Journal of the American Heart Association

ORIGINAL RESEARCH

Machine Learning–Based Risk Model for Predicting Early Mortality After Surgery for Infective Endocarditis

Li Luo, MD*; Sui-qing Huang, MD*; Chuang Liu, MCS; Quan Liu, MD; Shuohui Dong, MD; Yuan Yue, MD; Kai-zheng Liu, MD; Lin Huang, MD; Shun-jun Wang, MD; Hua-yang Li, MD; Shaoyi Zheng, PhD, MD; Zhong-kai Wu, PhD, MD

BACKGROUND: The early mortality after surgery for infective endocarditis is high. Although risk models help identify patients at high risk, most current scoring systems are inaccurate or inconvenient. The objective of this study was to construct an accurate and easy-to-use prediction model to identify patients at high risk of early mortality after surgery for infective endocarditis.

METHODS AND RESULTS: A total of 476 consecutive patients with infective endocarditis who underwent surgery at 2 centers were included. The development cohort consisted of 276 patients. Eight variables were selected from 89 potential predictors as input of the XGBoost model to train the prediction model, including platelet count, serum albumin, current heart failure, urine occult blood ≥(++), diastolic dysfunction, multiple valve involvement, tricuspid valve involvement, and vegetation >10 mm. The completed prediction model was tested in 2 separate cohorts for internal and external validation. The internal test cohort consisted of 125 patients independent of the development cohort, and the external test cohort consisted of 75 patients from another center. In the internal test cohort, the area under the curve was 0.813 (95% CI, 0.670–0.933) and in the external test cohort the area under the curve was 0.812 (95% CI, 0.606–0.956). The area under the curve was significantly higher than that of other ensemble learning models, logistic regression model, and European System for Cardiac Operative Risk Evaluation II (all, $P<0.01$). This model was used to develop an online, open-access calculator (http://42.240.140.58:1808/).

CONCLUSIONS: We constructed and validated an accurate and robust machine learning–based risk model to predict early mortality after surgery for infective endocarditis, which may help clinical decision-making and improve outcomes.

Key Words: cardiac surgery ■ infective endocarditis ■ machine learning ■ prognosis ■ risk model

Infective endocarditis (IE) is one of the most life-threatening cardiac diseases with a rising prevalence rate.[1] According to the 2016 Global Burden of Disease study,[2] the prevalence rate in China increased by 26.7% from 1990 to 2016. Globally, IE is still an important cause of death and disability, and in 2017, IE was associated with 2.23 million disability-adjusted life-years; an increase of 17.1% from that in 2007.[3]

Approximately 50% to 60% of patients with IE received surgical treatment.[4] Studies have shown that patients who meet the surgical indications recommended by guidelines have overall good outcomes.[5–7] However, surgery for IE is high risk, and the mortality rate within 30 days after surgery ranges from 8% to 30%.[8,9]

Risk models can efficiently identify patients with a high postoperative mortality risk and are useful for decision-making and patient counseling. Although recommended by recent guidelines, non-IE-specific scoring systems such as the Society of Thoracic

Correspondence to: Zhong-kai Wu, PhD, MD, 58 Zhongshan II Road, Guangzhou, 510080, Guangdong Province, P. R. China. Email: wuzhk@mail.sysu.edu.cn and Shaoyi Zheng, PhD, MD, No.1838 North Guangzhou Avenue, Baiyun District, 510000 Guangzhou, P. R. China. Email: zhsy@smu.edu.cn

*L. Luo and S.-Q. Huang are co-first authors.

Supplemental Material for this article is available at https://www.ahajournals.org/doi/suppl/10.1161/JAHA.122.025433

For Sources of Funding and Disclosures, see page 12.

© 2022 The Authors. Published on behalf of the American Heart Association, Inc., by Wiley. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

JAHA is available at: www.ahajournals.org/journal/jaha

CLINICAL PERSPECTIVE

What Is New?

- The infective endocarditis-specific, machine learning–based, clinical prediction model, Sun Yat-sen University Prediction Model for Infective Endocarditis, had a satisfactory predictive value across independent study samples and can be used to predict early mortality after surgery for IE and perform risk stratification.
- The model outperformed the European System for Cardiac Operative Risk Evaluation II in patients with infective endocarditis.
- The Sun Yat-sen University Prediction Model for Infective Endocarditis is the first infective endocarditis-specific risk model that uses urine occult blood, diastolic dysfunction, and tricuspid valve involvement as predictors.

What Are the Clinical Implications?

- The model can be used to identify patients with infective endocarditis at high risk of early mortality after surgery, and thus may help clinical decision-making and improve outcomes.

Nonstandard Abbreviations and Acronyms

EuroSCORE	European System for Cardiac Operative Risk Evaluation
FAH-SYSU	The First Affiliated Hospital of Sun Yat-sen University
LASSO	least absolute shrinkage and selection operator
NFH	Nanfang Hospital
STS	Society of Thoracic Surgeons
SYSUPMIE	Sun Yat-sen University Prediction Model for Infective Endocarditis
UOB	urine occult blood

Surgeons (STS) score and European System for Cardiac Operative Risk Evaluation II (EuroSCORE II) perform unsatisfactorily in patients with IE.[9–12] Patients with IE can have many associated comorbidities during the disease course, such as cardiovascular system disorders, infections, inflammation, and complications of other organ systems. Although risk scoring systems specific for IE surgery have recently been developed, their accuracy is somewhat low. In addition, information required for these methods (such as blood culture results) may not be easy to obtain in a short time, which limits their clinical application.[11,13,14] Therefore, there is a need for a user-friendly risk scoring system specifically for IE surgery.[15]

Machine learning is very useful for discovering hidden structures and patterns in intricate high-dimensional data and has been applied in many medical fields.[16–18] Machine learning algorithms are different from linear models, such as logistic regression, because they are computational methods that efficiently navigate the space of free parameters to arrive at a model with high predictive value.[19] In this study, we used machine learning to develop a simple and easy-to-use risk model to identify patients with IE who are at high risk of early mortality after surgery. The accuracy and robustness of the model were evaluated in separate cohorts, and the predictive value of the model was compared with that of commonly used scoring systems.

METHODS

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Data Sources and Extraction

This study was approved by the Institutional Review Board of The First Affiliated Hospital of Sun Yat-sen University (approval number: 2018[100]). Because of the use of de-labeled retrospective data, the requirement of informed consent from patients was waived.

Data of hospitalized patients who met the definite modified Duke criteria for IE, including native valve endocarditis and prosthetic valve endocarditis, were included in the present study as previously described.[20] Exclusion criteria were (1) did not have heart surgery, (2) pregnancy, (3) malignancy or life-threatening diseases, and (4) missing data of interest. Data extracted from electronic medical records were (1) patient demographic characteristics, (2) medical history; (3) clinical symptoms and signs, (4) laboratory and echocardiographic findings, (5) surgery-related information; and (6) discharge status and short-term (30 days after surgery) follow-up results. The EuroSCORE II score (http://www.euroscore.org/calc.html) and Charlson index[21] were calculated for each patient using online calculators. All data were reviewed, extracted, and cross-checked independently by 2 experienced clinicians. Disagreements between 2 reviewers were resolved by consultation with a third reviewer.

Patient data from 2 centers were used in this study. Data of patients with IE consecutively admitted to the First Affiliated Hospital of Sun Yat-sen University (FAH-SYSU) from October 2013 to August 2021 were categorized as the FAH-SYSU cohort, and the data were split into the training-validation cohort and the FAH-SYSU test cohort. Data of patients admitted from October 2013 through June 2019 were used as the

training-validation cohort for model development and selection, and data of patients admitted from July 2019 to August 2021 were used for internal testing (FAH-SYSU test cohort). Data of patients with IE consecutively admitted to Nanfang Hospital (NFH) of Southern Medical University from July 2018 to September 2021 were used for external testing (NFH cohort) (Figure 1).

Definitions

Early mortality was defined as all-cause mortality that occurred in hospital or within 30 days after surgery. Current heart failure was defined as New York Heart Association class III/IV at admission. Diastolic dysfunction and pulmonary hypertension were defined and graded according to echocardiography. Valve involvement referred to an infected valve(s). Unless otherwise stated, other variables and cutoffs were defined according to the *International Classification of Diseases, Ninth Revision, Clinical Modification* codes (*ICD-9-CM*). Potential predictive variables were patient characteristics at hospital admission and surgical data as indicated previously (Table 1 and Table 2).

Variable Selection and Model Construction

Potential predictors with more than 5% missing records were excluded before any analysis. For other variables, we used the mean of individual continuous variables to replace the missing records, and a negative result was used to replace the missing records of individual categorical variable. The point-biserial correlation coefficient was calculated for all continuous variables, and those with a 2-tailed value of $P<0.01$ were selected for further model development. To eliminate the adverse effects of potential multicollinearity within the clinical data, all categorical variables were input into a least absolute shrinkage and selection operator (LASSO) regression model for multivariate analysis. The penalty weight (lambda) was selected with minimum mean-square error based on 3-fold

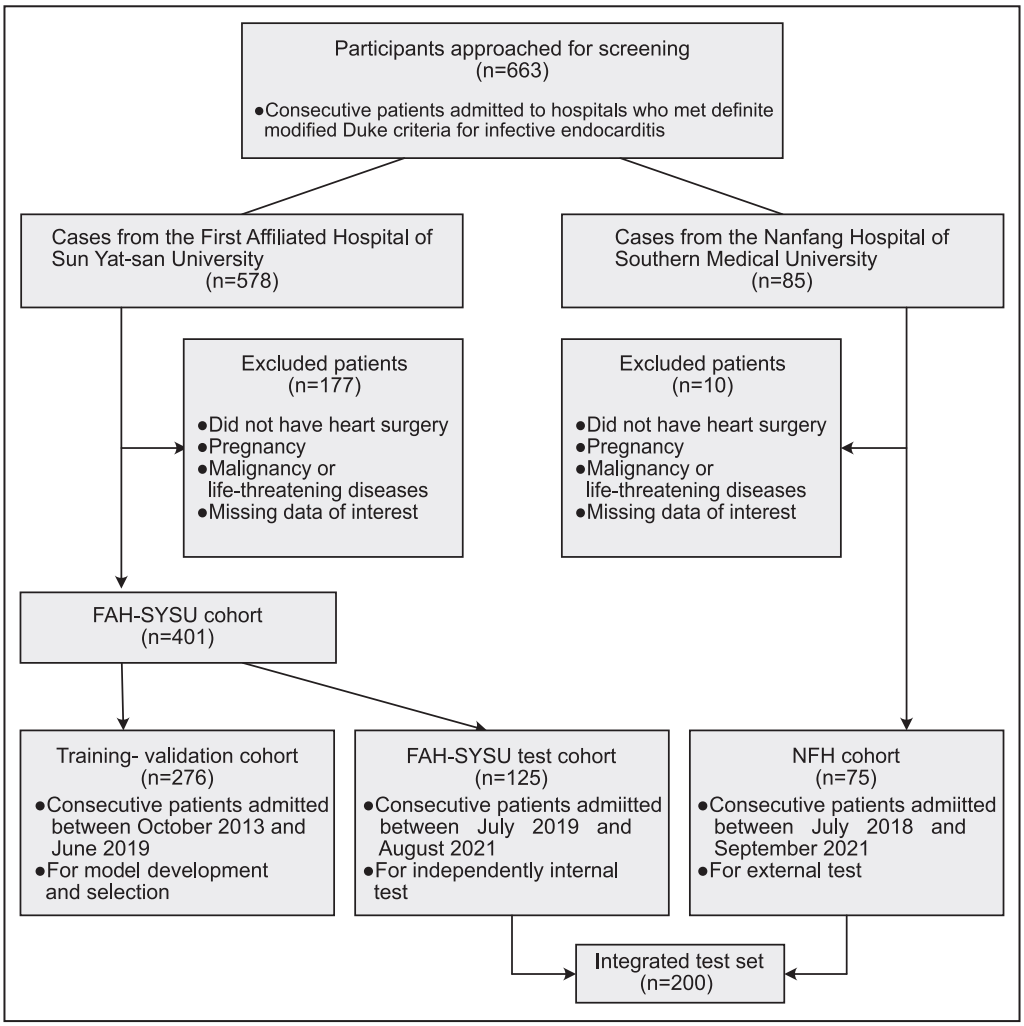


Figure 1. Flow chart of infective endocarditis patients.
FAH-SYSU indicates the First Affiliated Hospital of Sun Yat-sen University; and NFH, Nanfang Hospital.

Table 1. Demographics and Clinical Characteristics of Patients in the Training-Validation Cohort

Variable	No./No. (%)
Early mortality	20/276 (7.2)
Demographic characteristic	
Age, y, mean (SD) [range], y	43.70 (15.32) [13–84]
Age older than 60 y	50/276 (18.1)
Age older than 70 y	13/276 (4.7)
Female sex	82/276 (29.7)
Current smoker	37/276 (13.4)
Medical history	
Hypertension	40/276 (14.5)
Coronary heart disease	13/276 (4.7)
Diabetes	10/276 (3.6)
Chronic obstructive pulmonary disease	4/276 (1.4)
Preoperative continuous renal replacement therapy	3/276 (1.1)
Peripheral vascular disease	9/276 (3.3)
Congenital heart disease	53/276 (19.2)
Percutaneous coronary intervention history	2/276 (0.7)
Pacemaker/implantable cardioverter-defibrillator implantation	1/276 (0.4)
Preoperative cardiac arrest or cardiopulmonary resuscitation	4/276 (1.4)
Recurrence or previous infective endocarditis	1/276 (0.4)
Previous cardiac surgery	17/276 (6.2)
Prosthetic valve endocarditis	9/276 (3.3)
Preoperative mechanical ventilation or respiratory failure	6/276 (2.2)
Preoperative central neurological complications	63/276 (22.8)
Cerebral infarction within 3 mo	24/276 (8.7)
Preoperative embolization event	76/276 (27.5)
Moderate or severe anemia*	46/276 (16.7)
Renal insufficiency	17/276 (6.2)
Malnutrition†	
Moderate	35/276 (12.7)
Severe	8/276 (2.9)
Current heart failure‡	47/276 (17.0)
Symptom and sign	
Atrial fibrillation or atrial flutter	23/276 (8.3)
Cardiac conduction block	3/276 (1.1)
Surgery-related information	
Duration of preoperative treatment, mean (SD)[range], days	12.83 (8.53) [2–65]
Positive preoperative blood culture result	139/242 (57.4)
Biological valve replacement	61/262 (23.3)
Mechanical valve replacement	201/262 (76.7)
Type of operation	
Isolated valvuloplasty	14/276 (5.1)
Valve replacement	262/276 (94.9)
Valvuloplasty and valve replacement	66/276 (23.9)
Right-side heart surgery	73/276 (26.4)
Pulmonary vale	7/276 (2.5)
Valvuloplasty	3/276 (1.1)
Valve replacement	4/276 (1.4)
Tricuspid valve	66/276 (23.9)
Valvuloplasty	56/276 (20.3)
Valve replacement	10/276 (3.6)
Entire heart surgery§	20/276 (7.2)

*Moderate or severe anemia referred to hemoglobin <90 g/L.
†Malnutrition was diagnosed based on the serum albumin (ALB) level, moderate malnutrition referred to 25≤ALB <30 g/L, and severe malnutrition referred to ALB <25 g/L.
‡Current heart failure referred to New York Heart Association class III/IV at admission.
§Entire heart surgery referred to the situation in which the patient received both left- and right-side cardiac surgery.

cross-validation. With a larger penalty, the estimates of weaker predictors shrink toward zero and were thus excluded from the model. Variables remaining in the LASSO model were selected for further model development. All variables selected based on the training-validation cohort were used as input for XGBoost[22] model development. For hyperparameter selection, a 3-fold cross-validation and grid search was performed. The hyperparameters used for modeling are shown in Table S1. An online calculator based on the completed risk model was developed and published. A classical logistic regression model, and other machine learning models, including gradient boosting decision tree, light gradient boosting machine, random forest, and extra trees, were constructed using the same variables. The Python (version 3.7.6) package "scikit-learn" was used to perform LASSO regression, logistic regression, and gradient boosting decision tree, light gradient boosting machine, random forest, and extra trees modeling; the XGBoost (version 1.2.1) Python package was used to build the XGBoost model.

Statistical Analysis

Normally distributed continuous variables were expressed as mean (SD) [range], and nonnormally distributed continuous variables were expressed as median (interquartile range). The distribution of continuous variables was assessed using the Shapiro–Wilk normality test. Categorical variables were expressed as count and percentage. The comparison of continuous variables was performed using Student's t test and Mann-Whitney U test. The χ^2 test or Fisher's exact test was used for comparisons of categorical variables. A 2-tailed value of $P<0.01$ was considered statistically significant.

Table 2. Laboratory Findings and Echocardiographic Characteristics of Patients in the Training-Validation Cohort

Variable	Mean (SD) [range]
Early mortality, No./No. (%)	20/276 (7.2)
Laboratory finding	
White blood cell counts, $\times 10^9$/L	8.85 (3.26) [1.72–22.01]
Platelet count, $\times 10^9$/L	250.43 (93.55) [53–574]
Hemoglobin, g/L	112.99 (24.25) [58–228]
Hematocrit, %	34.31 (6.95) [20.70–71.70]
Red cell distribution width, %	15.38 (2.32) [11–25]
N-terminal pro-B-type natriuretic peptide, pg/mL	2914.37 (10860.59) [16.10–160559.00]
Serum creatinine*, μmol/L	91.59 (62.01) [36–602]
Renal insufficiency compensation phase, No./No. (%)	28/276 (10.1)
Renal insufficiency decompensated phase, No./No. (%)	15/276 (5.4)
Kidney failure phase, No./No. (%)	3/276 (1.1)
Uremia phase, No./No. (%)	0/276 (0.0)
Blood urea nitrogen, mmol/L	6.57 (4.55) [2.00–40.60]
Uric acid, μmol/L	379.23 (139.39) [112–798]
Serum albumin, g/L	36.18 (5.53) [20.9–50.2]
Total bilirubin, μmol/L	14.96 (9.79) [1.29–107.60]
Aspartate transaminase, U/L	28.01 (22.41) [4.50–254.00]
Fibrin, g/L	3.78 (1.27) [1.06–7.78]
Blood glucose†, mmol/L	5.04 (4.55) [1.4–13.4]
Mildly elevated, No./No. (%)	20/276 (7.2)
Moderately elevated, No./No. (%)	5/276 (1.8)
Severely elevated, No./No. (%)	2/276 (0.7)
Urine occult blood ≥ (++), No./No. (%)	118/265 (44.5)
Blood culture result, No./No. (%)	
Negative	103/242 (42.6)
Streptococcus viridans	44/242 (18.2)
Other *Streptococcus*	6/242 (2.5)
Staphylococcus aureus	15/242 (6.2)
Coagulase-negative *Staphylococcus*	2/242 (0.8)
Other bacteria	71/242 (29.3)
Fungus	1/242 (0.4)
Echocardiographic characteristic, No./No. (%)	
Valve involved‡	
Left heart	257/276 (93.1)
Aortic valve	141/276 (51.1)
Mitral valve	165/276 (59.8)
Aortic and mitral valve	48/276 (17.8)
Right heart	24/276 (8.7)
Pulmonary valve	8/276 (2.9)
Tricuspid valve	16/276 (5.8)
Single valve involved	221/276 (80.1)
Multiple valves involved	55/276 (19.9)
Vegetation formation	257/276 (93.1)
Size of vegetation§, mean (SD) [range], mm	11.20 (7.13) [0–45.00]
Larger than 10 mm	152/276 (55.1)
Larger than 15 mm	88/276 (31.9)
Larger than 20 mm	31/276 (11.2)
LVEF, mean (SD)[range], %	66.77 (8.44) [30–97]
LVEF<40%	2/276 (0.7)
Pulmonary artery pressure, mean (SD)[range]	40.52 (13.63) [16–86]
Moderate or higher pulmonary hypertension	53/276 (19.2)
Severe pulmonary hypertension	16/276 (5.8)
Diastolic dysfunction	174/276 (63.0)
Class I	121/276 (43.8)
Class II	41/276 (14.9)
Class III	12/276 (4.3)
Medium or more pericardial effusion	3/276 (1.1)
Heart abscess	29/276 (10.5)
Paravalvular abscess	23/276 (8.3)
Left atrial thrombus	2/276 (0.7)

LVEF indicates left ventricular ejection fraction.

*Patients with serum creatine levels of 133–177 μmol/L were considered at the renal insufficiency compensation phase, 178–442 μmol/L were considered at renal insufficiency decompensated phase, 443–707 μmol/L were considered at the kidney failure phase, and over 707 μmol/L were considered at the uremia phase.

†Fasting blood glucose >7 mmol/L was considered as mildly elevated, >8.4 mmol/L was considered as moderately elevated, and >11.1 mmol/L was considered as severely elevated.

‡Valve involvement referred to valve affected by infection.

§Size of vegetation referred to the largest diameter of vegetation measured by echocardiography.

The models were evaluated in 2 cohorts that were independent from the training-validation cohort (internal and external test). The area under the receiver operating curve (AUC) was used to assess the accuracy of the models, and the 95% CI of the AUC was calculated by 1000 bootstrap resamples. We evaluated the calibration of the XGBoost model and performed decision curve analysis[23] based on the integrated test set, which contained all the cases in the internal and external test sets. We also compared the accuracy of the ensemble learning models, the classic logistic regression model, and EuroSCORE II. Statistical analysis was performed with R software (version 4.1.2, R Foundation).

Feature Contribution Analysis

We used the permutation-based importance (function PermutationImportance) of the ELI5 Python package (https://eli5.readthedocs.io/) to report the feature contribution for the XGBoost model.

RESULTS

Baseline Characteristics of the Training-Validation Cohort

Data of 276 consecutive cases from the FAH-SYSU cohort formed the training-validation cohort, in which early mortality occurred in 20 patients (7.2%). The baseline characteristics of patients are summarized in Table 1 and Table 2. The mean (SD) age of the patients was 43.70 (15.32) years, and 82 patients (29.7%) were female. A total of 267 patients (96.7%) were diagnosed with native valve endocarditis, and 9 (3.3%) with prosthetic valve endocarditis. Fifty-five patients (19.9%) had multiple valves involvement. Seventy-three patients (26.4%) underwent right-side heart surgery. Preoperative embolization events occurred in 76 patients (27.5%) and were the most common complication. A total of 47 patients (17%) had current heart failure. Various degrees of diastolic dysfunction were found in more than half (174; 63%) of the patients. Urine occult blood (UOB)≥(++) was present in 44.5% of the patients. Blood culture results were available for 242 patients (87.7%) and were negative in 103 patients (42.6%). *Streptococcus viridans* was the most prevalent bacteria among patients with positive blood culture results.

Variable Selection and Model Construction

A total of 89 potential predictive variables were entered in the feature screening process. To avoid the adverse effects caused by the linear correlation between some continuous and categorical variables, they were analyzed separately. Among the 20 continuous variables, 3 variables were statistically significant predictors of early mortality after surgery for IE. These variables were platelet count (R=−0.21, $P<0.01$), serum albumin (R=−0.21, $P<0.01$), and Charlson score (R=0.20, $P<0.01$) (Table S2). Among the 69 categorical variables, 7 potential predictors were retained in the LASSO regression model. These variables included current heart failure, moderate or severe malnutrition, UOB≥(++), diastolic dysfunction, multiple valve involvement, tricuspid valve involvement, and vegetation >10 mm. The dynamic process of LASSO regression is shown in Figure S1.

The Charlson index is a composite score that covers multiple characteristics of patients. We discarded this variable from the final prediction model to ensure the model was simple and to avoid model distortion due to repeated inclusion of features. The diagnosis of malnutrition was based on serum albumin level. Considering that continuous variables may contain more information than categorical variables, the continuous variable (serum albumin) rather than the categorical variable (malnutrition) was included in the prediction model. Finally, 8 variables from the training-validation cohort were input into the XGBoost model for model training and validation: platelet count, serum albumin, current heart failure, UOB≥(++), diastolic dysfunction, multiple valve involvement, tricuspid valve involvement, and vegetation >10 mm.

After training, a machine learning–based risk model was constructed and named the Sun Yat-sen University Prediction Model for Infective Endocarditis (SYSUPMIE). We further developed an online calculator based on SYSUPMIE (http://42.240.140.58:1808/; Figure 2). By entering all values of the required 8 variables, users can easily calculate the probability of early mortality and obtain information on risk stratification.

Performance of SYSUPMIE

The mean AUC based on 5 times 3-fold cross-validation of SYSUPMIE in the training-validation cohort was 0.810±0.073. The results of the receiver-operator characteristic curve analysis of SYSUPMIE and other algorithms in the training-validation cohort are shown in Figure S2. The performance of SYSUPMIE was internally evaluated in the FAH-SYSU test cohort to illustrate the accuracy of the prediction model. The FAH-SYSU test cohort consisted of 125 consecutive patients with IE who underwent surgery at the FAH-SYSU from July 2019 to August 2021, and it was time independent from the training-validation cohort. There were 11 cases of early mortality (8.8%) in this cohort. The variables used for the internal test are shown in Table S3. The AUC of the model in the FAH-SYSU test cohort was 0.813 (95% CI, 0.670–0.933; Figure 3A), which indicated that the model had satisfactory discriminatory power.

External Test of SYSUPMIE

To evaluate the generalizability of SYSUPMIE, an external test was performed using the NFH cohort. A total of 75 consecutive patients with IE who underwent surgery at NFH of Southern Medical University from July 2018 to September 2021 comprised the NFH cohort. The mean (SD) [range] age of the patients was 46.67 (14.67) years (range, 14–70 years), 28 patients (37.3%) were female, 3 patients (4%) were diagnosed with prosthetic valve endocarditis, and 7 patients (9.3%) died within 30 days or in hospital after surgery. The variables used for external test are shown in Table S3. Variables, including vegetation >10 mm, multiple valves involved, diastolic dysfunction, and serum albumin, in the NFH cohort were significantly different from those in the FAH-SYSU test cohort ($P<0.05$; Figure S3). The performance of SYSUPMIE in the NFH cohort reached a predictive value similar to that of the FAH-SYSU validation cohort, with an AUC of 0.812 (95% CI, 0.606–0.956;

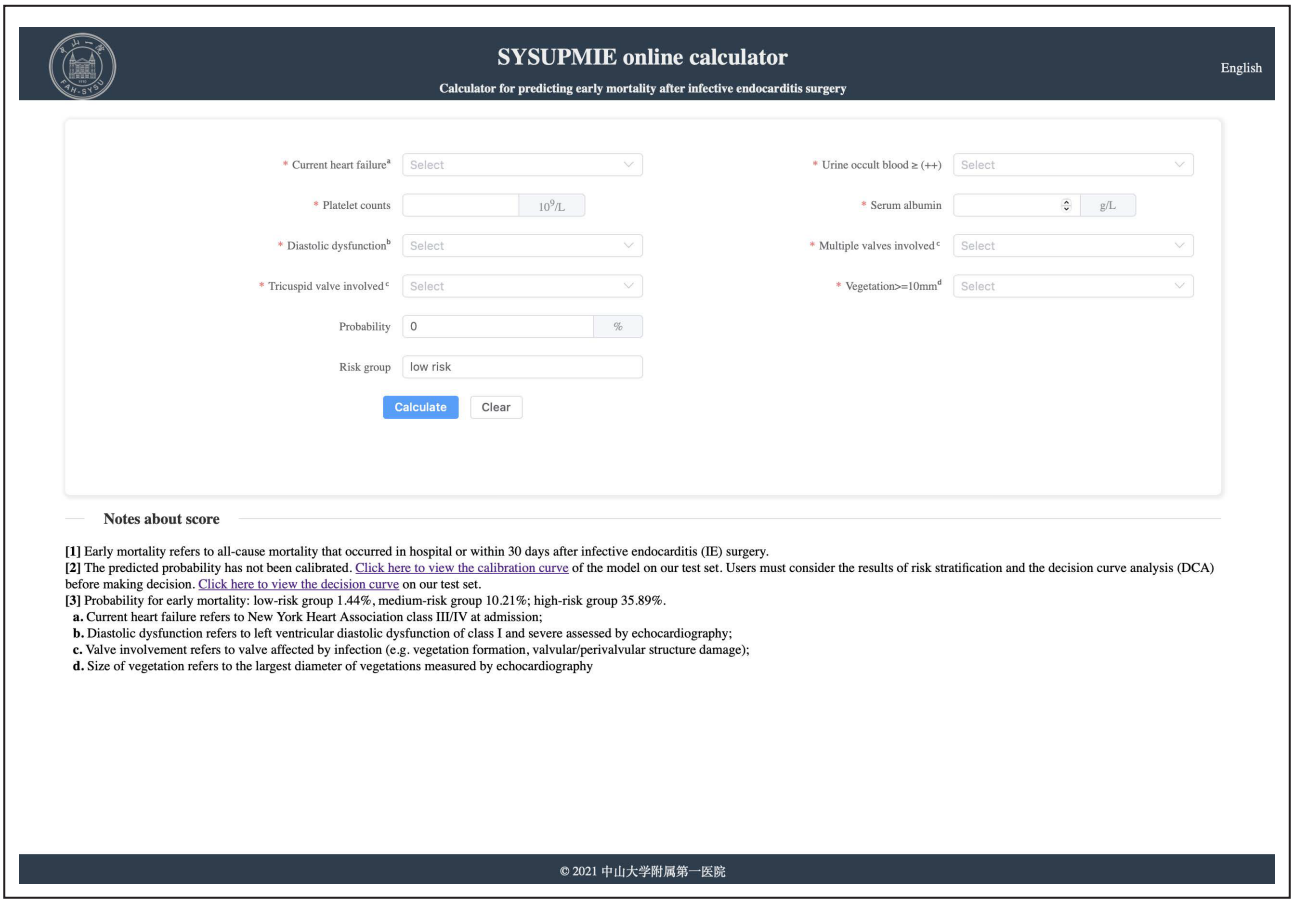


Figure 2. The online calculator* for predicting early mortality after infective endocarditis surgery.
*The Department of Cardiac Surgery of the First Affiliated Hospital of Sun Yat-sen University is responsible for the online calculator (http://42.240.140.58:1808). SYSUPMIE indicates Sun Yat-sen University Prediction Model for Infective Endocarditis.

Figure 3B), which proved that SYSUPMIE had good generalizability.

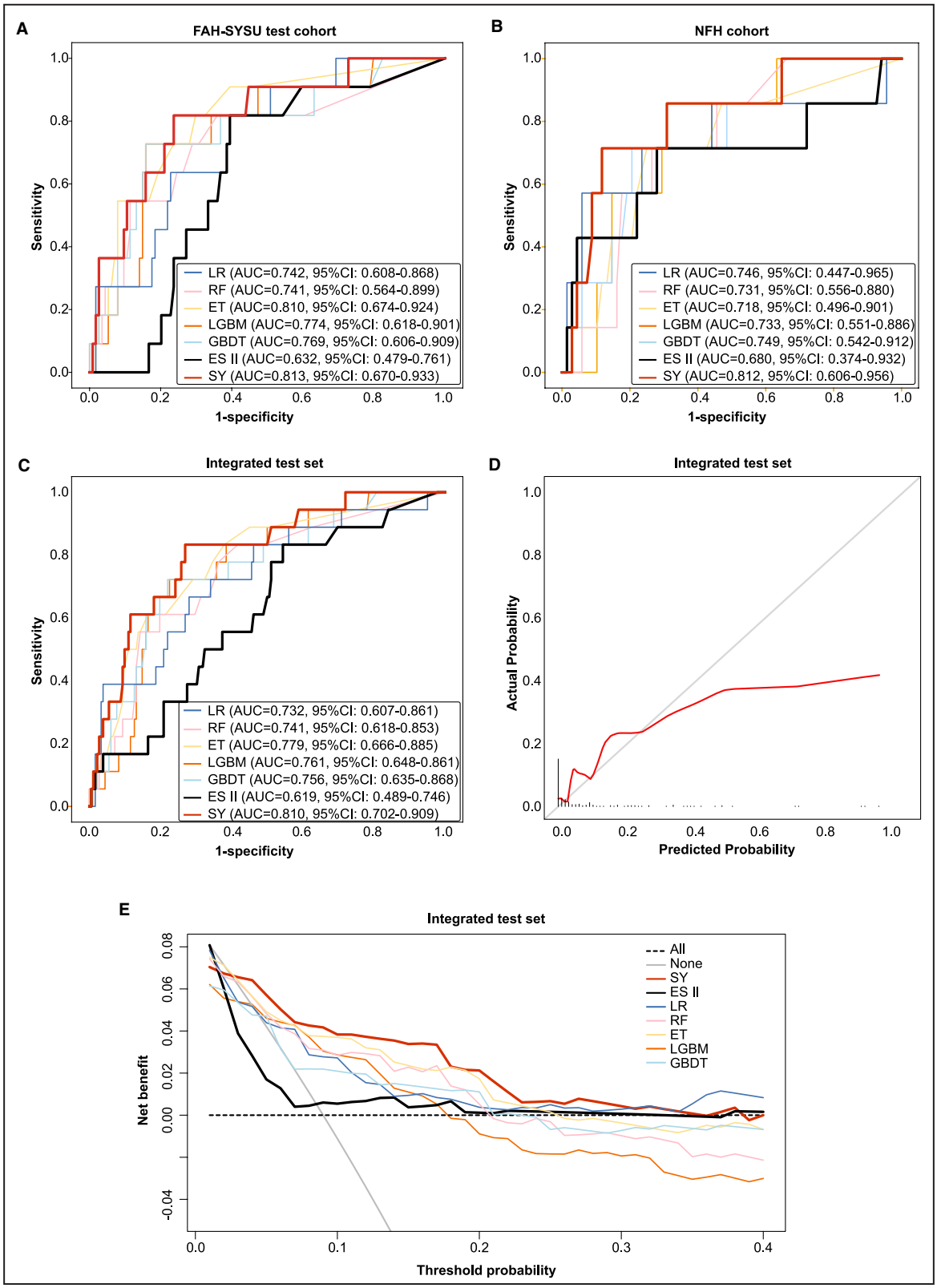
Comparisons Between Models

A comparison between the performance of SYSUPMIE and already established risk score models was carried out in the test cohorts. Receiver-operator characteristic curve analysis indicated that SYSUPMIE significantly outperformed EuroSCORE II in the FAH-SYSU test cohort (AUC, 0.813; 95% CI, 0.670–0.933 versus AUC, 0.632; 95% CI, 0.479–0.761, $P<0.01$), the NFH cohort (AUC, 0.812; 95% CI, 0.606–0.956 versus AUC, 0.680; 95% CI, 0.374–0.932, $P<0.01$), and the integrated test set (AUC, 0.810; 95% CI, 0.702–0.909 versus AUC, 0.619; 95% CI, 0.489–0.746, $P<0.01$) (Figure 3A through 3C).

To illustrate the superiority of XGBoost algorithm, the 8 variables were used to build a classical logistic regression model and ensemble learning models, including gradient boosting decision tree, light gradient boosting machine, random forest, and extra trees, based on the training-validation cohort. The predictive value of those models was significantly lower than that of SYSUPMIE (all, $P<0.01$) but higher than that of EuroSCORE II (all, $P<0.01$) (Figure 3A through 3C).

Decision curve analysis based on the integrated test set indicated that SYSUPMIE had a better clinical application value than EuroSCORE II, other ensemble learning models, and the classical logistic regression model within a probability threshold ranging from 0.02 to 0.40. Notably, we found that in the data sets from both centers EuroSCORE II had proved to be of minimal value with respect to guiding clinical decision-making (Figure 3E).

A small number of missing data in the training set was imputed in the preprocessing stage (Table S4). To demonstrate the impact of imputation strategy on model performance, a comparison of performance between SYSUPMIE and an XGBoost model trained using unimputed data was performed. The performance of SYSUPMIE, which was trained using imputed data, was slightly better than that of the XGBoost model, which was trained using unimputed data in the in the FAH-SYSU test cohort (AUC, 0.813; 95% CI, 0.670–0.933 versus AUC, 0.803; 95% CI, 0.652–0.928, $P<0.001$), the NFH cohort (AUC, 0.812; 95%



CI, 0.606–0.956 versus AUC, 0.796; 95% CI, 0.583–0.954, *P*<0.001), and the integrated test set (AUC, 0.810; 95% CI, 0.702–0.909 versus AUC, 0.795; 95% CI, 0.680–0.902, *P*<0.001) (Figure S4).

Risk Stratification

The calibration of the new model was further evaluated based on the integrated test set. The results showed that SYSUPMIE might overestimate the probability

Figure 3. Model performance in the test sets.
A through **C**, ROC analysis of Sun Yat-sen University Prediction Model for Infective Endocarditis (SYSUPMIE), classic logistic regression model, random forest model, extra trees model, LightGBM model, GBDT model and EuroSCORE II for predicting early mortality after surgery for IE in the internal test set of the First Affiliated Hospital of Sun Yat-sen University, in the external test set of Nanfang Hospital, and in the integrated test set. **D**, Calibration curve of SYSUPMIE in the integrated test set. **E**, Decision curve analysis (DCA) of SYSUPMIE, classic logistic regression, random forest model, extra trees model, LightGBM model, GBDT model and EuroSCORE II based on the integrated test set. AUC indicates area under the curve; EuroSCORE II, European System for Cardiac Operative Risk Evaluation II; ET, extra trees; FAH-SYSU, First Affiliated Hospital of Sun Yat-sen University; GBDT, gradient boosting decision trees; LGBM, light gradient boosting machine; LR, logistic regression; RF, random forest; ROC, receiver-operator characteristic curve; and SY, Sun Yat-sen University Prediction Model for Infective Endocarditis.

of early mortality when the predicted probability was >30% (Figure 3D). Therefore, a risk stratification system was developed in order to make SYSUPMIE more clinically applicable. Risk cutoff values of 0.055 and 0.35 were selected at 95% sensitivity and 95% specificity, respectively, according to 3-fold cross-validation of the training-validation cohort (Figure S5). Patients were divided into low-, medium-, and high-risk groups. The integrated test set was used to evaluate these risk cutoff values. The confusion matrices demonstrated that the risk stratification system possessed satisfactory discriminatory power for low-, medium-, and high-risk patients (Figure 4A). The probability of early mortality after surgery for low-risk patients was 1.44% (95% CI, 0–0.044), for medium-risk patients was 10.21% (95% CI, 0.051–0.162), and for high-risk patients was 35.89% (95% CI, 0.071–0.571) (Figure 4B).

Feature Contribution Analysis

The completed machine learning–based risk model, SYSUPMIE, which was a "black box" model, integrated 8 preoperative features to predict the risk of early mortality. Comparison with the classical logistic regression model indicated that there was not simply a linear correlation between these features and the risk of mortality. Although there was a degree of correlation between each feature and the predicted risk (Figure 4C), it was apparent that the model did not completely rely on the strength of the correlations when making predictions. Thus, we quantified how much these various features contributed to explaining patient-to-patient variation in risk (Figure 4D). The results showed that serum albumin levels had the greatest effect, followed by platelet count and current heart failure. Patients with hypoalbuminemia, a decreased platelet count, and current heart failure were more likely to have a higher risk of early mortality (Figure 4C and 4D).

DISCUSSION

The early mortality rate of surgery for IE remains the greatest among all valve operations.[24] Risk scores are useful tools to aid decision-making and guide perioperative management, and an accurate and scientific risk assessment plays a vital role in the management of IE and may affect the outcome of treatment. In this study, we built a machine learning–based prediction model, SYSUPMIE, for predicting early mortality after surgery for IE. A comprehensive evaluation of the model (SYSUPMIE) was performed with internal and external validation using independent cohorts. The evaluation showed that SYSUPMIE exhibited satisfactory accuracy and generalizability. We further developed an open-access online calculator based on SYSUPMIE. By providing the values of 8 variables, users can easily and quickly calculate the probability of early mortality after surgery for IE and obtain information for risk stratification of patients with IE.

EuroSCORE II and STS score are recommended by clinical guidelines and the most used scoring systems, but they are neither accurate nor easy to use. Several previous studies[10,12] have found that EuroSCORE II performed unsatisfactorily in patients with IE. The usability of STS score is also limited because it is applicable to only 7 fixed procedure types, and it requires more than 40 items. We attempted to retrospectively calculate STS scores for the patients include in this study but stopped because of mismatched procedures and not being able to determine all the variables required. Although EuroSCORE II and STS score perform well in cardiac diseases that do not involve infection or vegetation formation on heart valves, their accuracy with respect to IE is very low. This is likely because the models were not specifically developed for IE, and IE patients accounted for only a very small portion of the development cohorts and thus the effects of characteristics that are directly associated with early mortality after surgery for IE were diluted. Taking EuroSCORE II as an example, there were only 497 patients with active endocarditis in its development cohort, which accounted for only 2.2% of all cases (497/22381).[25]

STS-IE (2010) is an IE-specific risk score that was developed using the largest series available of IE treated surgically. STS-IE uses 13 variables, which greatly simplifies calculation of STS score and avoids being applicable only to fixed surgical procedures. However, the development of STS-IE score used medical records from 2002 to 2008. As guidelines have been updated since then,[26] it is not clear whether STS-IE is still useful. Moreover, STS-IE has a C-index of 0.75784, which is relatively low and needs to be improved. Although

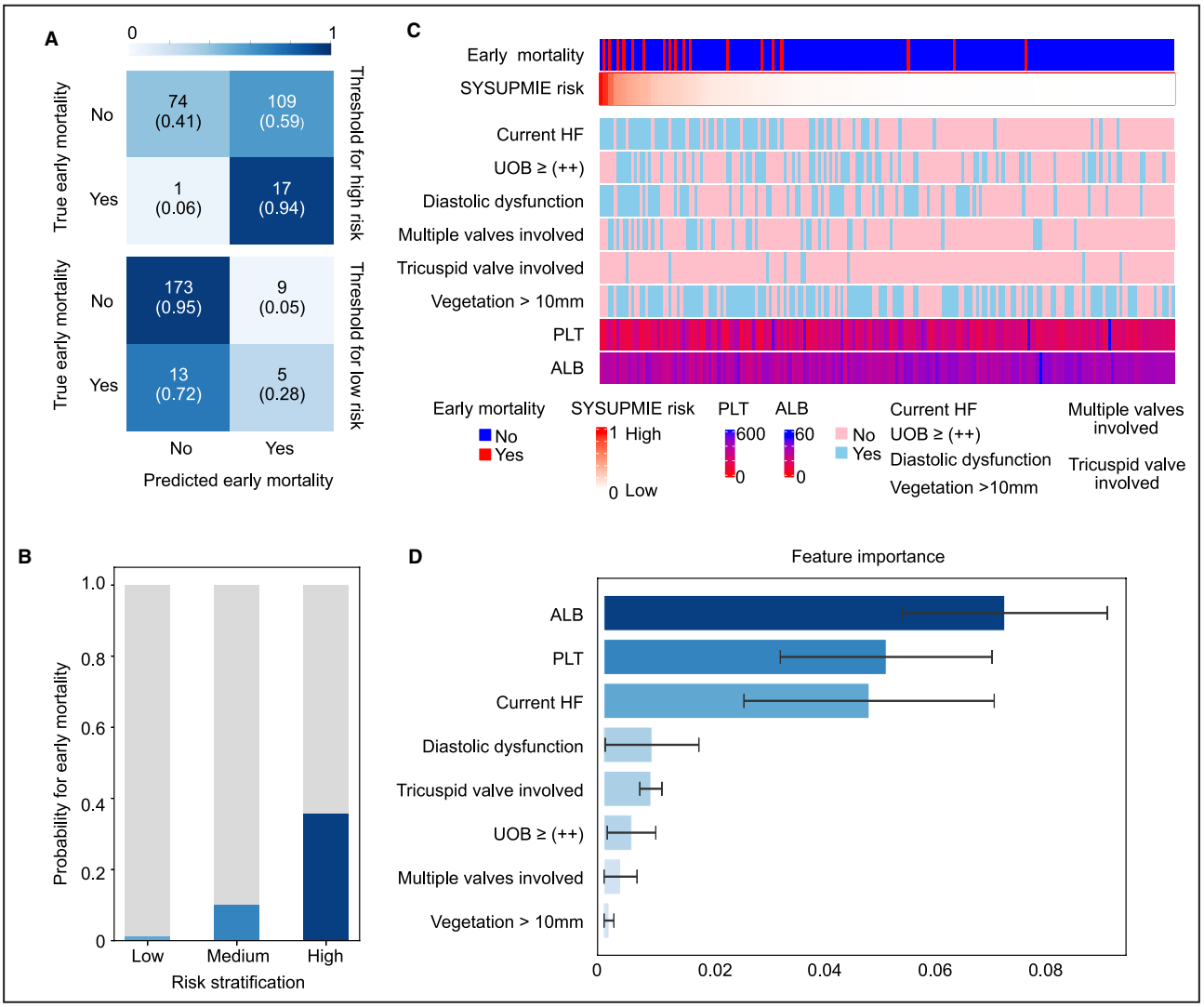


Figure 4. Risk stratification and feature contribution analysis for SYSUPMIE.
A, Confusion matrices at cutoffs of 0.35 and 0.055 for the high- and low-risk groups, based on the integrated test set. **B**, The actual probability of early mortality in low-, medium-, and high-risk groups, based on the integrated test set. **C**, Relevance of SYSUPMIE risk to early mortality and features. **D**, Feature contribution analysis, based on the integrated test set. The error bars denote SD of feature contribution. ALB indicates serum albumin; HF, heart failure; PLT, platelet count; SYSUPMIE, Sun Yat-sen University Prediction Model for Infective Endocarditis; and UOB, urine occult blood.

risk score systems that were developed after STS-IE, such as Defeo score (2012),[15] PALSUSE (2014),[13] AEPEI (2017),[27] and RISK-E (2017)[14] provide higher accuracy and greater value, all have some limitations.

The SYSUPMIE model does not have many of the limitations of prior systems. First, we did not include composite scores as in PALSUSE, in which "logistic EuroSCORE ≥ 10" is a predictor. Including composite scores put the model at risk of being too complex and increases the chances of overfitting because of the introduction of unknown multicollinearity. Second, we did not make a fine distinction between subtypes of IE, as do Defeo score and RISK-E. Because infection can be heterogeneous, a finer distinction will narrow the clinical value of score. The SYSUPMIE model incudes left- and right-side involvement, native valve endocarditis and prosthetic valve endocarditis, and acute and nonacute IE; thus, the model is applicable to various clinical scenarios. Finally, SYSUPMIE is practical. It does not require variables that have a high likelihood of being inaccurate or cannot be obtained in a short time, such as blood culture results, which are required for RISK-E, Defeo, and PALSUSE. Additionally, the online calculator provides quick access to the model using readily obtainable data that have a low likelihood of being inaccurate. In our development cohort, blood culture results were not available for about 12% of the patients. Furthermore, the negative result in 103 patients was based on only a single 3-day blood culture and thus has a high likelihood of being inaccurate.

Indeed, it usually takes 3 to 7 days to obtain the reliable blood culture results, and most patients might have received empirical antibiotics or completed surgical procedures during this time frame.

The selection of predictors is a necessary compromise of feasibility and effectiveness. Although previous studies have demonstrated the potential value of several variables (eg, high-sensitivity C-reactive protein, red cell distribution width, high-sensitivity troponin T, creatinine, and N-terminal of the prohormone brain natriuretic peptide) for predicting early mortality after cardiac surgery,[28] introducing every conceivable risk factor does not necessarily lead to increased performance of a model.[29] Of note, age was not a variable used in SYSUPMIE. Undoubtedly, older age increases the risk of cardiac surgery and aging-associated frailty syndrome has been shown to be significantly correlated with poor outcomes after cardiac surgery.[30] Furthermore, age is used in most cardiac surgery risk score systems.[13,15,25,31,32] However, overall patients with IE are younger than most patients requiring cardiac surgery.[33,34] In our development cohort, although the mean age of survivors was slightly less than that of nonsurvivors (43.36±15.119 versus 48.05±17.509 years), there was no significant difference (P=0.1877) between the 2 groups (Figure S6). Patients with perivalvular complications in our cohorts did not have a higher risk, although perivalvular complications were once considered to be a high-risk factor.[31] Thus, it was also not included in SYSUPMIE. In fact, whether perivalvular abscesses increase the perioperative mortality rate is inconclusive. Another prospective study[35] found that although paravalvular abscesses increase the need for surgical treatment to correct hemodynamic disorders and clear the infection, they do not increase the mortality rate.

To our knowledge, SYSUPMIE is the first risk model that uses UOB, diastolic dysfunction, and tricuspid valve involvement as predictors. Positive UOB results may be an indication of occult renal impairment caused by infection, immune response, or embolism. UOB≥(++) was very common in our IE patients, with an incidence as high as 40% (186/465). Previous studies have found a strong association between postoperative mortality and diastolic dysfunction, as it can further aggravate the circulatory disturbance caused by heart failure and may cause postoperative atrial fibrillation.[36–38] Contrary to the results of other studies that suggest the prognosis of right-sided IE is better than that of left-sided IE,[39,40] we found that tricuspid valve involvement was a risk factor for early mortality. Although the effect of tricuspid valve dysfunction on hemodynamics and systemic embolism is smaller than that of mitral and aortic valve dysfunction, right-side IE is associated with a high recurrence rate, especially in intravenous drug users, and its impact on mortality should not be ignored.[26]

The robust performance of SYSUPMIE, which is superior to that of classical logistic regression models, shows that machine learning algorithms, especially gradient boosting algorithms like XGBoost, are a good approach to this kind of classification task using structured, clinical tabular data[41–43] and can be used when the number of cases is limited. Traditional logistic regression modeling has higher requirements with respect to the scale of data, which limits its application when the amount of available data is small and when events have a low incidence. Moreover, methods have been developed to derive information about the importance of individual variables for a specific prediction, which makes it possible to reveal how machine learning models leverage patient characteristics for risk prediction. Different from using the coefficients of variables to explain how variables affect the prediction in a linear model, this kind of interpretability can even yield medical insights.[41]

Feature contribution analysis indicated that hypoalbuminemia and thrombocytopenia were strong predictors of early mortality after surgery for IE. Hypoalbuminemia reflects systemic inflammation, impaired antioxidant function, poor nutrient reserves, and potential kidney damage.[20] Thrombocytopenia reflects the severity of underlying sepsis and an elevated risks of hemorrhage.[44] Both are indicators of disease severity and the nature of the host response. Our results emphasize the possible risks of antiplatelet therapy in IE, as has been shown in previous studies.[26] Although many IE patients have indications for antithrombotic therapy, particularly those with mechanical prosthetic valves and coronary stents, the potential risks and benefits of antithrombotic therapy must be carefully weighed. Note that both serum albumin level and platelet count reflected pretreatment status, as the data were from admission before the patient received any interventions. Thus, using SYSUPMIE to dynamically evaluate patients receiving exogenous albumin supplementation and platelet transfusion therapy may result in an inaccurate risk estimate, and neither can the model indicate an effect of treatment when used in this manner.

This study had several limitations. Limited by the size of the study population, SYSUPMIE was not calibrated based on our data because this would be unreasonable. Clinicians must be very cautious in understanding the probability given by the model, and should comprehensively consider risk stratification and decision curve analysis before any decision-making. All data were from southern China, which could potentially limit the generalizability of SYSUPMIE in other areas of the world. Bias inherent to observational studies still cannot be avoided because our study used retrospectively collected data. Larger, multicenter external data are needed to further validate and calibrate the model.

In addition, no model is future proof as epidemiology changes and medical treatments evolve.

CONCLUSIONS

In this study, we developed and validated a machine learning–based prediction model (SYSUPMIE) to predict early mortality after surgery for IE. The model was shown to be accurate and robust in identifying patients with IE at high risk of early mortality after surgery and thus may help clinical decision-making and improve outcomes.

ARTICLE INFORMATION

Received January 17, 2022; accepted April 25, 2022.

Affiliations

Department of Cardiac Surgery, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, P. R. China (L.L., S.-Q H., Q.L., Y.Y., K.-z.L., L.H., S.-j.W., H.-y.L., Z.-k.W.); School of Computer Science and Technology, Xidian University, Xi'an, P. R. China (C.L.); Department of General Surgery, Qianfoshan Hospital, Shandong University, Jinan, P. R. China (S.D.); and Department of Cardiovascular Surgery, Nanfang Hospital, Southern Medical University, Guangzhou, P. R. China (S.Z.).

Acknowledgments

We gratefully thank all the colleagues and the participants in the study. Prof. Z.W. and S.Z. administrate and directed the project. L.L. conceived the goals and designed the methodology for the study. Q.L., K.L., S.H., H.L., L.H. and S.W. collected medical records and extracted and cross-checked the data. L.L., C.L. conducted the development and evaluation of the XGBoost model. C.L. developed the online calculator. C.L., S.D., L.L. and S.H. performed statistical analysis and visualized the results. L.L. wrote the article with the critical revisions proposed by S.H., and Y.Y.

Sources of Funding

This work was supported by the National Science Foundation of China (Grant numbers 81770319 and 82070297) and the Natural Science Funds of Guangdong Province (Grant number 2019A1515010218).

Disclosures

None.

Supplemental Material

Tables S1–S4
Figures S1–S6

REFERENCES

1. Pant S, Patel NJ, Deshmukh A, Golwala H, Patel N, Badheka A, Hirsch GA, Mehta JL. Trends in infective endocarditis incidence, microbiology, and valve replacement in the United States from 2000 to 2011. *J Am Coll Cardiol*. 2015;65:2070–2076. doi: 10.1016/j.jacc.2015.03.518
2. Liu S, Li Y, Zeng X, Wang H, Yin P, Wang L, Liu Y, Liu J, Qi J, Ran S, et al. Burden of cardiovascular diseases in China, 1990–2016: findings from the 2016 Global Burden of Disease Study. *JAMA Cardiol*. 2019;4:342–352. doi: 10.1001/jamacardio.2019.0295
3. Kyu HH, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, Abbastabar H, Abd-Allah F, Abdela J, Abdelalim A, Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2018;392:1859–1922. doi: 10.1016/S0140-6736(18)32335-3
4. Cahill TJ, Baddour LM, Habib G, Hoen B, Salaun E, Pettersson GB, Schäfers HJ, Prendergast BD. Challenges in infective endocarditis. *J Am Coll Cardiol*. 2017;69:325–344. doi: 10.1016/j.jacc.2016.10.066
5. Manne MB, Shrestha NK, Lytle BW, Nowicki ER, Blackstone E, Gordon SM, Pettersson G, Fraser TG. Outcomes after surgical treatment of native and prosthetic valve infective endocarditis. *Ann Thorac Surg*. 2012;93:489–493. doi: 10.1016/j.athoracsur.2011.10.063
6. Chu VH, Park LP, Athan E, Delahaye F, Freiberger T, Lamas C, Miro JM, Mudrick DW, Strahilevitz J, Tribouilloy C, et al. Association between surgical indications, operative risk, and clinical outcome in infective endocarditis: a prospective study from the International Collaboration on Endocarditis. *Circulation*. 2015;131:131–140. doi: 10.1161/CIRCULATIONAHA.114.012461
7. Vallejo Camazón N, Cediel G, Núñez Aragón R, Mateu L, Llibre C, Sopena N, Gual F, Ferrer E, Quesada MD, Berastegui E, et al. Short- and long-term mortality in patients with left-sided infective endocarditis not undergoing surgery despite indication. *Rev Esp Cardiol (Engl Ed)*. 2020;73:734–740. doi: 10.1016/j.rec.2019.09.011
8. Said SM, Abdelsattar ZM, Schaff HV, Greason KL, Daly RC, Pochettino A, Joyce LD, Dearani JA. Outcomes of surgery for infective endocarditis: a single-centre experience of 801 patients. *Eur J Cardiothorac Surg*. 2018;53:435–439. doi: 10.1093/ejcts/ezx341
9. Sevilla T, López J, Gómez I, Vilacosta I, Sarriá C, García-Granja PE, Olmos C, Di Stefano S, Maroto L, San Román JA. Evolution of prognosis in left-sided infective endocarditis: a propensity score analysis of 2 decades. *J Am Coll Cardiol*. 2017;69:111–112. doi: 10.1016/j.jacc.2016.10.052
10. Patrat-Delon S, Rouxel A, Gacouin A, Revest M, Flécher E, Fouquet O, Le Tulzo Y, Lerolle N, Tattevin P, Tadié J-M. EuroSCORE II underestimates mortality after cardiac surgery for infective endocarditis. *Eur J Cardiothorac Surg*. 2016;49:944–951. doi: 10.1093/ejcts/ezv223
11. Wang TKM, Wang MTM, Pemberton J. Risk scores and surgery for infective endocarditis: a meta-analysis. *Int J Cardiol*. 2016;222:1001–1002. doi: 10.1016/j.ijcard.2016.08.054
12. Grant SW, Hickey GL, Dimarakis I, Cooper G, Jenkins DP, Uppal R, Buchan I, Bridgewater B. Performance of the EuroSCORE models in emergency cardiac surgery. *Circ Cardiovasc Qual Outcomes*. 2013;6:178–185. doi: 10.1161/CIRCOUTCOMES.111.000018
13. Martínez-Sellés M, Muñoz P, Arnáiz A, Moreno M, Gálvez J, Rodríguez-Roda J, de Alarcón A, García Cabrera E, Fariñas MC, Miró JM, et al. Valve surgery in active infective endocarditis: a simple score to predict in-hospital prognosis. *Int J Cardiol*. 2014;175:133–137. doi: 10.1016/j.ijcard.2014.04.266
14. Olmos C, Vilacosta I, Habib G, Maroto L, Fernández C, López J, Sarriá C, Salaun E, Di Stefano S, Carnero M, et al. Risk score for cardiac surgery in active left-sided infective endocarditis. *Heart*. 2017;103:1435–1442. doi: 10.1136/heartjnl-2016-311093
15. De Feo M, Cotrufo M, Carozza A, De Santo LS, Amendolara F, Giordano S, Della Ratta EE, Nappi G, Della CA. The need for a specific risk prediction system in native valve infective endocarditis surgery. *Sci World J*. 2012;2012:1–8. doi: 10.1100/2012/307571
16. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521:436–444. doi: 10.1038/nature14539
17. Delen D, Walker G, Kadam A. Predicting breast cancer survivability: a comparison of three data mining methods. *Artif Intell Med*. 2005;34:113–127. doi: 10.1016/j.artmed.2004.07.002
18. Shi J-Y, Wang X, Ding G-Y, Dong Z, Han J, Guan Z, Ma L-J, Zheng Y, Zhang L, Yu G-Z, et al. Exploring prognostic indicators in the pathological images of hepatocellular carcinoma based on deep learning. *Gut*. 2021;70:951–961. doi: 10.1136/gutjnl-2020-320930
19. Deo RC. Machine learning in medicine. *Circulation*. 2015;132:1920–1930. doi: 10.1161/CIRCULATIONAHA.115.001593
20. Huang S, Zhou Z, Luo LI, Yue Y, Liu Q, Feng K, Hou J, Wang K, Chen J, Li H, et al. Preoperative serum albumin: a promising indicator of early mortality after surgery for infective endocarditis. *Ann Transl Med*. 2021;9:1445. doi: 10.21037/atm-21-3913
21. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40:373–383. doi: 10.1016/0021-9681(87)90171-8
22. Chen T, Guestrin C. Xgboost: A scalable tree boosting system. Paper/Poster presented at: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining; 2016.
23. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. 2006;26:565–574. doi: 10.1177/0272989X06295361
24. Aksoy O, Sexton DJ, Wang A, Pappas PA, Kourany W, Chu V, Fowler VG, Woods CW, Engemann JJ, Corey GR, et al. Early surgery in patients

with infective endocarditis: a propensity score analysis. *Clin Infect Dis*. 2007;44:364–372. doi: 10.1086/510583
25. Nashef SAM, Roques F, Sharples LD, Nilsson J, Smith C, Goldstone AR, Lockowandt U, EuroSCORE II. *J Cardiothorac Surg*. 2012;41. doi: 10.1093/ejcts/ezs043
26. Habib G, Lancellotti P, Antunes MJ, Bongiorni MG, Casalta J-P, Del Zotti F, Dulgheru R, El Khoury G, Erba PA, Iung B, et al. 2015 ESC guidelines for the management of infective endocarditis: the task force for the management of infective endocarditis of the European Society of Cardiology (ESC). Endorsed by: European Association for Cardio-Thoracic Surgery (EACTS), the European Association of Nuclear Medicine (EANM). *Eur Heart J*. 2015;36:3075–3128. doi: 10.1093/eurheartj/ehv319
27. Gatti G, Perrotti A, Obadia J-F, Duval X, Iung B, Alla F, Chirouze C, Selton-Suty C, Hoen B, Sinagra G, et al. Simple scoring system to predict in-hospital mortality after surgery for infective endocarditis. *J Am Heart Assoc*. 2017;6. doi: 10.1161/JAHA.116.004806
28. Duchnowski P, Hryniewiecki T, Kuśmierczyk M, Szymański P. The usefulness of selected biomarkers in patients with valve disease. *Biomark Med*. 2018;12:1341–1346. doi: 10.2217/bmm-2018-0101
29. Li B, Zhou H, He J, Wang M, Yang Y, Li L. On the sentence embeddings from pre-trained language models. *arXiv preprint arXiv:201105864*. 2020.
30. Duchnowski P, Szymański P, Kuśmierczyk M, Hryniewiecki T. Usefulness of FRAIL scale in heart valve diseases. *Clin Interv Aging*. 2020;15:1071–1075. doi: 10.2147/CIA.S239054
31. Jakuska P, Ereminiene E, Muliuolyte E, Kosys V, Pavlavičius L, Zukovas G, Karciauskas D, Benetis R. Predictors of early mortality after surgical treatment of infective endocarditis: a single-center experience. *Perfusion*. 2020;35:290–296. doi: 10.1177/0267659119872345
32. Gaca JG, Sheng S, Daneshmand MA, O'Brien S, Rankin JS, Brennan JM, Hughes GC, Glower DD, Gammie JS, Smith PK. Outcomes for endocarditis surgery in North America: a simplified risk scoring system. *J Thorac Cardiovasc Surg*. 2011;141:98–106. doi: 10.1016/j.jtcvs.2010.09.016
33. Ikama MS, Nkalla-Lambi M, Kimbally-Kaky G, Loumouamou ML, Nkoua JL. Profile of infective endocarditis at Brazzaville university hospital. *Med Sante Trop*. 2013;23:89–92. doi: 10.1684/mst.2013.0151
34. Letaief A, Boughzala E, Kaabia N, Ernez S, Abid F, Chaabane TB, Jemaa MB, Boujnah R, Chakroun M, Daoud M, et al. Epidemiology of infective endocarditis in Tunisia: a 10-year multicenter retrospective study. *Int J Infect Dis*. 2007;11:430–433. doi: 10.1016/j.ijid.2006.10.006
35. Graupner C, Vilacosta I, SanRomán J, Ronderos R, Sarriá C, Fernández C, Mújica R, Sanz O, Sanmartín JV, Pinto AG. Periannular extension of infective endocarditis. *J Am Coll Cardiol*. 2002;39:1204–1211. doi: 10.1016/S0735-1097(02)01747-3
36. Kaw R, Hernandez AV, Pasupuleti V, Deshpande A, Nagarajan V, Bueno H, Coleman CI, Ioannidis JPA, Bhatt DL, Blackstone EH. Effect of diastolic dysfunction on postoperative outcomes after cardiovascular surgery: a systematic review and meta-analysis. *J Thorac Cardiovasc Surg*. 2016;152:1142–1153. doi: 10.1016/j.jtcvs.2016.05.057
37. Melduni RM, Suri RM, Seward JB, Bailey KR, Ammash NM, Oh JK, Schaff HV, Gersh BJ. Diastolic dysfunction in patients undergoing cardiac surgery: a pathophysiological mechanism underlying the initiation of new-onset post-operative atrial fibrillation. *J Am Coll Cardiol*. 2011;58:953–961. doi: 10.1016/j.jacc.2011.05.021
38. Sastry P, Theologou T, Field M, Shaw M, Pullan DM, Fabri BM. Predictive accuracy of EuroSCORE: is end-diastolic dysfunction a missing variable? *Eur J Cardiothorac Surg*. 2010;37:261–266. doi: 10.1016/j.ejcts.2009.05.059
39. Stavi V, Brandstaetter E, Sagy I, Sapunar S, Nevzorov R, Bartal C, Barski L. Comparison of clinical characteristics and prognosis in patients with right- and left-sided infective endocarditis. *Rambam Maimonides Med J*. 2019;10:e0003. doi: 10.5041/RMMJ.10338
40. Akinosoglou K, Apostolakis E, Marangos M, Pasvol G. Native valve right sided infective endocarditis. *Eur J Intern Med*. 2013;24:510–519. doi: 10.1016/j.ejim.2013.01.010
41. Rousset A, Dellamonica D, Menuet R, Lira Pineda A, Sabatine MS, Giugliano RP, Trichelair P, Zaslavskiy M, Ricci L. Can machine learning bring cardiovascular risk assessment to the next level? A methodological study using FOURIER trial data. *Eur Heart J - Digital Health*. 2021. doi: 10.1093/ehjdh/ztab093
42. Desai RJ, Wang SV, Vaduganathan M, Evers T, Schneeweiss S. Comparison of machine learning methods with traditional models for use of administrative claims with electronic medical records to predict heart failure outcomes. *JAMA Netw Open*. 2020;3:e1918962. doi: 10.1001/jamanetworkopen.2019.18962
43. Akyea RK, Qureshi N, Kai J, Weng SF. Performance and clinical utility of supervised machine-learning approaches in detecting familial hypercholesterolaemia in primary care. *NPJ Digit Med*. 2020;3:142. doi: 10.1038/s41746-020-00349-5
44. Ferreira FL, Bota DP, Bross A, Mélot C, Vincent JL. Serial evaluation of the SOFA score to predict outcome in critically ill patients. *JAMA*. 2001;286:1754–1758. doi: 10.1001/jama.286.14.1754

SUPPLEMENTAL MATERIAL

Table S1. Hyperparameters used for XGBoost modeling

Hyperparameter	value
n_estimators	160
eta	0.1
max_depth	5
gamma	0
min_child_weight	2
subsample	0.8
colsample_bytree	0.8
scale_pos_weight	1

Table S2. Univariate analysis of continuous variables in training-validation cohort

Variable	R*	*p*
Age	0.7954	0.1877
White blood cell counts	-0.2021	0.8546
Hemoglobin	-0.1154	0.0555
Hematocrit	-0.1068	0.0766
Platelet counts	**-0.2087**	**0.0005**
Fibrin	0.0067	0.9117
NT-proBNP	0.0535	0.3755
Blood sugar	0.0177	0.7693
Blood urea nitrogen	0.0731	0.2261
Serum creatinine	0.1441	0.0166
Uric acid	0.0321	0.5950
Serum albumin	**-0.2111**	**0.0004**
Aspartate transaminase	0.0134	0.8247
Total bilirubin	0.0528	0.3823
Pulmonary artery pressure	0.0920	0.1272
LVEF	-0.0206	0.7329
Size of vegetation	0.0674	0.2641
Duration of preoperative treatment	-0.0346	0.5670
RDW	0.0806	0.1817
Charlson score	**0.2028**	**0.0007**

* Point-biserial correlation coefficient

Abbreviations: NT-proBNP, N-terminal pro-B-type natriuretic peptide; LVEF, Left ventricular ejection fraction; RDW, Red cell distribution width.

Table S3. Laboratory findings and clinical characteristics of patients in two test sets

Variable, No./N0. (%)	FAH-SYSU test cohort	NFH cohort
Early mortality	11/125 (8.8)	7/75 (9.3)
Current heart failure	37/125 (29.6)	32/75 (42.7)
Urine occult blood ≥ (++)	45/125 (36.0)	23/75 (30.7)
Diastolic dysfunction	53/125 (42.4)	11/75 (14.7)
Multiple valves involvement	8//125 (6.4)	27/75 (36.0)
Tricuspid valve involvement	5/125 (4.0)	4/75 (5.3)
Vegetation ≥ 10mm	64/125 (51.2)	54/75 (72.0)
Serum albumin, mean (SD) [range], g/L	35.07 (4.14) [25.7-44.7]	32.68 (6.08) [20.6-50.8]
Platelet counts, mean (SD) [range], 10^9/L	245.50 (103.31) [52-554]	235.33 (106.08) [46-598]

Abbreviation: SD, Standard deviation.

Table S4. Missing data in the training-validation cohort*.

Variable	Number of missing data No. (%)	Imputation strategy
Positive blood culture result	34 (12.32%)	Exclude the variable†
Pathogen type	34 (12.32%)	Exclude the variable†
Urine Occult Blood ≥ (++)	11 (3.99%)	'0'
NT-proBNP	7 (2.54%)	mean
Uric acid	7 (2.54%)	mean

* No missing data in the test cohorts.

† This variable was excluded prior to any analysis.

Abbreviation: NT-proBNP, N-terminal pro-B-type natriuretic peptide.

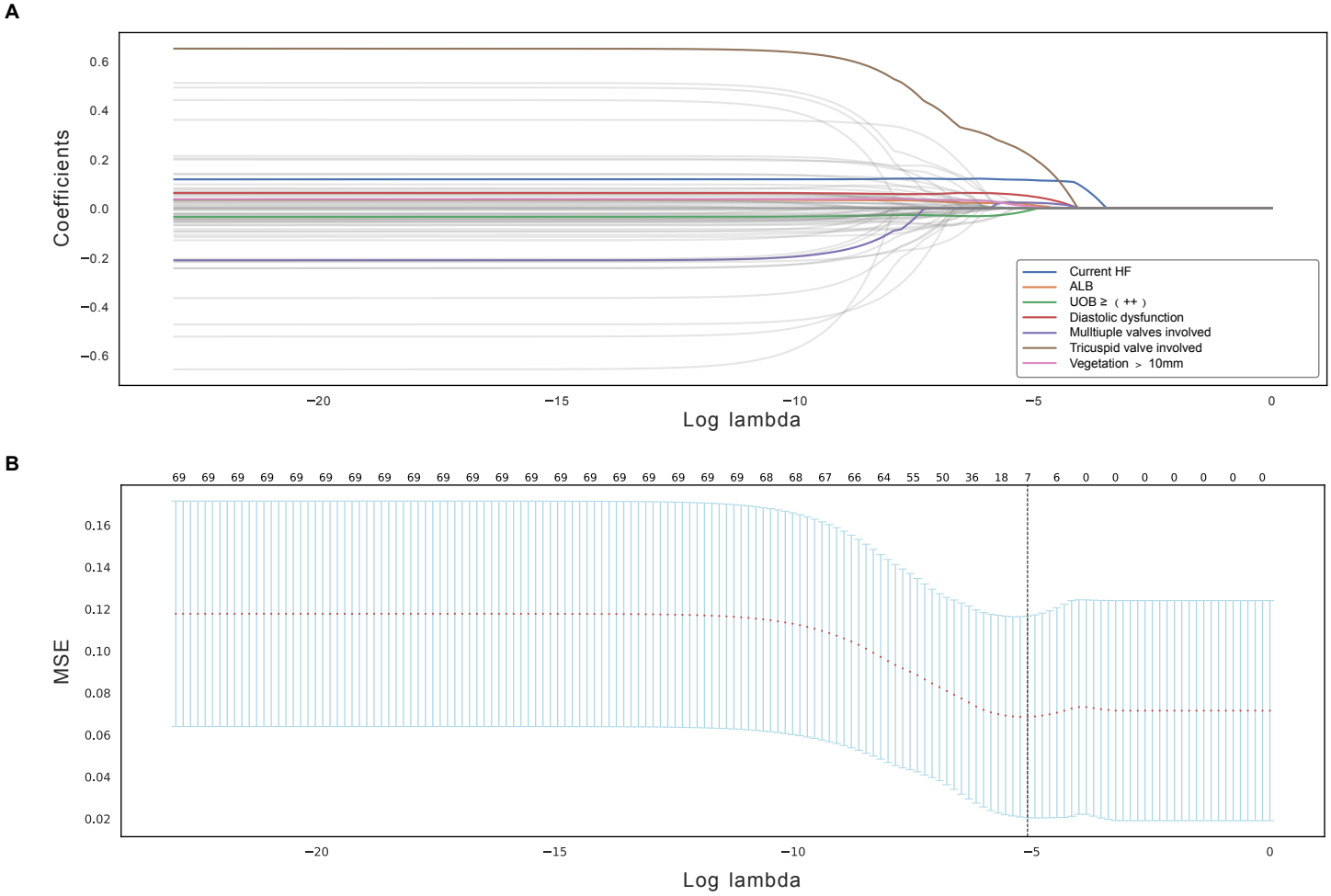


Figure S1. Feature selection process using the least absolute shrinkage and selection operator (LASSO) model

(A) LASSO coefficient profiles of the 69 categorical potential predictors. **(B)** Tuning parameter (λ) selection in the LASSO model used 3-fold cross-validation via minimum mean-squared error.

Abbreviation: HF, heart failure; ALB, serum albumin; UOB, urine occult blood.

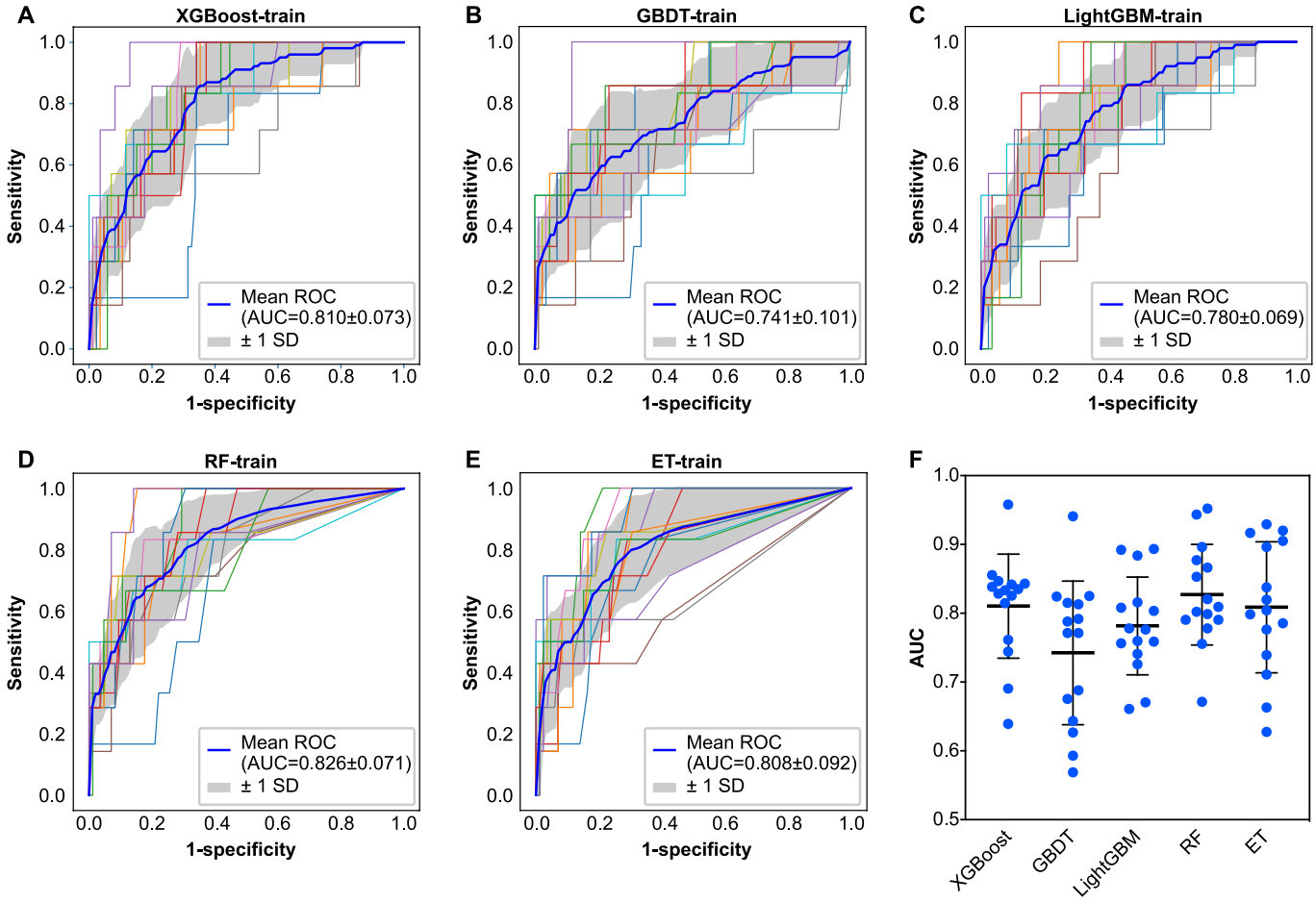


Figure S2. ROC analysis of different algorithms in the training-validation cohort for predicting early mortality after surgery for IE

(A) ROC analysis of XGBoost model (SYSUPMIE) in the training-validation cohort. **(B)** ROC analysis of GBDT model in the training-validation cohort. **(C)** ROC analysis of LightGBM model in the training-validation cohort. **(D)** ROC analysis of Random Forest model in the training-validation cohort. **(E)** ROC analysis of Extra Trees model in the training-validation cohort. The mean AUCs were the mean of the AUCs in 15 folds of validation sets in the 5 times 3-fold cross-validation which were denoted by the thinner colored lines. **(F)** AUCs of XGBoost model (SYSUPMIE), GBDT model, LightGBM model, Random Forest model, Extra Trees model in the training-validation cohort. The error bars denote standard deviation of AUCs.

Abbreviations: XGBoost, eXtreme Gradient Boosting; GBDT, gradient boosting decision trees; LightGBM, light gradient boosting machine; RF, random forest; ET, extra trees; ROC, receiver-operator characteristic curve; AUC, area under the curve; SD, standard deviation.

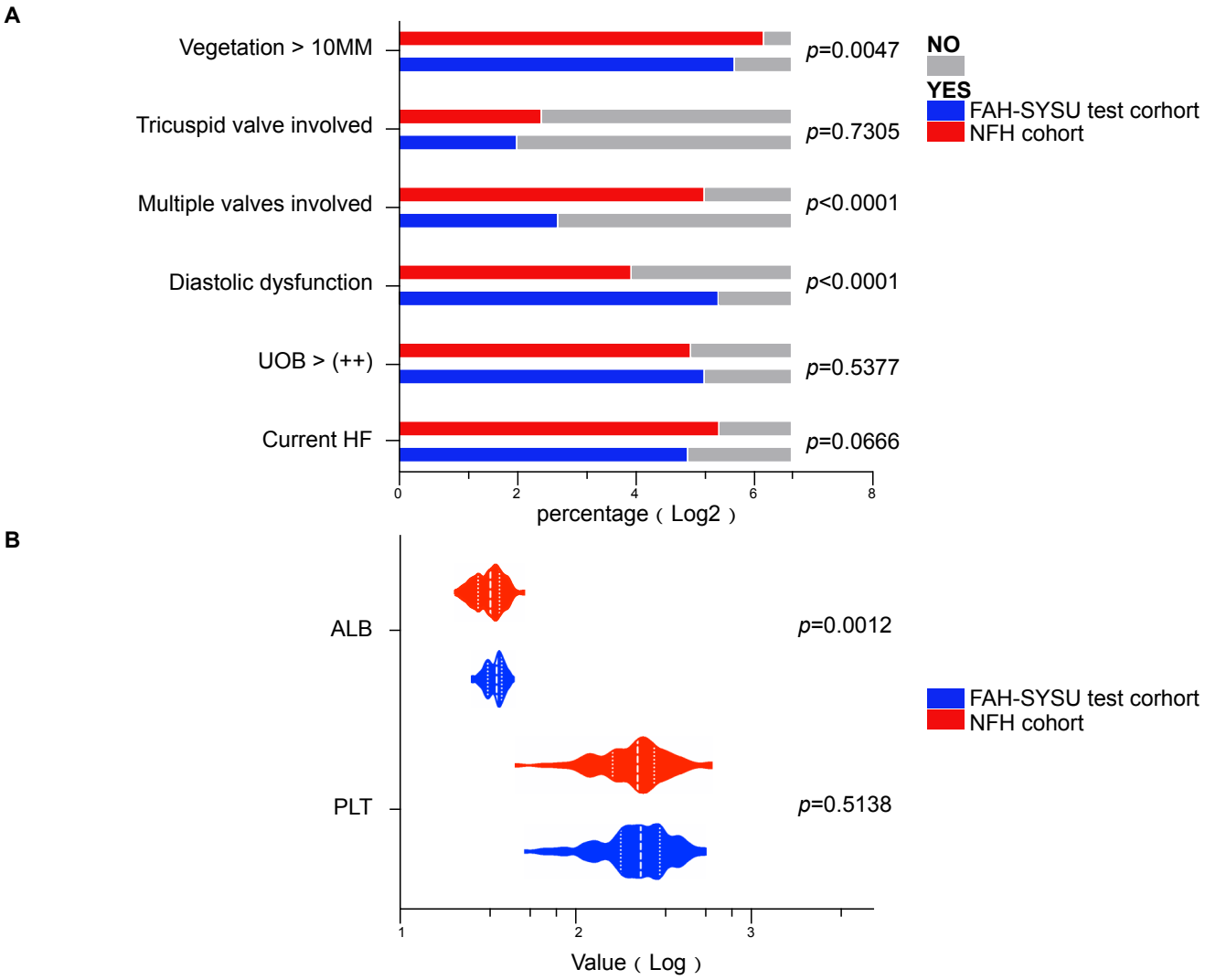


Figure S3. Comparison of features composited two independent test sets

(A) Comparison of categorical variables composited the FAH-SYSU test cohort and the NFH cohort. Two-tailed *p*-value corresponds to the results of χ^2 test. **(B)** Comparison of continuous variables composited the FAH-SYSU test cohort and the NFH cohort. Two-tailed *p*-value corresponds to the results of Student's t test.

Abbreviations: HF, heart failure; ALB, serum albumin; UOB, urine occult blood; PLT, platelet count; FAH-SYSU, the First Affiliated Hospital of Sun Yat-sen University; NFH, Nanfang Hospital.

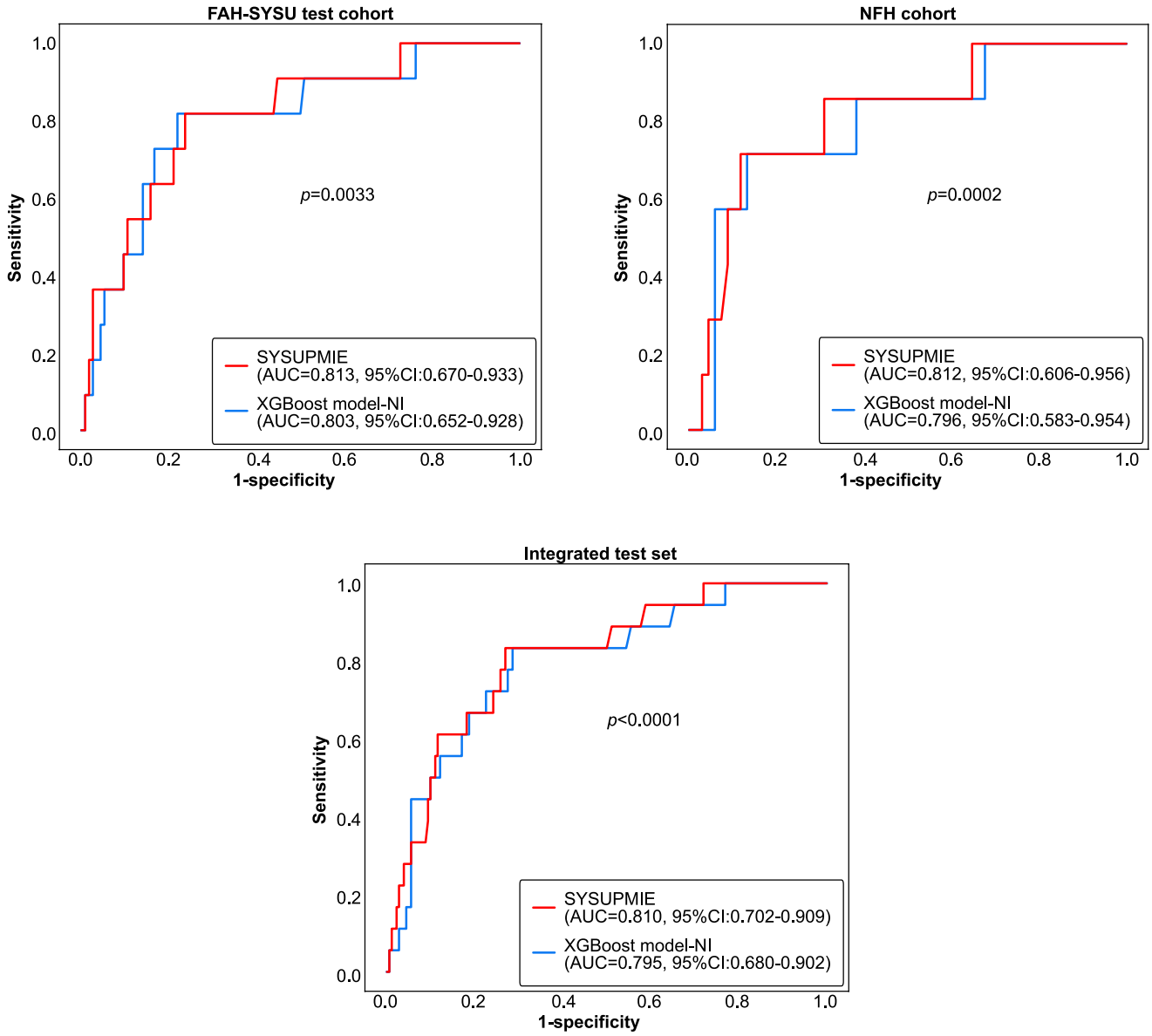


Figure S4. Comparison of model performance between models trained using imputed and unimputed data

XGBoost model-NI denote XGBoost model trained using unimputed data; two-tailed *p*-value corresponds to the results of Mann-Whitney U test of AUCs calculated by 1,000 bootstrap resamples.

Abbreviations: AUC, area under the curve; FAH-SYSU, the First Affiliated Hospital of Sun Yat-sen University; NFH, Nanfang Hospital; SYSUPMIE, Sun Yat-sen University prediction model for infective endocarditis.

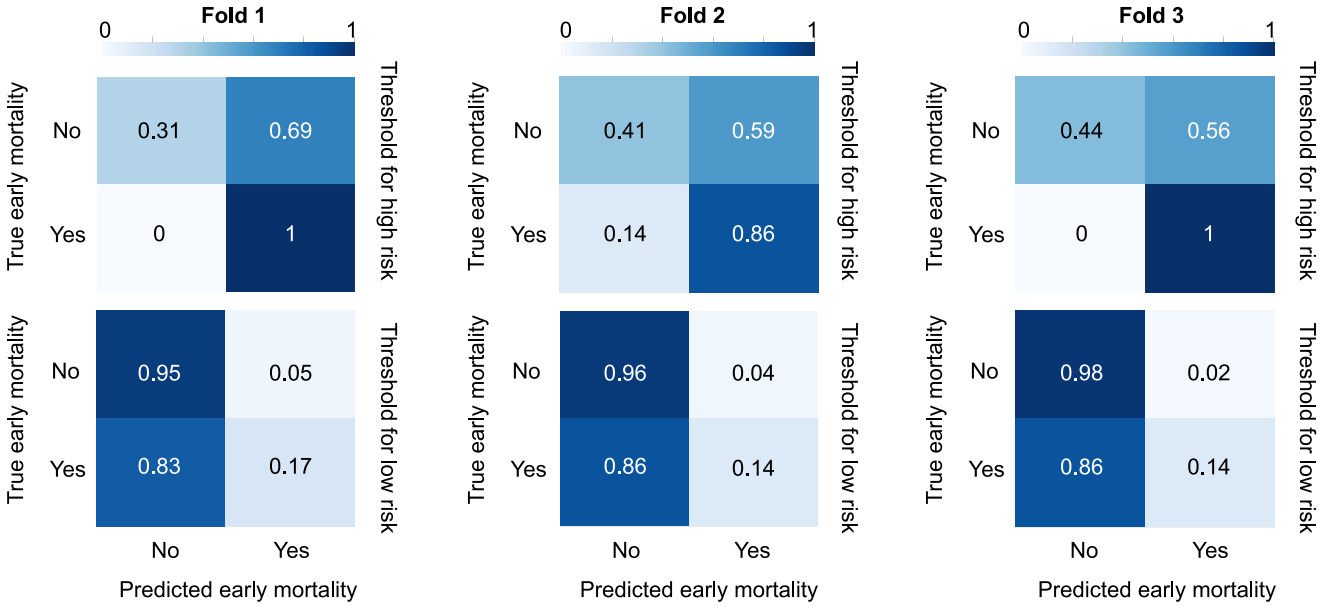


Figure S5. Cut-offs selection at 95% sensitivity and 95% specificity in the training-validation cohort via 3-fold cross validation

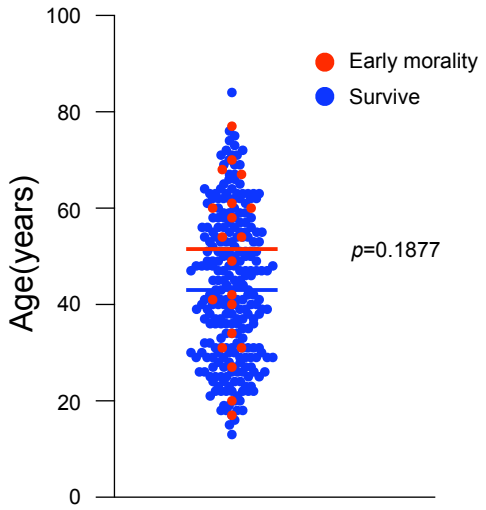


Figure S6. Comparison of age between survivor and non-survivor groups in the developing cohort.

Two-tailed *p*-value corresponds to the results of Student's t test.