Supplementary

Table S1 Detailed parameter settings of 5 ML model

ML model	Detailed parameter settings
Logistic regression	C-index: 1, max_iter: 100, penalty: l2, tol: 0.0001
Gaussian NB	Priors: None, var_smoothing: 1e-07
Multilayer perceptron	Activation: logistic, hidden_layer_sizes: (60, 10), max_iter: 200
Support vector machine	C-index: 1.0, kernel: rbf, tol: 0.001
K-nearest neighbor	n-neighbors: 6, weights: uniform

ML, machine learning.

Table S2 Detailed feature importance by LASSO regression

Feature	Value
SPO_2	0.084
CPB-time	0.063
Gender	0.038
ACC-time	0.026
EF	0.021
Weight	0.010
PV-leaflet	0.004
Z-index	<0.001
LVEDI	≤0.001
DP	0.078
TP repair	0.057
Age	0.038
Annulus	0.024
PvO	0.013
VSD	0.007
HCT	0.002
M-index	≤0.001
Branch	≤0.001

CPB, cardiopulmonary bypass; ACC, aortic cross-clamp; EF, ejection fraction; PV-leaflet, pulmonary valve leaflet; Z-index, pulmonary valve annulus Z score; LVEDI, left ventricular end-diastolic volume index; RVOTDP or DP, differential pressure of the right ventricular outflow tract; TP repair, transannular patch repair; PvO, opening of the pulmonary valve; VSD, ventricular septal defect; HCT, haematocrit; M-index, McGoon index.

https://dx.doi.org/10.21037/tp-22-246

Table S3 Details of results in the training sets

Factors	AUC	Accuracy	Sensitivity	Specificity	Positive predictive value	Negative predictive value	F1-SCORE
LR mean	0.760	0.708	0.657	0.760	0.684	0.729	0.668
LR SD	0.010	0.012	0.038	0.048	0.036	0.019	0.008
GNB mean	0.730	0.667	0.733	0.630	0.614	0.750	0.658
GNB SD	0.012	0.023	0.113	0.117	0.058	0.048	0.030
MLP mean	0.513	0.536	0.691	0.435	0.518	0.642	0.541
MLP SD	0.069	0.064	0.302	0.336	0.081	0.115	0.106
SVM mean	0.624	0.637	0.602	0.664	0.615	0.703	0.607
SVM SD	0.155	0.032	0.305	0.179	0.074	0.065	0.124
KNN mean	0.779	0.665	0.660	0.793	0.825	0.638	0.728
KNN SD	0.035	0.009	0.068	0.022	0.067	0.002	0.034

AUC, the area under the curve; LR, logistic regression; SD, standard deviation; GNB, Gaussian Naive Bayes; MLP, multilayer perceptron; SVM, support vector machine; KNN, K-nearest neighbour.

Table S4 Details of results in the testing sets

Factors	AUC	Accuracy	Sensitivity	Specificity	Positive predictive value	Negative predictive value	F1-SCORE
LR mean	0.701	0.664	0.612	0.741	0.626	0.705	0.598
LR SD	0.062	0.033	0.180	0.134	0.124	0.055	0.105
GNB mean	0.707	0.664	0.671	0.724	0.674	0.732	0.628
GNB SD	0.012	0.022	0.204	0.202	0.173	0.114	0.044
MLP mean	0.450	0.536	0.792	0.355	0.456	0.689	0.574
MLP SD	0.065	0.040	0.199	0.215	0.034	0.073	0.083
SVM mean	0.607	0.586	0.692	0.597	0.615	0.748	0.641
SVM SD	0.180	0.066	0.368	0.219	0.117	0.164	0.153
KNN mean	0.523	0.545	0.314	0.795	0.523	0.549	0.352
KNN SD	0.139	0.020	0.341	0.262	0.125	0.040	0.114

AUC, the area under the curve; LR, logistic regression; SD, standard deviation; GNB, Gaussian Naive Bayes; MLP, multilayer perceptron; SVM, support vector machine; KNN, K-nearest neighbour.

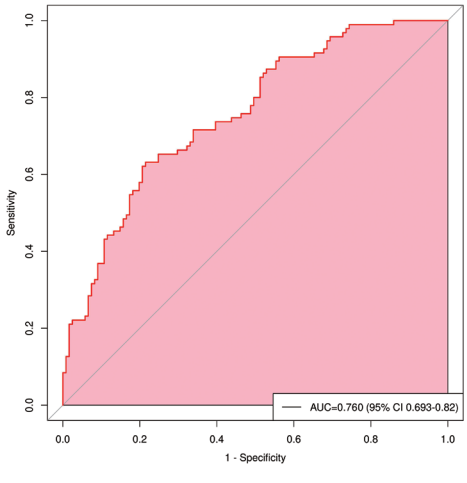

Figure S1 AUC of the LR model. AUC, the area under the curve; CI, confidence interval; LR, logistic regression.

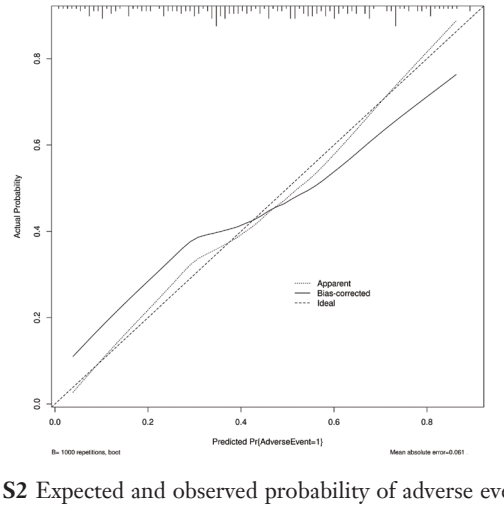

Figure S2 Expected and observed probability of adverse events by the Hosmer-Lemeshow test.

https://dx.doi.org/10.21037/tp-22-246