BMC Medical Informatics and
Decision Making

RESEARCH

Machine learning model predicts airway stenosis requiring clinical intervention in patients after lung transplantation: a retrospective case-controlled study

Dong Tian[1,2*†], Yu-Jie Zuo[1,3†], Hao-Ji Yan[4†], Heng Huang[1], Ming-Zhao Liu[2], Hang Yang[2], Jin Zhao[2], Ling-Zhi Shi[2*] and Jing-Yu Chen[2*]

Abstract

Background Patients with airway stenosis (AS) are associated with considerable morbidity and mortality after lung transplantation (LTx). This study aims to develop and validate machine learning (ML) models to predict AS requiring clinical intervention in patients after LTx.

Methods Patients who underwent LTx between January 2017 and December 2019 were reviewed. The conventional logistic regression (LR) model was fitted by the independent risk factors which were determined by multivariate LR. The optimal ML model was determined based on 7 feature selection methods and 8 ML algorithms. Model performance was assessed by the area under the curve (AUC) and brier score, which were internally validated by the bootstrap method.

Results A total of 381 LTx patients were included, and 40 (10.5%) patients developed AS. Multivariate analysis indicated that male, pulmonary arterial hypertension, and postoperative 6-min walking test were significantly associated with AS (all $P < 0.001$). The conventional LR model showed performance with an AUC of 0.689 and brier score of 0.091. In total, 56 ML models were developed and the optimal ML model was the model fitted using a random forest algorithm with a determination coefficient feature selection method. The optimal model exhibited the highest AUC and brier score values of 0.760 (95% confidence interval [CI], 0.666–0.864) and 0.085 (95% CI, 0.058–0.117) among all ML models, which was superior to the conventional LR model.

Conclusions The optimal ML model, which was developed by clinical characteristics, allows for the satisfactory prediction of AS in patients after LTx.

Keywords Airway stenosis, Lung transplantation, Machine learning, Logistic regression, Prediction model

---

†Dong Tian, Yu-Jie Zuo and Hao-Ji Yan contributed equally to this work.

*Correspondence:
Dong Tian
22tiandong@163.com
Ling-Zhi Shi
shilingzhi1979@126.com
Jing-Yu Chen
chenjy@wuxiph.com
Full list of author information is available at the end of the article

Tian *et al. BMC Medical Informatics and Decision Making*   (2024) 24:229

Page 2 of 13

Introduction

Lung transplantation (LTx) has been considered the only effective therapeutic option for end-stage lung diseases. The number of lung transplants has been increasing over the last two decades, with approximately 70,000 adult lung transplants performed worldwide thus far [1]. Since the first clinical LTx in 1963, airway complications (AC) have resulted in substantial mortality and clinical LTx failure [2]. In recent years, the occurrence of AC has tended to decrease with improvements in surgical techniques, immunosuppression, and patient allocation [3]. Nevertheless, large studies have reported that the prevalence of AC remains high.

Airway stenosis (AS) refers to a fixed reduction in the caliber of the airway and is the most common AC after LTx with a reported occurrence rate ranging from 1.6%–32.0% in previous studies [4–9]. The onset of AS usually occurs between 2 and 9 months after LTx [10, 11]. A reduction in the cross-sectional area > 50% is confirmation of severe AS, which reduces the quality of life and increases the morbidity and mortality of patients [12]. Severe AS requires timely clinical intervention to prevent further progression of AS [13]. Early detection of AS and treatment by balloon dilation can achieve good efficacy [14]. However, the early stages of AS are difficult to detect since they often present without specific clinical symptoms. Bronchoscopy is the gold standard for diagnosis, but it is usually used in patients who present with clinical symptoms [15]. Therefore, early and accurate detection of AS requiring clinical intervention is crucial to guide clinical decision-making about subsequent treatment.

Although the published 2018 International Society for Heart and Lung Transplantation (ISHLT) consensus statement reported risk factors for AC, the risk factors for AS remain unclear [4]. The risk factors for AS are still controversial due to the inconsistency of risk factors among different institutions [16, 17]. In addition, the occurrence of AS is difficult to accurately predict by independent risk factors. Identification of AS status requiring clinical intervention using an accurate prediction model could be valuable to conduct optimal treatment and improve outcomes for LTx patients. However, there has been no satisfactory tool to accurately predict AS requiring clinical intervention. Machine learning (ML) algorithms, a branch of artificial intelligence, can integrate clinical characteristics to achieve accurate predictive outcomes [18]. Our prior research underscored the efficacy of ML algorithms in predicting survival outcomes in LTx patients. Building on this foundation, we endeavored to extend the application of ML models to address the prediction of AS requiring clinical intervention after LTx [19]. No published research has reported using ML algorithms to predict AS requiring clinical intervention.

In this study, we assessed the clinical characteristics of patients and developed ML models to predict AS requiring clinical intervention. Moreover, the conventional logistic regression (LR) model was fitted by independent risk factors and compared in performance to the optimal ML model.

Methods

Patients

Patients who underwent LTx in Wuxi People's Hospital affiliated with Nanjing Medical University between January 2017 and December 2019 were included. The study excluded 3 retransplant patients, 3 pediatric lung transplant patients, 2 patients who were lost to follow-up, and 2 patients with incomplete clinical records. Figure 1 shows a flow chart of the included and excluded patients. All the research procedures were consistent with the ISHLT Ethics statement. The Institutional Review Board of Wuxi People's Hospital affiliated with Nanjing Medical University approved this study (No. 2020 [374]). Patient consent was waived due to the retrospective nature of the study.

Parameter measurements

The following variables were extracted from the database: age, body mass index (BMI), sex, diagnosis, surgical type, extracorporeal membrane oxygenation (ECMO) type, ECMO support, preoperative hormone use, grade 3 primary graft dysfunction at 72 h (72 h PGD 3), operation time, postoperative ventilator time, intensive care unit (ICU) stay, postoperative 6-minute walking test (6MWT), cold-ischemia time, and arterial oxygen tension/inspired oxygen fraction (PaO_2/FiO_2). Diagnoses included interstitial lung disease (ILD), chronic obstructive pulmonary disease (COPD), pulmonary arterial hypertension (PAH), and others. By definition, 72 h PGD 3 refers to the syndrome of acute lung injury over the first 72 h after LTx and is clinically manifested by diffuse alveolar infiltration on chest radiographs with $PaO_2/FiO_2 < 200$ mmHg (10 mmHg = 1.33 kPa) [20]. Cold-ischemia time in single lung transplantation (SLTx) was defined as the interval between the beginning of cold perfusion of the donor lung and blood reperfusion during LTx surgery. For double lung transplantation (DLTx), the cold-ischemia time was determined at the end of reperfusion of the second lung.

Surgery and perioperative management

Since January 1, 2015, China has stopped using organs from executed prisoners, and voluntary organ donation has become the only legal source. Each bronchial anastomosis was performed in an "end-to-end" technique avoiding telescoping during LTx surgery. All recipients
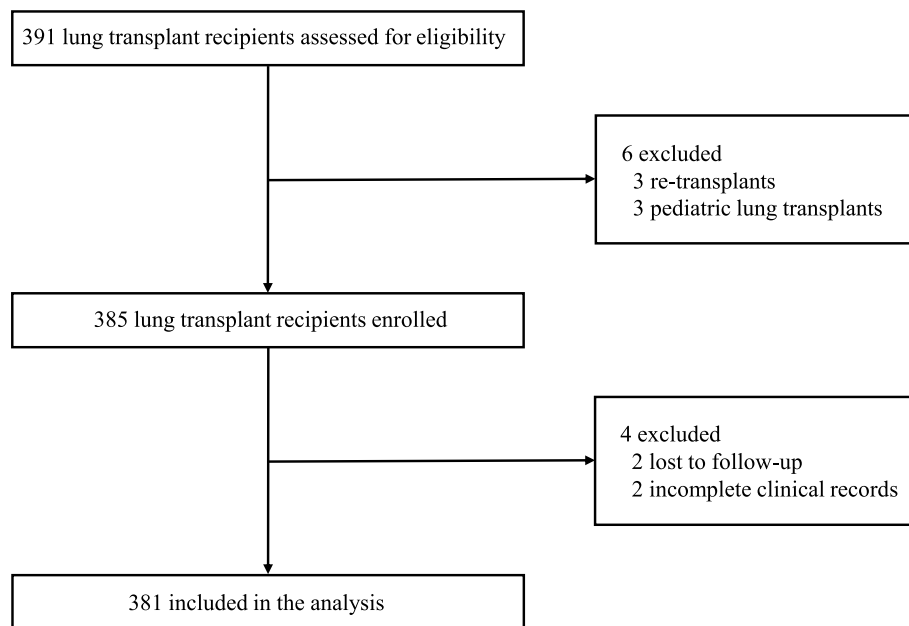
Tian *et al. BMC Medical Informatics and Decision Making*   (2024) 24:229

Page 3 of 13

```
┌──────────────────────────────────────────────┐
│ 391 lung transplant recipients assessed for   │
│ eligibility                                     │
└──────────────────────────────────────────────┘
                  │
                  │          ┌──────────────────────────────┐
                  ├─────────►│ 6 excluded                    │
                  │          │   3 re-transplants            │
                  │          │   3 pediatric lung transplants│
                  │          └──────────────────────────────┘
                  ▼
┌──────────────────────────────────────────────┐
│ 385 lung transplant recipients enrolled        │
└──────────────────────────────────────────────┘
                  │
                  │          ┌──────────────────────────────┐
                  ├─────────►│ 4 excluded                    │
                  │          │   2 lost to follow-up         │
                  │          │   2 incomplete clinical records│
                  │          └──────────────────────────────┘
                  ▼
┌──────────────────────────────────────────────┐
│ 381 included in the analysis                   │
└──────────────────────────────────────────────┘
```

Fig. 1 Flow diagram for selection of lung transplant recipients. A total of 391 lung transplant recipients were assessed for eligibility. Of this cohort, patients with re-transplant, pediatric lung transplant, lost follow-up, and incomplete clinical records were excluded from the study leaving 381 patients available for the analysis

were treated with regular triple immunosuppressive therapy. Patients underwent routine bronchoscopy after LTx, prior to extubation and prior to discharge to assess the condition of the bronchial anastomoses, and the examination frequency was adjusted according to the actual situation. If patients have obvious airflow limitations such as respiratory distress and wheezing, relevant clinical intervention will be activated. An experienced physician (MZL) evaluated the classification of AS based on all definitions and grading systems of AS in the 2018 ISHLT consensus statement [4].

Development of the LR model and ML model

Univariate LR was used to select factors associated with AS based on our cohort. Multivariate LR included only factors with a $P < 0.05$ in univariate LR. A conventional LR model of AS was developed by LR using independent risk factors. For feature selection, three types of methods were used: filtering, wrapping and embedding, which aim to reduce dimension and avoid overfitting of ML models. Within these three categories of feature selection methods, seven methods were utilized, including LR, determination coefficient (DC), Relief, recursive feature elimination (RFE), Boruta, random forest (RF), and least absolute shrinkage and selection operator (LASSO). Finally, 7 groups of features were determined for the subsequent modeling. For the development of ML model, we applied eight ML algorithms, LR, decision tree (DT),

k-nearest neighbors (KNN), naïve bayes (NB), support vector machine (SVM), generalized boosted regression modeling (GBRM), random forest (RF), and extreme gradient boosting (XGB). A total of 56 ML models were developed based on the 8 ML algorithms with 7 feature selection methods for predicting AS requiring clinical intervention. The model with the highest the area under the curve (AUC) was identified as the optimal ML model.

Predictive performance of the LR model and ML model

We compared the predictive performance of the conventional LR model with the optimal ML model for AS requiring clinical intervention. The performance of all models was evaluated in terms of discrimination and calibration. The AUC of the receiver operating characteristic (ROC) curve was used to evaluate the discrimination of the model. The brier score was used to assess the calibration of the model. The brier score ranges from 0 to 1; a score that is close to 0 indicates excellent calibration. Moreover, the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were also evaluated. All statistics were internally validated by the bootstrap method with 1000 resamples.

Patients were stratified into high- and low-risk groups in the optimal ML model based on the threshold determined by ROC. Mean decrease accuracy measures the extent to which each feature's contribution to the model affects the accuracy of the prediction. It was used to

Tian *et al. BMC Medical Informatics and Decision Making*   (2024) 24:229

Page 4 of 13

identify features that contributed most significantly to the optimal ML model performance. In addition, the relative importance scores of each predictor in the optimal RF model were assessed using two metrics: Percentage Increase in MSE (percentage increase in mean square error) and Increase in Node Purity. Percentage Increase in MSE measures the impact of the variable on the prediction performance, while Increase in Node Purity measures the contribution of the variable to the purity of the decision tree nodes.

Statistical analysis

Statistical analysis was performed using IBM SPSS Statistics (version 22.0 Inc., Chicago, IL, USA), R programming language (version 4.2.1, Vienna, Austria) and GraphPad Prism (version 10.1.2, CA, USA). Patient demographics and clinical parameters were summarized as the means ± standard deviations for continuous variables and numbers with percentages for categorical variables. The odds ratio (OR) and 95% confidence interval (CI) were calculated. A value of $P < 0.05$ was considered statistically significant in all analyses.

Results

Clinical characteristics

The clinical characteristics of the LTx patients are summarized in Table 1. A total of 381 patients with 244 males and 137 females were enrolled, and the median age of patients was 57 (range, 19–82) years. In the cohort, most of the indications for LTx were ILD (N = 214) and COPD (N = 67). Regarding surgical type, the numbers of patients with SLTx and DLTx were 201 (52.8%) and 180 (47.2%), respectively. The ECMO type was venoarterial (VA) in 120 cases (31.5%) and venovenous (VV) in 150 cases (39.4%); there were 111 cases (29.1%) that did not involve ECMO. In addition, the operation time, postoperative ventilator time, ICU stay, postoperative 6MWT, cold-Ischemia time and PaO_2/FiO_2 were 327.76 ± 98.39 min, 5.76 ± 12.42 days, 7.78 ± 10.20 days, 460.84 ± 80.58 m, 7.31 ± 2.05 h and 443.55 ± 66.40, respectively. In this study, forty (10.5%) patients encountered AS requiring clinical intervention during the follow-up period.

Development of the LR model and ML model

Univariate analysis indicated that male (OR = 3.535, 95% CI, 1.445–8.650, $P = 0.006$), PAH (OR = 9.651, 95% CI, 2.828–32.930, $P < 0.001$), VV-ECMO (OR = 0.267, 95% CI, 0.100–0.711, $P = 0.008$), and postoperative 6MWT (OR = 0.995, 95% CI, 0.991–0.998, $P = 0.006$) were significantly associated with AS requiring clinical intervention. The multivariate analysis further revealed that male (OR = 7.034, 95% CI, 2.232–22.170, $P < 0.001$), PAH (OR = 11.249, 95% CI, 2.554–49.549, $P < 0.001$), and postoperative 6MWT (OR = 0.993, 95% CI, 0.988–0.997, $P < 0.001$) were independent risk factors for AS requiring clinical intervention (Table 2). Conventional LR models were established based on independent risk factors. For the ML model, a total of 5, 5, 7, 8, 7, and 7 features were selected for modeling in the DC, Relief, RF, RFE, Boruta, and LASSO methods, respectively (Table 3). The combination of 7 feature selection methods and 8 ML algorithms (56 ML models) is shown in a heatmap (Fig. 2). The heatmap shows the AUC for the 56 ML models with a median bootstrapped AUC of 0.679 (range 0.569–0.760). The ML model using an RF algorithm with the DC feature selection method exhibited the highest bootstrapped AUC of 0.760 among the models and was confirmed to be the optimal ML model.

Predictive performance of the LR model and ML model

The model performance for the prediction of AS requiring clinical intervention is summarized in Table 4. The differences emerged in the predicted values of the conventional LR and optimal ML models. The bootstrapped AUC of the optimal ML model was 0.760 (95% CI, 0.666–0.864), which was superior to the conventional LR model of 0.689 (95% CI, 0.545–0.803). The brier score of the optimal ML models was 0.085 (95% CI, 0.058–0.117), outperforming the conventional LR models of 0.091 (95% CI, 0.064–0.125). Furthermore, the sensitivity of the optimal ML model versus the conventional LR model was 0.782 (95% CI, 0.526–1.000) versus 0.680 (95% CI, 0.350–1.000). The specificity of the optimal ML model versus the conventional LR model was 0.689 (95% CI, 0.424–0.917) versus 0.623 (95% CI, 0.305–0.956). The PPV of the optimal ML model versus the conventional LR model was 0.252 (95% CI, 0.133–0.429) versus 0.236 (95% CI, 0.105–0.500). The NPV of the optimal ML model versus the conventional LR model was 0.965 (95% CI, 0.927–1.000) versus 0.952 (95% CI, 0.905–1.000).

A histogram established by the optimal threshold of 0.163 indicates different distributions in the optimal ML model between patients in the high- and low-risk groups (Fig. 3). The majority of patients in the high-risk groups stratified by the optimal ML model presented with AS requiring clinical intervention, while the majority of patients in the low-risk group presented without AS requiring clinical intervention.

Figure 4 illustrates the ranking of features by importance in the optimal ML model for predicting AS requiring clinical intervention. Mean decrease accuracy was calculated over the optimal ML model for the features considered in the model. The five features of the DC feature selection method were postoperative 6MWT, diagnosis, sex, ECMO type, and preoperative hormone use, with postoperative 6MWT being the most significant.

Tian *et al. BMC Medical Informatics and Decision Making*   (2024) 24:229

Page 5 of 13

Table 1 Summary of clinical characteristics of lung transplant recipients with and without AS requiring clinical intervention

Variables	All patients (*n* = 381)	AS (*n* = 40)	Non-AS (*n* = 341)
Age (years)	55.13 ± 12.44	53.95 ± 12.19	55.26 ± 12.48
BMI (kg/m^2)	20.59 ± 3.52	20.27 ± 3.55	20.63 ± 3.51
Sex			
Female	137 (36%)	6 (15%)	131 (38.4%)
Male	244 (64%)	34 (85%)	210 (61.6%)
Diagnosis			
Others	83 (21.8%)	7 (17.5%)	76 (22.3%)
ILD	214 (56.1%)	13 (32.5%)	201 (60.0%)
COPD	67 (17.6%)	12 (30%)	55 (16.1%)
PAH	17 (4.5%)	8 (20%)	9 (2.6%)
Surgical type			
DLTx	180 (47.2%)	18 (45%)	162 (47.5%)
SLTx	201 (52.8%)	22 (55%)	179 (52.5%)
ECMO type			
None	111 (29.1%)	15 (37.5%)	96 (28.2%)
VA	120 (31.5%)	19 (47.5%)	101 (29.6%)
VV	150 (39.4%)	6 (15%)	144 (42.2%)
ECMO support			
No	111 (29.1%)	15 (37.5%)	96 (28.2%)
Yes	270 (70.9%)	25 (62.5%)	245 (71.8%)
Preoperative hormone use			
No	217 (57%)	27 (67.5%)	190 (55.7%)
Yes	164 (43%)	13 (32.5%)	151 (44.3%)
72 h PGD 3			
No	285 (74.8%)	29 (72.5%)	256 (75.1%)
Yes	96 (25.2%)	11 (27.5%)	85 (24.9%)
Operation time (min)	327.76 ± 98.39	328.55 ± 96.20	327.67 ± 98.78
Postoperative ventilator time (days)	5.76 ± 12.42	4.75 ± 11.24	5.88 ± 12.56
ICU stay (days)	7.78 ± 10.20	8.25 ± 11.70	7.72 ± 10.03
Postoperative 6MWT (m)	460.84 ± 80.58	427.52 ± 105.24	464.75 ± 76.42
Cold-ischemia time (hours)	7.31 ± 2.05	7.33 ± 1.95	7.31 ± 2.06
PaO$_2$/FiO$_2$	443.55 ± 66.40	445.61 ± 60.41	443.31 ± 67.15

AS Airway stenosis, *BMI* Body mass index, *ILD* Interstitial lung disease, *COPD* Chronic obstructive pulmonary disease, *PAH* Pulmonary arterial hypertension, *SLTx* Single lung transplantation, *DLTx* Double lung transplantation, *ECMO* Extracorporeal membrane oxygenation, *VA* Venoarterial, *VV* Venovenous, *72 h PGD 3* grade 3 primary graft dysfunction at 72 h, *ICU* Intensive care unit, *6MWT* 6-minute walking test, *PaO$_2$/FiO$_2$* Arterial oxygen tension/inspired oxygen fraction

Figure 5 illustrates the relative importance scores of the predictor variables in the optimal RF model. Postoperative 6MWT showed the highest Percentage Increase in MSE with Increase in Node Purity, implying that it had the greatest impact on the predictive performance of the model and contributed the most to the purity of the decision tree nodes.

Discussion

Considering the significant value of predicting AS requiring clinical intervention in patients after LTx for treatment guidance, we sought to evaluate the clinical characteristics of the patients and further construct prediction models. The following major findings were revealed in this study: (a) Postoperative 6MWT, diagnosis, sex, ECMO type, and preoperative hormone use are five important features of the optimal ML model. (b) Compared with the conventional LR model, the optimal ML model showed better performance in the prediction of AS requiring clinical intervention. (c) The predictive values of the optimal ML model could obviously distinguish patients with AS requiring clinical intervention. Our study suggests that the optimal ML model may become an effective method for predicting AS requiring clinical intervention.

Tian *et al. BMC Medical Informatics and Decision Making*   (2024) 24:229

Page 6 of 13

Table 2 Univariate and multivariate analysis for risk factors of AS requiring clinical intervention after LTx

Variables	Univariate analysis			Multivariate analysis		
	OR	95% CI	*P*	OR	95% CI	*P*
Age (years)	0.992	0.967—1.018	0.527			
BMI (kg/m^2)	0.971	0.883—1.068	0.550			
Sex						
Female	Ref			Ref		
Male	3.535	1.445—8.650	0.006*	7.034	2.232—22.170	0.001*
Diagnosis						
Others	Ref			Ref		
ILD	0.702	0.270—1.827	0.469	0.476	0.172—1.316	0.153
COPD	2.369	0.876—6.405	0.089	1.770	0.566—5.536	0.327
PAH	9.651	2.828—32.930	<0.001*	11.249	2.554—49.549	0.001*
Surgical type						
DLTx	Ref					
SLTx	1.106	0.573—2.136	0.764			
ECMO type						
None	Ref			Ref		
VA	1.204	0.579—2.504	0.619	1.531	0.595—3.938	0.377
VV	0.267	0.100—0.711	0.008*	0.523	0.172—1.591	0.254
ECMO support	0.653	0.330—1.292	0.221			
Preoperative hormone use	0.606	0.302—1.214	0.158			
72 h PGD 3	1.142	0.547—2.385	0.723			
Operation time (min)	1.000	0.997—1.003	0.957			
Postoperative ventilator time (days)	0.991	0.958—1.025	0.590			
ICU stay (days)	1.005	0.976—1.035	0.758			
Postoperative 6MWT (m)	0.995	0.991—0.998	0.006*	0.993	0.988—0.997	0.001*
Cold-ischemia time (hours)	1.003	0.854—1.177	0.971			
PaO$_2$/FiO$_2$	1.001	0.996—1.005	0.835			

AS Airway stenosis, *LTx* Lung transplantation, *BMI* Body mass index, *ILD* Interstitial lung disease, *COPD* Chronic obstructive pulmonary disease, *PAH* Pulmonary arterial hypertension, *SLTx* Single lung transplantation, *DLTx* Double lung transplantation, *ECMO* Extracorporeal membrane oxygenation, *VA* Venoarterial, *VV* Venovenous, *72 h PGD 3* grade 3 primary graft dysfunction at 72 h, *ICU* Intensive care unit, *6MWT* 6-minute walking test, *PaO$_2$/FiO$_2$* arterial oxygen tension/inspired oxygen fraction, *Ref* Reference, *OR* Odds ratio, *CI* Confidence interval

* *P* < 0.05

The 6MWT is used to quantify the functional exercise capacity of patients with moderate to severe lung disease [21]. The negative correlation between the postoperative 6MWT and AS has been described in previous literature [22]. In our study, postoperative 6MWT was the feature with the highest importance in the optimal ML model, indicating the importance of the postoperative 6MWT in predicting AS requiring clinical intervention. PAH is a progressive hemodynamic disease characterized by proliferation and remodeling of small pulmonary arteries [23]. We confirmed that PAH is significantly associated with AS requiring clinical intervention. Patients with PAH are prone to hemodynamic instability in the early postoperative period, which may exacerbate the ischemic condition after LTx by limiting collateral blood flow and lead to development of AS. Sex was usually regarded as a potential contributor to posttransplant complications in LTx patients. The present study found that males were related to an increased probability of AS. Castleberry et al. [24] also reported similar findings. However, Van De Wauwer et al. [25] concluded that males have no negative impact on AS since the sex of the donor and recipient generally overlap. In our opinion, males, with higher levels of PGD after LTx, can have an inadequate anastomotic blood flow supply, which may induce abnormal airway remodeling and increase the occurrence of AS [26]. Additionally, lower estrogen levels in males may lack the protective effect on the airway [27]. VA-ECMO is the bridging modality for patients with respiratory failure awaiting LTx [28]. However, patients on VA-ECMO inherently demonstrate a higher risk of AS episodes since VA-ECMO is more likely to result in bleeding and thrombotic complications compared to VV-ECMO [29]. Our study emphasized the necessity of appropriate use of

Table 3 Features screened by each different feature selection method

Features	DC	Relief	RF	RFE	Boruta	LASSO
Age (years)						X
BMI (kg/m^2)		X	X	X		
Sex	X	X	X	X	X	X
Diagnosis	X		X	X	X	X
Surgical type		X	X			
ECMO type	X		X	X	X	X
ECMO support						
Preoperative hormone use	X					X
72 h PGD 3		X		X	X	
Operation time (min)			X			
Postoperative ventilator time (days)				X	X	X
ICU stay (days)				X	X	
Postoperative 6MWT (m)	X	X	X	X	X	X
Cold-ischemia time (hours)						
PaO$_2$/FiO$_2$						

Features that remained after each feature selection method for each of the measured outcomes are indicated with an X

DC Determination coefficient, *RF* Random forest, *RFE* Recursive feature elimination, *LASSO* Least absolute shrinkage and selection operator, *BMI* Body mass index, *ECMO* Extracorporeal membrane oxygenation, *72 h PGD 3* grade 3 primary graft dysfunction at 72 h, *ICU* Intensive care unit, *6MWT* 6-minute walking test, *PaO$_2$/FiO$_2$* Arterial oxygen tension/inspired oxygen fraction
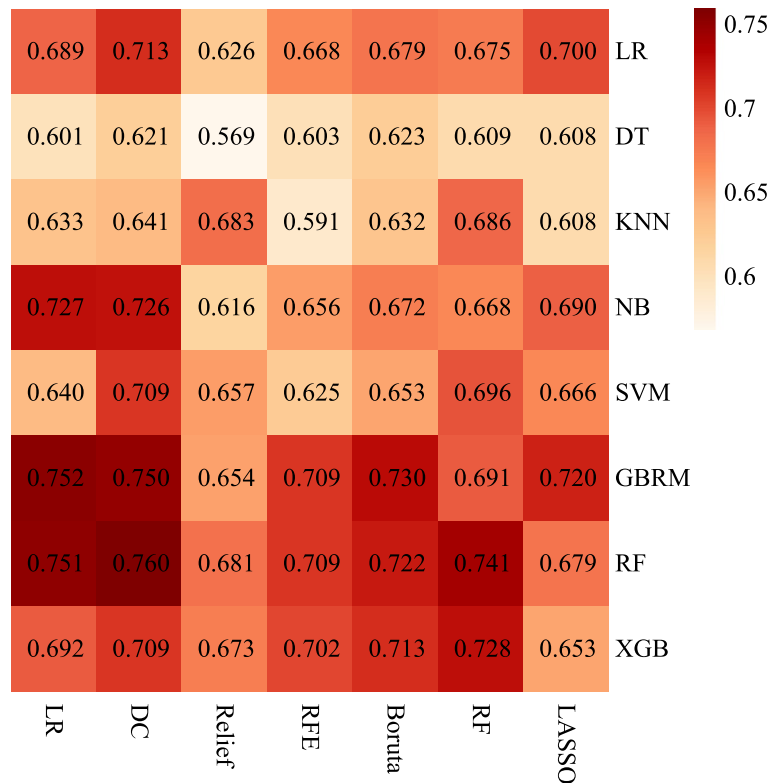
Fig. 2 Heatmaps of the ML models for predicting AS requiring clinical intervention after LTx. Heatmaps illustrated the performance of each ML algorithm (columns) with each feature selection method (rows), measured by AUC. LR, logistic regression; DT, decision tree; KNN, k-nearest neighbors; NB, naïve bayes; SVM, support vector machine; GBRM, generalized boosted regression modeling; RF, random forest; XGB, extreme gradient boosting; LASSO, least absolute shrinkage and selection operator; RFE, recursive feature elimination; DC, determination coefficient; ML, machine learning; AS, airway stenosis; LTx, lung transplantation; AUC, the area under the curve

Table 4 Performance of the optimal RF and conventional LR model

Models	AUC	Brier score	Sensitivity	Specificity	PPV	NPV
RF model[a]	0.760 (0.666–0.864)	0.085 (0.058–0.117)	0.782 (0.526–1.000)	0.689 (0.424–0.917)	0.252 (0.133–0.429)	0.965 (0.927–1.000)
LR model[b]	0.689 (0.545–0.803)	0.091 (0.064–0.125)	0.680 (0.350–1.000)	0.623 (0.305–0.956)	0.236 (0.105–0.500)	0.952 (0.905–1.000)

All statistics were internally validated by bootstrap resampling and the 95% confidence interval was shown in parentheses

RF Random forest, *LR* Logistic regression;, *AUC* the area under the curve, *PPV* Positive predictive value, *NPV* Negative predictive value, *DC* determination coefficient

[a] Feature selection was performed by the DC method

[b] Feature selection was performed by the univariate and multivariate LR
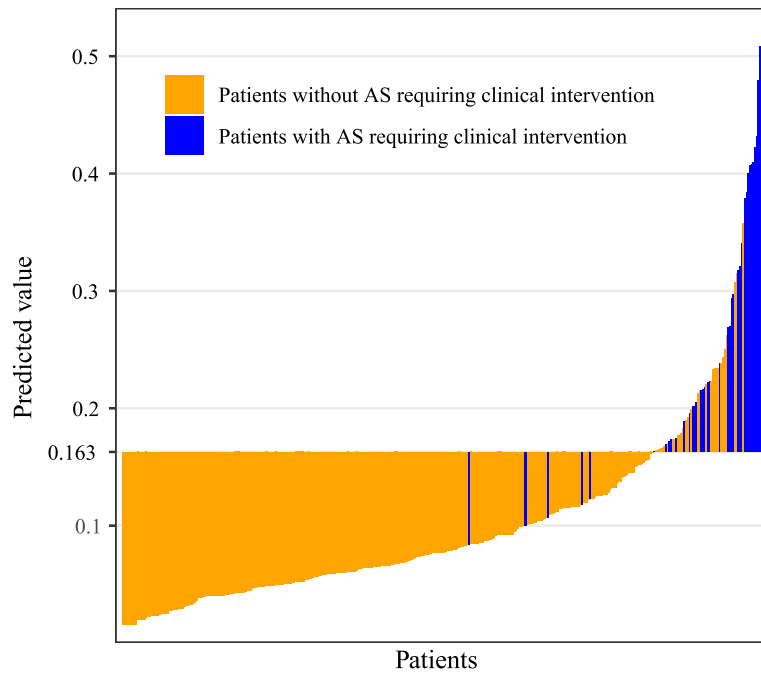
Fig. 3 Histogram of the predicted values in patients with and without AS requiring clinical intervention after LTx. Patients were divided into high- and low-risk patients with a cut-off value of 0.163. Most of the high-risk patients presented with AS requiring clinical intervention, while most of the low-risk patients presented without AS requiring clinical intervention. AS, airway stenosis; LTx, lung transplantation
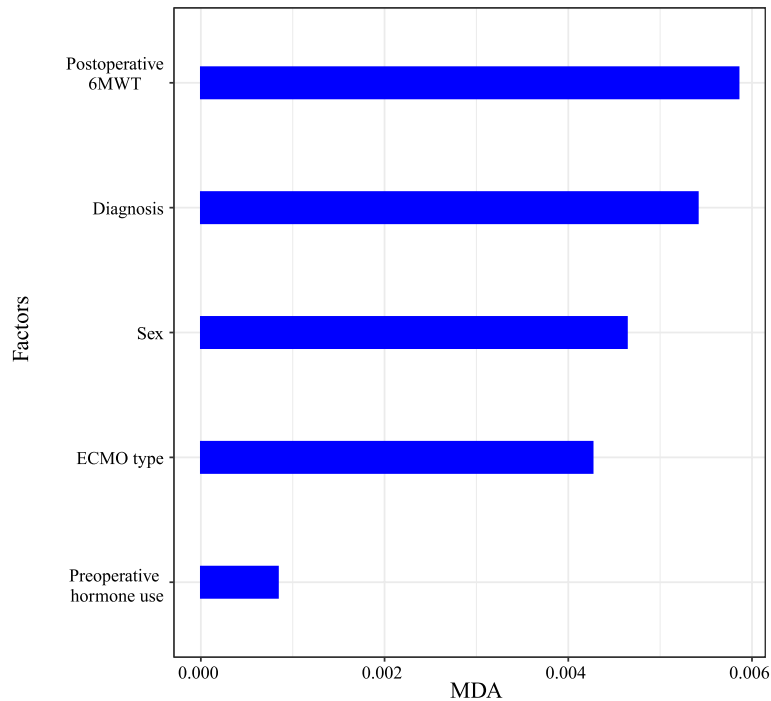
Fig. 4 Variable importance in the optimal RF model. Mean decrease accuracy calculated over the optimal RF model for the features considered in the model. 6MWT: 6-minute walking test; ECMO, extracorporeal membrane oxygenation; RF, random forest
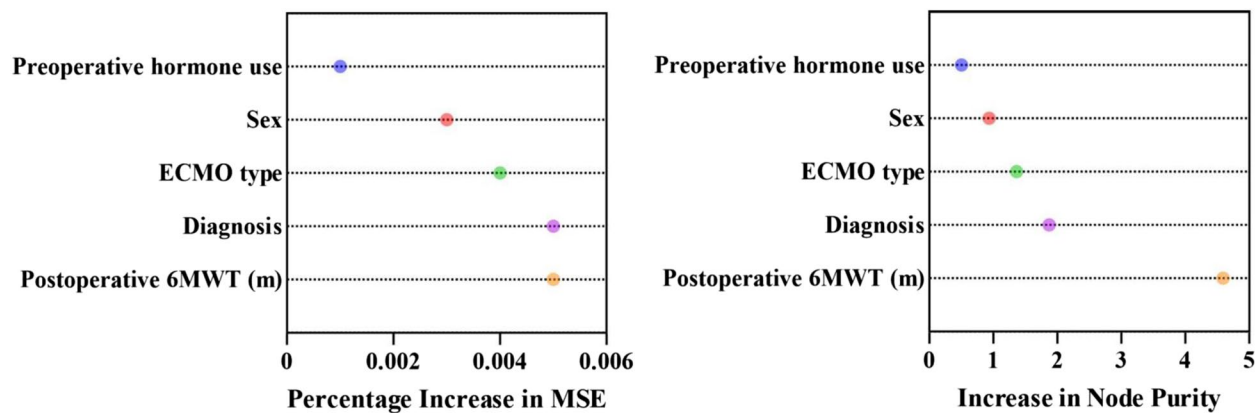
Fig. 5 Relative importance score in the optimal RF model. In the optimal RF model, Percentage Increase in MSE measures the impact of the variable on the prediction performance, while Increase in Node Purity measures the contribution of the variable to the purity of the decision tree nodes. 6MWT, 6-minute walking test; ECMO, extracorporeal membrane oxygenation; Percentage Increase in MSE, percentage increase in mean square error; RF, random forest

VV-ECMO rather than VA-ECMO in the LTx perioperative period. The present study also found that preoperative hormone use (prednisone) increased the incidence of AS, which is consistent with the study by Park et al. [30]. Kim et al. [31] reported that the AC rate did not vary significantly with preoperative hormone use. Nevertheless, they found that the incidence of AC in the first postoperative year remains high after receiving high doses of preoperative prednisone. Hence, the effects of receiving high doses of prednisone preoperatively cannot be ignored. McAnally et al. [32] concluded that preoperative hormone use may induce related complications, such as poor bronchial anastomotic healing and severe infections, which may be the reason for the increased risk of AS episodes. Therefore, reducing the preoperative dose of prednisone or discontinuing prednisone may be a feasible way to reduce the risk of AS episodes.

ML algorithm is a scientific tool that focuses on how computers learn from data [33]. It can be applied to clinical characteristics to develop robust risk prediction models and predict patient outcomes [34]. In previous studies, Hindocha et al. utilized clinical features to develop, validate, and externally test ML model. They found that the ML model might allow satisfactory predictions of survival after treatment for non-small cell lung cancer [18]. In this study, we constructed 56 ML models by clinical characteristics, and an optimal ML model was developed based on the most appropriate RF algorithm and DC feature selection method. A conventional LR model was constructed based on three independent risk factors. The discrimination, calibration, sensitivity, and specificity of the models highlighted their performance. Finally, the bootstrap method was used to internally validate the two models. The bootstrapped AUCs of the optimal ML

model were higher than 0.750, indicating that the optimal ML model had acceptable discrimination. A brier score of 0.085 proves the calibration of the optimal ML model. Both discrimination and calibration demonstrated that the optimal ML model had better performance in predicting AS requiring clinical intervention compared to the conventional LR model.

The optimal ML model has higher sensitivity and specificity than the conventional LR model, further proving that it is an effective prediction model. Our study is the first to assess the predictive value of the optimal ML model for AS requiring clinical intervention in patients after LTx. The important advantage of the optimal ML model is that it exhibits excellent performance and the application of this method does not require data to conform to statistical assumptions, such as the avoidance of independent variable multicollinearity. Although the optimal ML model exhibits the best performance, not all ML models outperform the conventional LR models. Only the ML model constructed with the most appropriate ML algorithm and feature selection method performed best. Additionally, the results of our study do not completely negate the performance of the conventional LR model since they are applicable to different scenarios respectively [35].

Historically, the conventional LR model is widely used to predict the effect of variables on disease [36]. Nevertheless, the conventional LR model assumes that the contribution of all clinical characteristics to the model is linear, which is not applicable to clinical practice. ML models can be better applied to deal with high-dimensional and nonlinear clinical characteristics. Therefore, it is more suitable for clinical practice to achieve good performance. Moreover, the histogram of predicted AS

requiring clinical intervention showed that the predicted outcomes and actual outcomes of the optimal ML model were approximately equal, indicating excellent performance. The majority of high-risk patients presented with AS requiring clinical intervention, and the most intensive follow-up could be performed for high-risk populations. In future studies, developing ML model by using large sample size data is warranted. The ML model could be used in clinical trials to help clinicians screen out high-risk patients and improve patient prognosis.

The limitations of this current study are presented as follows. First, being retrospective, the study had some inevitable selection bias and the results are less convincing than prospective studies. However, strict inclusion and exclusion criteria were used to control for bias. Second, we performed this study in a single center with a relatively small sample size, which limited the application of the model. Therefore, investigations with a large sample size are warranted in the future. Third, microbial infection, an important risk factor, was not evaluated in this study. As patients present with an infectious condition, they are administered the appropriate clinical intervention to suppress the infectious response, which would have an impact on our study results. Fourth, the dataset was imbalanced, with only 10% of patients developing AS. This imbalance may affect the results and the generalization ability of the ML model. Fifth, the study was limited by the absence of certain clinical characteristics such as lung function, imaging, or pathological data, which could potentially enhance the accuracy of predictions. Last, the validation process was conducted by bootstrap resampling instead of application of an independent validation set. Considering that the patient cohort consisted of only 381 individuals, we needed to keep as many samples as possible for model training in order to enhance the model's generalization. However, bootstrapping could not provide comprehensive validation for the model.

Conclusion
In this study, postoperative 6MWT, diagnosis, sex, ECMO type, and preoperative hormone use were identified as five important features of the optimal ML model. We constructed ML models that can effectively predict AS requiring clinical intervention for patients after LTx with good performance. The optimal ML model outperformed the conventional LR model in predicting AS requiring clinical intervention. Multicenter studies with large data samples are warranted to further validate the model. The obtained results may enable early and accurate prediction of AS requiring clinical intervention, guiding clinical decisions for subsequent treatment. Future multi-center studies with large data samples are anticipated to further validate the model. Moreover, the deep learning model could potentially be applied to the personalized treatment of LTx patients in the future.

Abbreviations

AS	Airway stenosis
AUC	The area under the curve
BMI	Body mass index
COPD	Chronic obstructive pulmonary disease
DLTx	Double lung transplantation
DT	Decision tree
DC	Determination coefficient
ECMO	Extracorporeal membrane oxygenation
GBRM	Generalized boosted regression modeling
ILD	Interstitial lung disease
ICU	Intensive care unit
ISHLT	International Society for Heart and Lung Transplantation
KNN	K-nearest neighbors
LTx	Lung transplantation
LASSO	Least absolute shrinkage and selection operator
LR	Logistic regression
ML	Machine learning
NB	Naïve bayes
NPV	Negative predictive value
PPV	Positive predictive value
Percentage Increase in MSE	Percentage increase in mean square error
PAH	Pulmonary arterial hypertension
PaO_2/FiO_2	Arterial oxygen tension/inspired oxygen fraction
RF	Random forest
RFE	Recursive feature elimination
ROC	Receiver operating characteristic
SLTx	Single lung transplantation
SVM	Support vector machine
VA	Venoarterial
VV	Venovenous
XGB	Extreme gradient boosting
72 h PGD 3	Grade 3 primary graft dysfunction at 72 h
6MWT	6- Minute walking test

Availability of data and materials
The data that support the findings of this study are available from the corresponding authors upon reasonable request.

Declarations

Author details
[1]Department of Thoracic Surgery, West China Hospital, Sichuan University, Chengdu 610041, China. [2]Wuxi Lung Transplant Center, Wuxi People's Hospital affiliated to Nanjing Medical University, Wuxi 214023, China. [3]Guangzhou Institute of Respiratory Health, the First Affiliated Hospital of Guangzhou Medical University, Guangzhou 510120, China. [4]Department of General Thoracic Surgery, Juntendo University School of Medicine, Tokyo 113-8431, Japan.

References
1. Chambers DC, Perch M, Zuckermann A, Cherikh WS, Harhay MO, Hayes D Jr, et al. The international thoracic organ transplant registry of the international society for heart and lung transplantation: Thirty-eighth adult lung transplantation report - 2021; Focus on recipient characteristics. J Heart Lung Transplant. 2021;40(10):1060–72.
2. Hardy JD, Webb WR, Dalton ML Jr, Walker GR Jr. Lung homotransplantation in man. Jama. 1963;186:1065–74.
3. Date H, Trulock EP, Arcidi JM, Sundaresan S, Cooper JD, Patterson GA. Improved airway healing after lung transplantation. An analysis of 348 bronchial anastomoses. J Thorac Cardiovasc Surg. 1995;110(5):1424–32 discussion 32–3.
4. Crespo MM, McCarthy DP, Hopkins PM, Clark SC, Budev M, Bermudez CA, et al. ISHLT Consensus Statement on adult and pediatric airway complications after lung transplantation: Definitions, grading system, and therapeutics. J Heart Lung Transplant. 2018;37(5):548–63.
5. Santacruz JF, Mehta AC. Airway complications and management after lung transplantation: ischemia, dehiscence, and stenosis. Proc Am Thorac Soc. 2009;6(1):79–93.
6. Machuzak M, Santacruz JF, Gildea T, Murthy SC. Airway complications after lung transplantation. Thorac Surg Clin. 2015;25(1):55–75.
7. Moreno P, Alvarez A, Algar FJ, Cano JR, Espinosa D, Cerezo F, et al. Incidence, management and clinical outcomes of patients with airway complications following lung transplantation. Eur J Cardiothorac Surg. 2008;34(6):1198–205.
8. Thistlethwaite PA, Yung G, Kemp A, Osbourne S, Jamieson SW, Channick C, et al. Airway stenoses after lung transplantation: incidence, management, and outcome. J Thorac Cardiovasc Surg. 2008;136(6):1569–75.
9. Redmond J, Diamond J, Dunn J, Cohen GS, Soliman AM. Rigid bronchoscopic management of complications related to endobronchial stents after lung transplantation. Ann Otol Rhinol Laryngol. 2013;122(3):183–9.
10. Felton TW, Roberts SA, Isalska B, Brennan S, Philips A, Whiteside S, et al. Isolation of Aspergillus species from the airway of lung transplant recipients is associated with excess mortality. J Infect. 2012;65(4):350–6.
11. Puchalski J, Lee HJ, Sterman DH. Airway complications following lung transplantation. Clin Chest Med. 2011;32(2):357–66.
12. Bin Saeedan M, Rizk A, Yadav R, Ghosh S. Imaging Evaluation of Airway Complications After Lung Transplant. J Comput Assist Tomogr. 2020;44(3):314–27.
13. Varela A, Hoyos L, Romero A, Campo-Cañaveral JL, Crowley S. Management of Bronchial Complications After Lung Transplantation and Sequelae. Thorac Surg Clin. 2018;28(3):365–75.
14. Chhajed PN, Malouf MA, Tamm M, Spratt P, Glanville AR. Interventional bronchoscopy for the management of airway complications following lung transplantation. Chest. 2001;120(6):1894–9.
15. Luecke K, Trujillo C, Ford J, Decker S, Pelaez A, Hazelton TR, et al. Anastomotic airway complications after lung transplant: clinical, bronchoscopic and CT correlation. J Thorac Imaging. 2016;31(5):W62-71.
16. Yserbyt J, Dooms C, Vos R, Dupont LJ, Van Raemdonck DE, Verleden GM. Anastomotic airway complications after lung transplantation: risk factors, treatment modalities and outcome-a single-centre experience. Eur J Cardiothorac Surg. 2016;49(1):e1-8.
17. Shofer SL, Wahidi MM, Davis WA, Palmer SM, Hartwig MG, Lu Y, et al. Significance of and risk factors for the development of central airway stenosis after lung transplantation. Am J Transplant Off J Am Soc Transplant Am Soc Transplant Surg. 2013;13(2):383–9.
18. Hindocha S, Charlton TG, Linton-Reid K, Hunter B, Chan C, Ahmed M, et al. A comparison of machine learning methods for predicting recurrence and death after curative-intent radiotherapy for non-small cell lung cancer: Development and validation of multivariable clinical prediction models. EBioMedicine. 2022;77: 103911.
19. Tian D, Yan HJ, Huang H, Zuo YJ, Liu MZ, Zhao J, et al. Machine Learning-Based Prognostic Model for Patients After Lung Transplantation. JAMA Netw Open. 2023;6(5): e2312022.
20. Snell GI, Yusen RD, Weill D, Strueber M, Garrity E, Reed A, et al. Report of the ISHLT working group on primary lung graft dysfunction, part I: definition and grading-A 2016 Consensus Group statement of the international society for heart and lung transplantation. J Heart Lung Transplant. 2017;36(10):1097–103.
21. Agarwala P, Salzman SH. Six-minute walk test: clinical role, technique, coding, and reimbursement. Chest. 2020;157(3):603–11.
22. Nęcki M, Latos M, Urlik M, Antończyk R, Gaweda M, Pandel A, et al. Number of Bronchoscopic Interventions in Lung Transplant Recipients Correlates with Respiratory Function Assessed by Pulmonary Function Tests. Ann Transplant. 2021;26: e927025.
23. Frost AE. The intersection of pulmonary hypertension and solid organ transplantation. Methodist Debakey Cardiovasc J. 2016;12(4 Suppl):10–3.
24. Castleberry AW, Worni M, Kuchibhatla M, Lin SS, Snyder LD, Shofer SL, et al. A comparative analysis of bronchial stricture after lung transplantation in recipients with and without early acute rejection. Ann Thorac Surg. 2013;96(3):1008–17 discussion 17–8.
25. Van De Wauwer C, Van Raemdonck D, Verleden GM, Dupont L, De Leyn P, Coosemans W, et al. Risk factors for airway complications within the first year after lung transplantation. Eur J Cardiothorac Surg. 2007;31(4):703–10.
26. Chacon-Alberty L, Ye S, Daoud D, Frankel WC, Virk H, Mase J, et al. Analysis of sex-based differences in clinical and molecular responses to ischemia reperfusion after lung transplantation. Respir Res. 2021;22(1):318.
27. Loor G, Brown R, Kelly RF, Rudser KD, Shumway SJ, Cich I, Holley CT, Quinlan C, Hertz MI. Gender differences in long-term survival posttransplant: A single-institution analysis in the lung allocation score era. Clin Transplant. 2017;31(3):10.1111/ctr.12889. https://doi.org/10.1111/ctr.12889. Epub 2017 Feb 8.
28. Faccioli E, Terzi S, Pangoni A, Lomangino I, Rossi S, Lloret A, et al. Extracorporeal membrane oxygenation in lung transplantation: Indications, techniques and results. World J Transplant. 2021;11(7):290–302.
29. Guimbretière G, Anselmi A, Roisne A, Lelong B, Corbineau H, Langanay T, et al. Prognostic impact of blood product transfusion in VA and VV ECMO. Perfusion. 2019;34(3):246–53.
30. Park SJ, Nguyen DQ, Savik K, Hertz MI, Bolman RM 3rd. Pre-transplant corticosteroid use and outcome in lung transplantation. J Heart Lung Transplant. 2001;20(3):304–9.
31. Kim HE, Paik HC, Kim SY, Park MS, Lee JG. Preoperative Corticosteroid Use and Early Postoperative Bronchial Anastomotic Complications after Lung Transplantation. Korean J Thorac Cardiovasc Surg. 2018;51(6):384–9.
32. McAnally KJ, Valentine VG, LaPlace SG, McFadden PM, Seoane L, Taylor DE. Effect of pre-transplantation prednisone on survival after lung transplantation. J Heart Lung Transplant. 2006;25(1):67–74.
33. Deo RC. Machine Learning in Medicine. Circulation. 2015;132(20):1920–30.
34. Greener JG, Kandathil SM, Moffat L, Jones DT. A guide to machine learning for biologists. Nat Rev Mol Cell Biol. 2022;23(1):40–55.

Tian *et al. BMC Medical Informatics and Decision Making* (2024) 24:229

Page 13 of 13

35. Chen Z, Luo H, Xu L. Machine learning models of ischemia/hemorrhage in moyamoya disease and analysis of its risk factors. Clin Neurol Neurosurg. 2021;209: 106919.
36. Nick TG, Campbell KM. Logistic regression. Methods Mol Biol. 2007;404:273–301.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.