82 (0.74–0.90)
KNN	0.89 (Na-Na)	0.66 (0.57–0.74)	0.89 (0.86–0.93)	0.84 (0.81–0.88)	0.63 (0.54–0.72)	0.91 (0.89–0.93)	0.64 (0.56–0.72)
AdaBoost	0.89 (0.86–0.93)	0.86 (0.82–0.90)	0.79 (0.76–0.83)	0.81 (0.78–0.84)	0.54 (0.49–0.59)	0.95 (0.94–0.97)	0.66 (0.62–0.70)
LightGBM	0.78 (0.73–0.82)	0.51 (0.28–0.73)	0.85 (0.77–0.92)	0.77 (0.75–0.79)	Na	0.87 (0.83–0.91)	Na
GNB	0.75 (0.69–0.80)	0.82 (0.77–0.87)	0.56 (0.50–0.61)	0.61 (0.58–0.65)	0.34 (0.33–0.36)	0.92 (0.91–0.93)	0.48 (0.47–0.49)
Logistic	0.72 (0.66–0.78)	0.78 (0.74–0.82)	0.60 (0.56–0.65)	0.64 (0.61–0.67)	0.35 (0.33–0.37)	0.91 (0.90–0.92)	0.48 (0.47–0.50)
SVM	0.72 (0.65–0.78)	0.58 (0.47–0.68)	0.77 (0.72–0.83)	0.73 (0.70–0.76)	0.42 (0.38–0.45)	0.87 (0.85–0.90)	0.47 (0.43–0.51)
CNB	0.70 (0.64–0.76)	0.78 (0.72–0.84)	0.58 (0.52–0.64)	0.62 (0.58–0.66)	0.33 (0.32–0.35)	0.91 (0.90–0.93)	0.46 (0.45–0.48)
MLP	0.60 (0.53–0.66)	0.73 (0.59–0.88)	0.48 (0.33–0.64)	0.53 (0.44–0.63)	0.29 (0.26–0.33)	0.88 (0.85–0.91)	0.39 (0.35–0.44)

Results were presented as mean and 95%CI

95%CI 95% confidence interval, *AUC* area under the receiver operating characteristic curve, *ACC* accuracy, *PPV* positive prediction value, *NPV* negative prediction value, *KNN* k-nearest neighbors, *GNB* gaussian naive bayes, *SVM*, support vector machine, *CNB* complement naive bayes, *MLP* multi-layer perceptron neural network, *Na* not applicable

Table 3 Models' performance in the validation subset

Models	AUC	Sensitivity	Specificity	ACC	PPV	NPV	F1-Score
XGBoost	0.71 (0.62–0.80)	0.76 (0.68–0.83)	0.84 (0.78–0.89)	0.71 (0.68–0.74)	0.37 (0.33–0.42)	0.83 (0.81–0.85)	0.38 (0.35–0.40)
KNN	0.64 (0.54–0.74)	0.28 (0.19–0.37)	0.79 (0.74–0.85)	0.71 (0.68–0.74)	Na	0.80 (0.78–0.82)	Na
AdaBoost	0.69 (0.59–0.79)	0.54 (0.48–0.60)	0.72 (0.68–0.76)	0.68 (0.65–0.71)	0.35 (0.31–0.38)	0.85 (0.82–0.87)	0.42 (0.38–0.45)
LightGBM	0.63 (0.53–0.72)	0.30 (0.16–0.43)	0.82 (0.72–0.91)	0.71 (0.66–0.76)	Na	0.82 (0.80–0.84)	Na
GNB	0.70 (0.60–0.80)	0.42 (0.35–0.48)	0.53 (0.47–0.59)	0.58 (0.54–0.61)	0.30 (0.27–0.33)	0.89 (0.87–0.92)	0.43 (0.40–0.46)
Logistic	0.68 (0.58–0.77)	0.72 (0.68–0.76)	0.56 (0.52–0.61)	0.60 (0.57–0.63)	0.32 (0.29–0.34)	0.88 (0.86–0.90)	0.44 (0.41–0.46)
SVM	0.62 (0.51–0.73)	0.49 (0.39–0.60)	0.72 (0.65–0.80)	0.67 (0.64–0.71)	0.34 (0.29–0.39)	0.84 (0.82–0.86)	0.39 (0.34–0.43)
CNB	0.67 (0.58–0.76)	0.75 (0.66–0.83)	0.53 (0.44–0.63)	0.58 (0.52–0.64)	0.33 (0.30–0.35)	0.89 (0.86–0.92)	0.45 (0.42–0.47)
MLP	0.54 (0.43–0.64)	0.68 (0.53–0.84)	0.42 (0.28–0.57)	0.49 (0.40–0.57)	0.25 (0.23–0.28)	0.84 (0.81–0.88)	0.36 (0.31–0.41)

Results were presented as mean and 95%CI

95%CI 95% confidence interval, *AUC* area under the receiver operating characteristic curve, *ACC* accuracy, *PPV* positive prediction value, *NPV* negative prediction value, *KNN* k-nearest neighbors, *GNB* gaussian naive bayes, *SVM* support vector machine, *CNB* complement naive bayes, *MLP* multi-layer perceptron neural network, *Na* not applicable

the risk of delirium and shorter LOS in patients following cardiac surgery [27]. Our findings of the SHAP analysis revealed that HDL-C was identified as the most crucial risk variable for prolonged LOS. These findings implied that HDL-C of patients should to be noticed by clinicians in clinical practice.

SBP is the highest blood pressure exerted on the active artery when the left ventricle contracts. A 2023 Mendelian randomization analysis demonstrated that elevated SBP was significantly associated with an increased risk of both aortic aneurysm and AD in patients [28]. Indeed, SBP can serve as a robust prognostic indicator for mortality in various cardiovascular conditions, including acute coronary syndrome, cardiogenic shock, and acute heart failure. Likewise, a study conducted by Bossone, E et al. revealed a logarithmic correlation between SBP and cardiovascular risk [29]. The inadequate control of preoperative SBP may elevate the susceptibility to perioperative nervous system complications [30]. Additionally, maintaining appropriate preoperative SBP is crucial for ensuring intraoperative hemodynamic stability, shortening OT, reducing the risk of stroke and postoperative delirium, as well as other neurologic complications and mortality [31]. Simultaneously, elevated SBP levels heighten susceptibility to arterial sclerosis, aneurysm formation, and acute aortic syndrome [32, 33]. Although there are limited studies that have included SBP as an independent indicator, the results of the SHAP analysis showed that elevated SBP levels were important predictors of prolonged LOS in patients with AD. These results might offer valuable insights for future blood pressure control programs.

LymP serves as a crucial indicator for monitoring immune function and reflects the level of inflammation in the body [34, 35]. Current studies in cardiovascular disease primarily focus on T lymphocytes, B lymphocytes, and NK cells. The

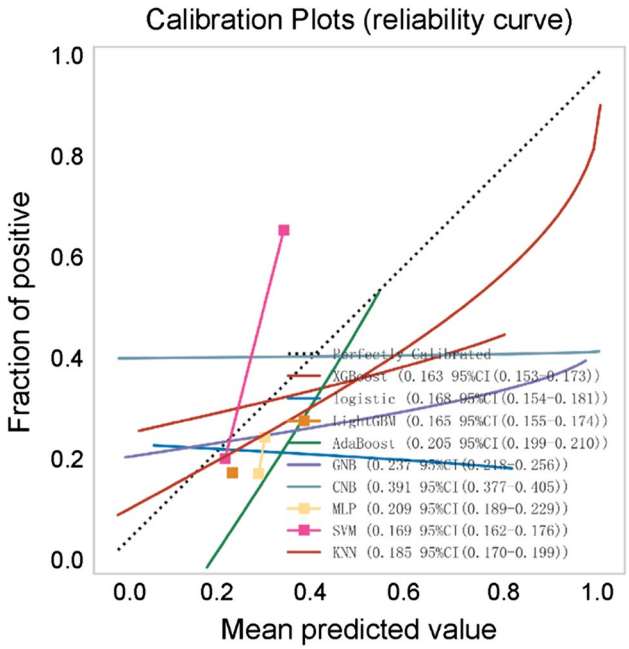

Fig. 4 Calibration plots of machine learning-based models. The comparison of machine learning-based models in calibration curve fitting is shown by Brier score and 95% confidence interval (CI). KNN, k-nearest neighbors; GNB, gaussian naive bayes; SVM, support vector machine; CNB, complement naive bayes; MLP, multilayer perceptron neural network

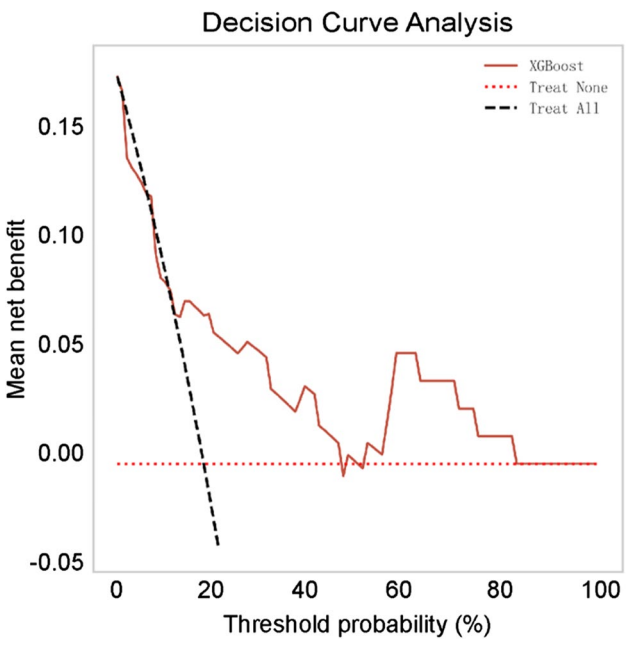

Fig. 5 Decision curve analysis. The decision curve analysis (DCA) for the XGBoost model predicting prolonged length of stay in patients with aortic dissection. The DCA curve shows that the miscalibration offsets improved discrimination: the XGBoost model has a higher net benefit compared to the clinical default strategies of "treat all" or "treat none", over the entire range of reasonable threshold probabilities

key role of LymP in cardiovascular disease is undisputed, as it accelerates the pathological process through the secretion of proinflammatory cytokines. Equally well-established is the fact that autoimmune diseases, mediated by autoreactive T cells, significantly increase the risk of developing cardiovascular disease [36]. NLR has been utilized as a prognostic indicator for cardiovascular surgery, with numerous studies highlighting its correlation with the duration of postoperative hospitalization [37]. In this study, NLR and LymP exhibited statistical significance between the prolonged and normal LOS groups in AD patients. However, the SHAP analysis suggested that LymP might have a more prominent role than NLR. Previous studies found that lymphopenia may lead to elevated creatinine levels, prolonged mechanical ventilation duration, increased risk of arrhythmia following cardiac surgery, and extended LOS [38–40]. Consistently, our study identified LymP as a significant variable for predicting prolonged LOS in patients with AD. With the emergence of novel indicators, greater attention should be devoted to the role of LymP in clinical practice in the future.

ALT, an enzyme primarily synthesized by the liver, serves as a clinical indicator for hepatic injury and an independent determinant of cardiovascular disease [41, 42]. A recent study involving 2,565 patients found that elevated levels of ALT were significantly associated with prolonged LOS and increased mortality in individuals undergoing cardiac surgery [43]. Furthermore, higher preoperative ALT levels are correlated with increased drainage volume following aortic arch surgery, highlighting the significant prognostic relevance of ALT levels for patient outcomes [44]. As ALT levels increased in our study, AD patients were more likely to experience a prolonged LOS, suggesting that monitoring this index could provide valuable guidance for future clinical treatment strategies. Notably, the SHAP analysis identified ALT as an important risk factor, following LymP.

OT in patients undergoing cardiac surgery is determined by multiple factors, primarily involving disease severity, extracorporeal circulation time, and surgical technique. Numerous studies have demonstrated that OT closely correlates with the occurrence of postoperative complications, and increased OT may raise the risk of postoperative infection, mechanical ventilation, delirium, and prolonged LICU [45–48]. Indeed, the specific timing for defining aortic dissection surgery lacks relevant research. Our study employed a machine learning prediction model to effectively examine risk factors associated with prolonged LOS, identifying OT as a key contributing factor. Prolonged OT was detrimental to the recovery of patients with AD, indicating the necessity for medical practitioners to be mindful of this aspect in clinical practice. The prolongation of OT is associated with

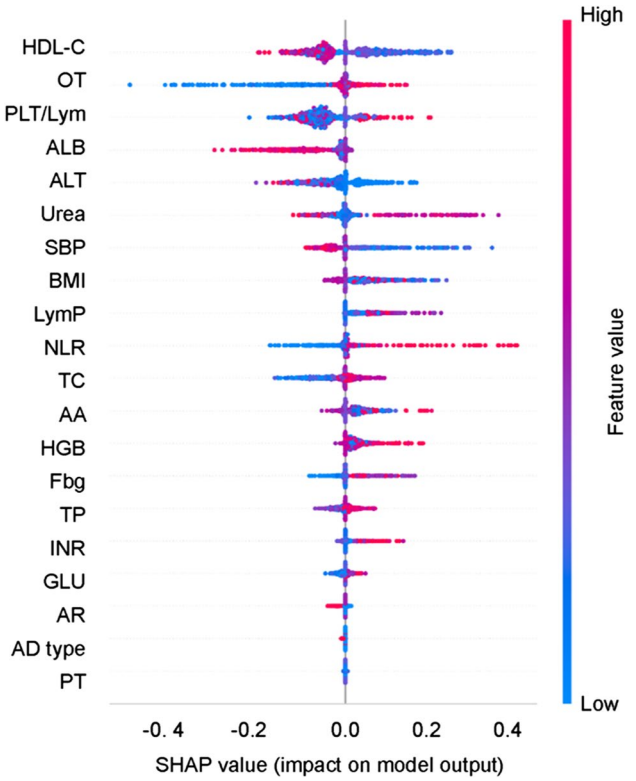

Fig. 6 The XGBoost model's interpretation based on the SHAP. The positive or negative effects of top twenty variables on prolonged length of stay in patients with aortic dissection (AD)

an elevated postoperative risk in open surgery, which hinders patients' recovery and leads to prolonged LOS and increased treatment expenses. Future treatment methods should prioritize tailoring the specific timing of operations based on individual conditions, thereby reducing OT, facilitating patient recovery, and minimizing the risk of postoperative complications.

AD patients undergoing surgery do not imply the completion of treatment program; instead, it means that they may be entering a period characterized by high risk and multifaceted challenges. These situations require us to be better prepared before surgery, and the lack of preoperative assessment of prolonged LOS may be a prominent issue. The application of AI-based machine learning in the medical field is important and necessary. It can not only help doctors improve their work efficiency but also provide more accurate diagnosis and treatment plans, thereby improving the quality of patients' life. Our study aimed to identify preoperative risk variables associated with prolonged LOS in AD patients using machine learning-based models. With the help of machine learning AI tools, the XGBoost model was successfully simplified into easy-to-understand graphs, which enhanced its practicality for clinicians.

However, although the machine learning models exhibited favorable predictive performance in this study, it is imperative to acknowledge certain limitations. Firstly, the data of patients in this study were obtained solely from a single center, leading to some biases in the results. This might affect the development of the models and risk identification to some extent. Secondly, surgical manners and therapeutic strategies adopted by the doctors might affect the final prognosis. Although all the subjects in this study were diagnosed with AD, different surgical manners and therapeutic strategies might lead to different LOS. Finally, there was a lack of utilization of the external data for model validation, potentially impacting the generalizability of the models. The above limitations might restrict the applicability and generalizability of machine learning models in individuals diagnosed with AD. Factually, machine learning-based AI tools primarily rely on computer systems, thereby exhibiting characteristics of simplicity and operability. This feature facilitates convenience for scientific research or clinical work. The future plans for this study involve expanding the sample size and establishing multi-center cooperation, we hope that external data can be obtained and used for the external validation. This will enhance the reliability of the relevant conclusions. Simultaneously, machine learning will be employed to further investigate clinical practice, including complementary imaging data and clinical intervention, with the aim of providing clinicians with a more robust decision-making framework and improving medical care quality.

Conclusion

In this study, the XGBoost model for predicting prolonged LOS in patients with AD was successfully established by machine learning algorithm. The SHAP analysis for the

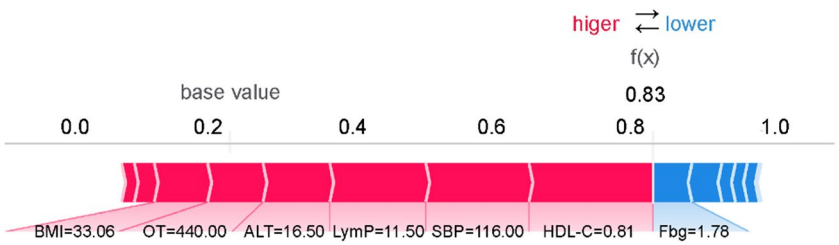

Fig. 7 Individual analysis of SHAP. The SHAP force plot for explaining results of individual's prediction in the validation cohort. SHAP, SHapley Additive explanation

model identified HDL-C, SBP, LymP, ALT, and OT as the top five crucial risk factors affecting prolonged LOS in AD patients. This may help clinicians enhance the preoperative management and thereby reduce LOS in patients with AD.

Data Availability Statement

All data can be obtained from the corresponding author for reasonable reasons.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12265-024-10565-z.

Authors Contributions Luo Li and Yihuan Chen collected data, analyzed the data, and wrote the draft manuscript. Hui Xie and Peng Zheng collected and supervised the data. Gaohang Mu and Qian Li managed the data. Haoyue Huang and Zhenya Shen were responsible for the overall project, designed the study, supervised the study, and reviewed the draft manuscript. All authors read and approved the final manuscript.

Funding No research funding was received for this study.

Declarations

Ethics Affirmation The protocol of this study was approved by the Ethics Review Committee of the First Affiliated Hospital of Soochow University (No. 2021–424), and the study protocol was in accordance with the Declaration of Helsinki (revised in 2013). Informed consent of the patients was waived by the First Affiliated Hospital of Soochow University.

Competing Interest The authors declare no competing interests.

References

1. Sacks D, et al. Multisociety Consensus Quality Improvement Revised Consensus Statement for Endovascular Therapy of Acute Ischemic Stroke. Int J Stroke. 2018;13(6):612–32.
2. Gawinecka J, Schönrath F, von Eckardstein A. Acute aortic dissection: pathogenesis, risk factors and diagnosis. Swiss Med Wkly. 2017;147: w14489.
3. Hu YY, et al. Non-coding RNAs Regulate the Pathogenesis of Aortic Dissection. Front Cardiovasc Med. 2022;9: 890607.
4. Salmasi MY, et al. The risk of misdiagnosis in acute thoracic aortic dissection: a review of current guidelines. Heart. 2020;106(12):885–91.
5. Malaisrie SC, et al. 2021 The American Association for Thoracic Surgery expert consensus document: Surgical treatment of acute type A aortic dissection. J Thorac Cardiovasc Surg. 2021;162(3):735-758.e2.
6. Evangelista A, et al. Insights From the International Registry of Acute Aortic Dissection: A 20-Year Experience of Collaborative Clinical Research. Circulation. 2018;137(17):1846–60.
7. Ramos Dos Santos PM, et al. Effects of early mobilisation in patients after cardiac surgery: a systematic review. Physiotherapy. 2017;103(1): 1–12.
8. Medeiros NB, et al. Forecasting the length-of-stay of pediatric patients in hospitals: a scoping review. BMC Health Serv Res. 2021;21(1):938.
9. AbuRuz ME, Momani A, Shajrawi A. The Association Between Depressive Symptoms and Length of Hospital Stay Following Coronary Artery Bypass Graft is Moderated by Perceived Control. Risk Manag Healthc Policy. 2021;14:1499–507.
10. Gupta D, et al. Prolonged hospital length of stay after pediatric heart transplantation: A machine learning and logistic regression predictive model from the Pediatric Heart Transplant Society. J Heart Lung Transplant. 2022;41(9):1248–57.
11. Greener JG, et al. A guide to machine learning for biologists. Nat Rev Mol Cell Biol. 2022;23(1):40–55.
12. Li B, et al. Machine learning in vascular surgery: a systematic review and critical appraisal. NPJ Digit Med. 2022;5(1):7.
13. Ahsan MM, Siddique Z. Machine learning-based heart disease diagnosis: A systematic literature review. Artif Intell Med. 2022;128: 102289.
14. Wang K, et al. Interpretable prediction of 3-year all-cause mortality in patients with heart failure caused by coronary heart disease based on machine learning and SHAP. Comput Biol Med. 2021;137: 104813.
15. Ali MM, et al. Heart disease prediction using supervised machine learning algorithms: Performance analysis and comparison. Comput Biol Med. 2021;136: 104672.
16. Isselbacher EM, et al. 2022 ACC/AHA Guideline for the Diagnosis and Management of Aortic Disease: A Report of the American Heart Association/American College of Cardiology Joint Committee on Clinical Practice Guidelines. Circulation. 2022;146(24):e334–482.
17. Al'Aref SJ, et al. Clinical applications of machine learning in cardiovascular disease and its relevance to cardiac imaging. Eur Heart J. 2019;40(24):1975–86.
18. Eraslan G, et al. Deep learning: new computational modelling techniques for genomics. Nat Rev Genet. 2019;20(7):389–403.
19. Heo J, et al. Machine Learning-Based Model for Prediction of Outcomes in Acute Stroke. Stroke. 2019;50(5):1263–5.
20. Chen B, et al. Differentiation of Low-Grade Astrocytoma From Anaplastic Astrocytoma Using Radiomics-Based Machine Learning Techniques. Front Oncol. 2021;11: 521313.
21. Kaplan E, et al. Automated BI-RADS classification of lesions using pyramid triple deep feature generator technique on breast ultrasound images. Med Eng Phys. 2022;108: 103895.
22. Zhou Y, et al. Machine learning predictive models for acute pancreatitis: A systematic review. Int J Med Inform. 2022;157: 104641.
23. Fan Y, Mao H, Li Q. A Model-Agnostic Feature Attribution Approach to Magnetoencephalography Predictions Based on Shapley Value. IEEE J Biomed Health Inform. 2023;27(5):2524–35.
24. Chen L, et al. Very high HDL-C (high-density lipoprotein cholesterol) is associated with increased cardiovascular risk in patients

with NSTEMI (non-ST-segment elevation myocardial infarction) undergoing PCI (percutaneous coronary intervention). BMC Cardiovasc Disord. 2023;23(1):357.
25. Ren X, Wang X. Association of the low-density lipoprotein cholesterol to high-density lipoprotein cholesterol ratio and major adverse cardiac and cerebrovascular events in patients with coronary heart disease undergoing percutaneous coronary intervention: a cohort study. Curr Med Res Opin. 2023;39(9):1175–81.
26. Zhong GC, et al. HDL-C is associated with mortality from all causes, cardiovascular disease and cancer in a J-shaped dose-response fashion: a pooled analysis of 37 prospective cohort studies. Eur J Prev Cardiol. 2020;27(11):1187–203.
27. Li J, et al. Risk factors for delirium after coronary artery bypass grafting in elderly patients. Ann Transl Med. 2021;9(22):1666.
28. Yang T, et al. Causal effect of hypertension and blood pressure on aortic diseases: evidence from Mendelian randomization. Hypertens Res. 2023;46(9):2203–12.
29. Bossone E, et al. Presenting Systolic Blood Pressure and Outcomes in Patients With Acute Aortic Dissection. J Am Coll Cardiol. 2018;71(13):1432–40.
30. Liu Y, Chen K, Mei W. Neurological complications after cardiac surgery: anesthetic considerations based on outcome evidence. Curr Opin Anaesthesiol. 2019;32(5):563–7.
31. Descamps R, et al. Perioperative individualized hemodynamic optimization according to baseline mean arterial pressure in cardiac surgery patients: Rationale and design of the OPTIPAM randomized trial. Am Heart J. 2023;261:10–20.
32. Falkner B, et al. Pediatric Primary Hypertension: An Underrecognized Condition: A Scientific Statement From the American Heart Association. Hypertension. 2023;80(6):e101–11.
33. Bossone E, Eagle KA. Epidemiology and management of aortic disease: aortic aneurysms and acute aortic syndromes. Nat Rev Cardiol. 2021;18(5):331–48.
34. Kvedaraite E. Neutrophil-T cell crosstalk in inflammatory bowel disease. Immunology. 2021;164(4):657–64.
35. Nguyen M, et al. Leukocyte cell population data in patients with cardiac surgery and cardiopulmonary bypass: A potential readily available tool to monitor immunity. Front Immunol. 2022;13:1101937.
36. Schwartz DM, et al. T Cells in Autoimmunity-Associated Cardiovascular Diseases. Front Immunol. 2020;11: 588776.
37. Mirna M, et al. Neutrophil-to-lymphocyte ratio and monocyte-to-lymphocyte ratio predict length of hospital stay in myocarditis. Sci Rep. 2021;11(1):18101.
38. Haran C, et al. Preoperative Neutrophil and Lymphocyte Ratio as a Predictor of Mortality and Morbidity After Cardiac Surgery. Heart Lung Circ. 2021;30(3):414–8.
39. Wheatley J, et al. The prognostic value of elevated neutrophil-lymphocyte ratio for cardiac surgery-associated acute kidney injury: A systematic review and meta-analysis. Acta Anaesthesiol Scand. 2023;67(2):131–41.
40. Charitakis E, et al. Risk and protective factors for atrial fibrillation after cardiac surgery and valvular interventions: an umbrella review of meta-analyses. Open Heart. 2022;9(2): e002074.
41. Hasan A, et al. Assessment of the Relationship Between Liver Enzymes and Cardiovascular Disease: A Study in Bangladeshi Adults. Endocrinol Diabetes Metab. 2024;7(2): e00481.
42. Bekkelund SI. Serum alanine aminotransferase activity and risk factors for cardiovascular disease in a Caucasian population: the Tromsø study. BMC Cardiovasc Disord. 2021;21(1):29.
43. Shang L, et al. Preoperative Liver Function Test Abnormalities Were Associated With Short-Term and Long-Term Prognosis in Cardiac Surgery Patients Without Liver Disease. Front Cardiovasc Med. 2021;8: 772430.
44. Yan W, et al. Elevated AST/ALT (De Ritis) Ratio is a Risk Factor of Drainage Volume after Aortic Arch Surgery. Heart Surg Forum. 2023;26(6):E755-e763.
45. Yu X, et al. Risk factors of nosocomial infection after cardiac surgery in children with congenital heart disease. BMC Infect Dis. 2020;20(1):64.
46. Shevchenko AA, et al. Surgical treatment of sternal osteomyelitis and sternomediastinitis following cardiac surgery. Khirurgiia (Mosk). 2021;9:34–9.
47. Aksoy R, et al. Predictive Factors of Prolonged Ventilation Following Cardiac Surgery with Cardiopulmonary Bypass. Braz J Cardiovasc Surg. 2021;36(6):780–7.
48. Habeeb-Allah A, Alshraideh JA. Delirium post-cardiac surgery: Incidence and associated factors. Nurs Crit Care. 2021;26(3):150–5.