



OPEN

## A predictive model for post-thoracoscopic surgery pulmonary complications based on the PBNN algorithm

Cheng-Mao Zhou<sup>1,2</sup>✉, Qiong Xue<sup>2</sup>, HuiJuan Li<sup>2</sup>, Jian-Jun Yang<sup>2</sup>✉ & Yu Zhu<sup>1,2</sup>✉

We constructed an early prediction model for postoperative pulmonary complications after thoracoscopic surgery using machine learning and deep learning algorithms. The artificial intelligence prediction models were built in Python, primarily using artificial intelligence algorithms including both machine learning and deep learning algorithms. Correlation analysis showed that postoperative pulmonary complications were positively correlated with age and surgery duration, and negatively correlated with serum albumin. Using the light gradient boosting machine (LGBM) algorithm, weighted feature engineering revealed that single lung ventilation duration, history of smoking, surgery duration, ASA score, and blood glucose were the main factors associated with postoperative pulmonary complications. Results of artificial intelligence algorithms for predicting pulmonary complications after thoracoscopy in the test group: In terms of accuracy, the two best algorithms were Logistic Regression (0.831) and light gradient boosting machine (0.827); in terms of precision, the two best algorithms were Gradient Boosting (0.75) and light gradient boosting machine (0.742); in terms of recall, the three best algorithms were gaussian naive bayes (0.581), Logistic Regression (0.532), and pruning Bayesian neural network (0.516); in terms of F1 score, the two best algorithms were Logistic Regression (0.589) and pruning Bayesian neural network (0.566); and in terms of Area Under Curve (AUC), the two best algorithms were light gradient boosting machine (0.873) and pruning Bayesian neural network (0.869). The results of this study suggest that pruning Bayesian neural network (PBNN) can be used to assess the possibility of pulmonary complications after thoracoscopy, and to identify high-risk groups prior to surgery.

**Keywords** LGBM, Machine learning, Deep learning, Prediction, PPCs

At present, thoracoscopic technique is widely used in thoracic surgery<sup>1</sup>. Thoracoscopic surgery has a much lower incidence of complications than earlier forms of open surgery<sup>2</sup>. Although there has been great progress in surgical techniques such as thoracoscopic approach and perioperative treatment, the incidence of complications after lobectomy remains high<sup>3,4</sup>. Despite this, pneumonectomy invariably results in impaired lung function as well as further pulmonary complications which have an incidence ranging from 15 to 37%<sup>5</sup>. The incidence of PPCs after thoracoscopic pneumonectomy also remains high, due to one-lung ventilation, nerve injury, lung disease and high concentrations of inhaled oxygen during surgery and stress reaction<sup>6,7</sup>.

Pulmonary complications are the leading cause of perioperative mortality during pneumonectomy<sup>8</sup>. They contribute to prolonged hospital stays, as well as increased ICU admissions and hospital costs<sup>9,10</sup>. Postoperative pulmonary complications (PPCs) may also be a driver of tumor-related deaths<sup>11</sup>. However, targeted measures can alleviate and prevent postoperative pulmonary complications<sup>12</sup>. At present, there are many studies on the prediction of pulmonary complications, but there remains no effective and feasible intelligence prediction research model<sup>11,13,14</sup>. Furthermore, there have been few studies on developing models to predict PPCs after thoracoscopic surgery.

At present, artificial intelligence (AI) such as machine learning plays a significant role in clinical diagnosis and treatment. In recent years, machine learning methods have attracted considerable attention, due to

<sup>1</sup>Big Data and Artificial Intelligence Research Group, Department of Anaesthesiology, Central People's Hospital of Zhanjiang, Zhanjiang, Guangdong, China. <sup>2</sup>Big Data and Artificial Intelligence Research Group, Department of Anesthesiology, Pain and Perioperative Medicine, First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China. ✉email: zhouchengmao187@foxmail.com; yjyangjj@126.com; zhuyu07090527@163.com

their superiority over traditional methods in predicting patient prognosis in a variety of settings and disease conditions<sup>15</sup>. A significant advantage of machine learning techniques lies in their ability to produce more stable predictions by handling complex nonlinear relationships between predictive variables. For example, a recent study has suggested that a variety of AI algorithms can be used to construct prediction models for difficult intubations<sup>16</sup>. Additionally, it has been demonstrated that intelligent algorithms can be used to predict the likelihood of pulmonary complications after emergency gastrointestinal procedures<sup>17</sup>. Machine learning and deep learning techniques can also predict intraoperative bleeding in patients undergoing hepatectomy<sup>18</sup>.

The aim of this study was to construct an early prediction model for PPCs after thoracoscopic surgery using machine learning and deep learning algorithms.

## Methods

### Study population

In this study, patients who had undergone thoracoscopic surgery according to the public BioStudies medical database were analyzed. A total of 905 patients who had undergone thoracoscopic surgery were included. Exclusion criteria: emergency and trauma patients; age < 18 years; patients with preoperative pulmonary infection and/or pleural effusion; patients who had undergone open thoracotomy or whose surgery had been canceled; patients who had undergone a second operation; and patients whose relevant data were incomplete or missing. This study was approved by the Ethics Committee at the First Affiliated Hospital of Zhengzhou University (2020-KY-130). It was exempted from informed consent because it was a retrospective study.

### Data collection

Demographic and clinical variables were collected from multiple patients who had undergone thoracoscopic surgery. General information included age, sex, American Society of Anesthesiologists (ASA) classification, body mass index (BMI), and history of hypertension, diabetes, stroke, heart disease, chronic obstructive pulmonary disease (COPD), smoking, and alcohol use. Preoperative laboratory tests included white blood cell, red blood cell, and platelet counts. Information on surgery and anesthetic management included surgery duration and single lung ventilation duration. The study focused on PPCs, which, according to the definition of European perioperative clinical outcomes, comprise respiratory infection, respiratory failure, pleural effusion, pulmonary atelectasis, pneumothorax, bronchospasm, and aspiration pneumonia<sup>19,20</sup>. In the event that one of these complications was detected within the first seven days after surgery, it was considered a PPC.

### AI algorithms

The AI prediction models were built in Python, primarily using AI algorithms. These included both machine learning and deep learning algorithms, such as Logistic Regression, Decision Tree, Random Forest, Gradient Boosting Decision Tree-Gradient Boosting, Extreme gradient boosting-XGB, light gradient boosting machine-LGBM, Linear Support Vector-LinearSVC, Multilayer Perceptron Classifier-MLPC, Gaussian naive Bayes-gnb, K-nearest neighbors-knn, AdaBoost-adab, Convolutional Neural Network-CNN, Long Short Term Memory-LSTM, Convolutional Neural Network + Recurrent Neural Networks-CNNRNN, Convolutional Neural Network + Long Short Term Memory-CNNLSTM and Pruning Bayesian neural network-PBNN. The dataset was first divided into training and test groups at a ratio of 7:3. Then the AI algorithms were used to build prediction models for data in the training group with fivefold cross-validation. Next, the model's performance was verified in the test group. The LGBM algorithm was used to analyze and rank the weights of each variable accounting for the PPC. Person correlation was used to analyze the relationship between the individual variables. The data were then normalized. Any missing data were processed using the SimpleImputer package. The model was evaluated based on its ROC curve, accuracy, precision, F1 score, recall, Matthews correlation coefficient(MCC), Specificity and MSE score.

### General statistical analysis

R was used to conduct general analysis. The count data were expressed as percentages, with a  $\chi^2$  test for group comparisons. Any measurement data conforming to a normal distribution were expressed as  $\bar{x} \pm s$ , with a *t*-test for group comparisons. A *P* value less than 0.05 was considered significant.

### Ethics approval and consent to participate

This study was approved by the Ethics Committee of the First Affiliated Hospital of Zhengzhou University (2020-KY-130), which exempted the informed consent form because it was a retrospective study.

## Results

### General information

A total of 905 post-thoracoscopic patients were included in this study. In neither the training nor the test groups was there any statistical difference in BMI between the PPC and non-PPC groups (Table 1).

### Correlation analysis between clinical PPC variables

Correlation analysis showed that PPCs were positively correlated with age and surgery duration, and negatively correlated with serum albumin (Fig. 1). Using the LGBM algorithm, weighted feature engineering revealed that single lung ventilation duration, history of smoking, surgery duration, ASA score, and blood glucose were the main factors associated with PPCs (Fig. 2).

PPCS	No	Yes	P value	No	Yes	P value
N	488	145		210	62	
Age (years)	54.4 ± 16.2	65.4 ± 10.7	<0.001	56.1 ± 13.6	66.7 ± 11.3	<0.001
MAP (mmHg)	101.3 ± 13.5	102.0 ± 12.5	0.694	102.5 ± 14.3	107.1 ± 15.7	0.031
Duration of operation	163.1 ± 98.0	255.7 ± 116.9	<0.001	151.7 ± 79.9	275.2 ± 120.3	<0.001
Duration_of one lung ventilation	99.2 ± 58.6	130.8 ± 54.0	<0.001	98.4 ± 51.8	140.7 ± 57.1	<0.001
BMI (kg/m <sup>2</sup> )	22.2 ± 3.6	22.1 ± 3.3	0.635	22.7 ± 3.7	23.1 ± 3.4	0.493
Leukocyte counts (× 10 <sup>9</sup> /L)	6.1 ± 2.2	6.2 ± 2.6	0.849	5.9 ± 1.7	6.0 ± 2.6	0.732
Red blood cell counts (× 10 <sup>12</sup> /L)	4.4 ± 0.6	4.3 ± 0.6	0.155	4.4 ± 0.5	4.2 ± 0.6	0.047
Platelet counts (× 10 <sup>9</sup> /L)	204.1 ± 61.9	196.3 ± 71.5	0.065	203.0 ± 66.6	186.4 ± 67.2	0.086
Aspartate aminotransferase (U/L)	20.9 ± 17.3	19.6 ± 14.4	0.336	20.7 ± 17.8	20.1 ± 16.1	0.800
Alanine aminotransferase (U/L)	22.1 ± 10.3	21.6 ± 8.0	0.586	21.9 ± 8.7	23.5 ± 13.0	0.282
Blood glucose (mmol/L)	5.3 ± 1.1	5.4 ± 1.4	0.656	5.4 ± 1.2	5.3 ± 1.3	0.879
Blood creatinine (μmol/L)	64.1 ± 17.0	67.8 ± 21.1	0.065	63.6 ± 19.9	68.5 ± 17.9	0.080
Blood urea nitrogen (mmol/L)	5.5 ± 1.8	5.7 ± 1.8	0.397	5.6 ± 1.4	5.8 ± 2.1	0.233
Serum sodium (mmol/L)	141.2 ± 6.4	141.9 ± 2.6	0.032	141.7 ± 2.0	141.9 ± 2.2	0.376
Serum potassium (mmol/L)	4.1 ± 0.4	4.0 ± 0.4	0.259	4.0 ± 0.4	4.0 ± 0.4	0.479
Serum albumin (g/L)	40.2 ± 4.1	38.9 ± 4.2	0.005	39.8 ± 4.2	37.8 ± 5.5	0.002
Sex			0.051			0.003
Female	188 (38.5%)	43 (29.7%)		87 (41.4%)	13 (21.0%)	
Male	300 (61.5%)	102 (70.3%)		123 (58.6%)	49 (79.0%)	
ASA			<0.001			<0.001
1	40 (8.2%)	1 (0.7%)		10 (4.8%)	0 (0.0%)	
2	419 (85.9%)	67 (46.2%)		182 (86.7%)	35 (56.5%)	
3	29 (5.9%)	75 (51.7%)		18 (8.6%)	27 (43.5%)	
4	0 (0.0%)	2 (1.4%)		0 (0.0%)	0 (0.0%)	
Airway management			<0.001			<0.001
Single-lumen	18 (3.7%)	1 (0.7%)		6 (2.9%)	4 (6.5%)	
Double-lumen	296 (60.7%)	71 (49.0%)		147 (70.0%)	25 (40.3%)	
Bronchial occluder	165 (33.8%)	73 (50.3%)		57 (27.1%)	33 (53.2%)	
LMA	9 (1.8%)	0 (0.0%)				
History of hypertension			0.021			0.688
No	379 (77.7%)	99 (68.3%)		161 (76.7%)	46 (74.2%)	
Yes	109 (22.3%)	46 (31.7%)		49 (23.3%)	16 (25.8%)	
History of diabetes mellitus			0.005			0.710
No	465 (95.3%)	129 (89.0%)		201 (95.7%)	60 (96.8%)	
Yes	23 (4.7%)	16 (11.0%)		9 (4.3%)	2 (3.2%)	
History of stroke			<0.001			0.115
No	470 (96.3%)	126 (86.9%)		205 (97.6%)	58 (93.5%)	
Yes	18 (3.7%)	19 (13.1%)		5 (2.4%)	4 (6.5%)	
History of COPD			<0.001			
No	481 (98.6%)	116 (80.0%)		202 (96.2%)	58 (93.5%)	0.373
Yes	7 (1.4%)	29 (20.0%)		8 (3.8%)	4 (6.5%)	
History of smoking			<0.001			<0.001
No	416 (85.2%)	68 (46.9%)		171 (81.4%)	32 (51.6%)	
Yes	72 (14.8%)	77 (53.1%)		39 (18.6%)	30 (48.4%)	

**Table 1.** Patient basic characteristic information.

### Results of AI algorithms for predicting pulmonary complications after thoracoscopy in the training group

In terms of accuracy, the two best algorithms were adab (0.915) and CNNRNN (0.912); in terms of precision, the best algorithm was RandomForest (0.944); in terms of recall, the two best algorithms were CNNRNN (0.752) and adab (0.745); in terms of F1 score, the best algorithm was CNNRNN (0.796); and in terms of AUC, the two best algorithms were CNNRNN (0.959) and RandomForest (0.918); MCC value greater than 0.6 includes the following algorithms: CNNRNN, adab, logistic regression, linear SVC, and PBNN; and except for the gnb algorithm, the specificity values of other algorithms are all greater than 0.900 (Table 2 and Fig. 3).

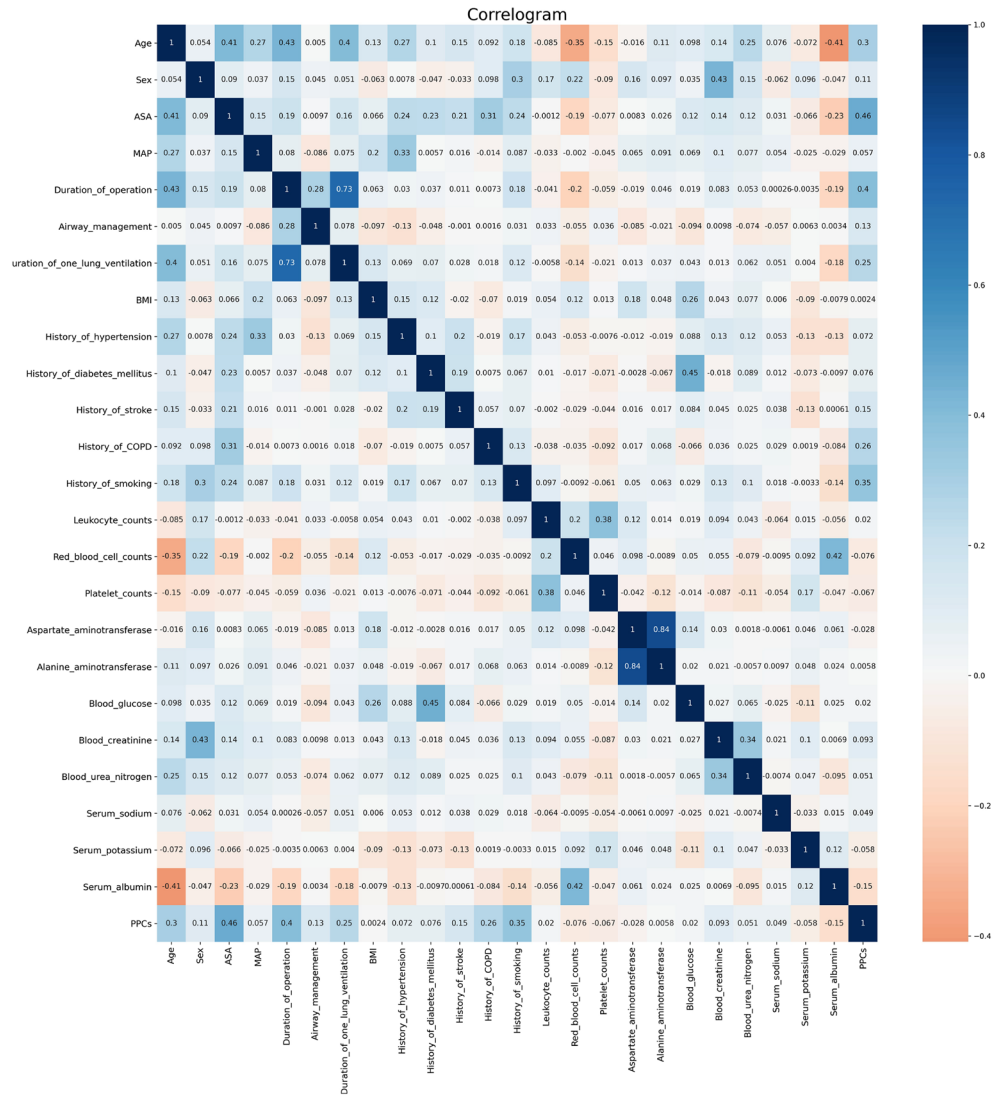


Figure 1. Correlation between variables.

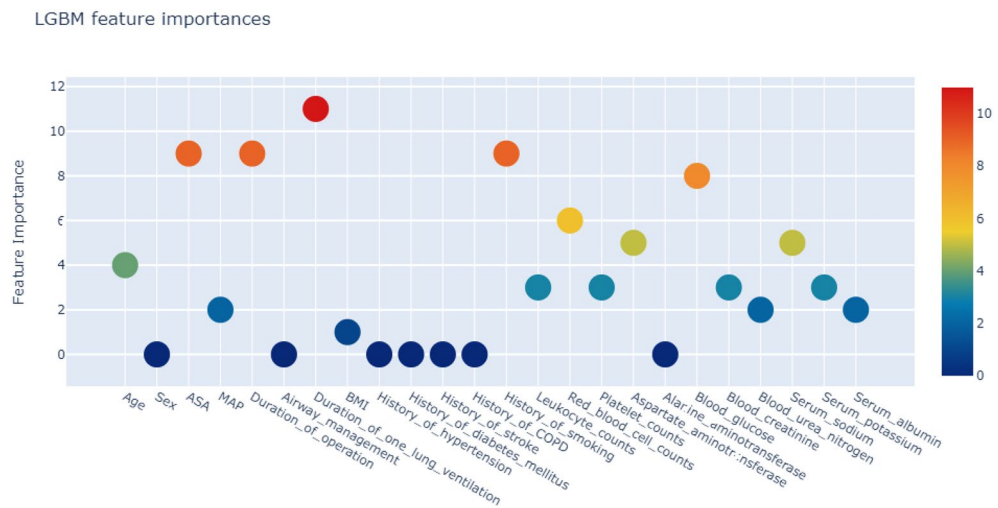
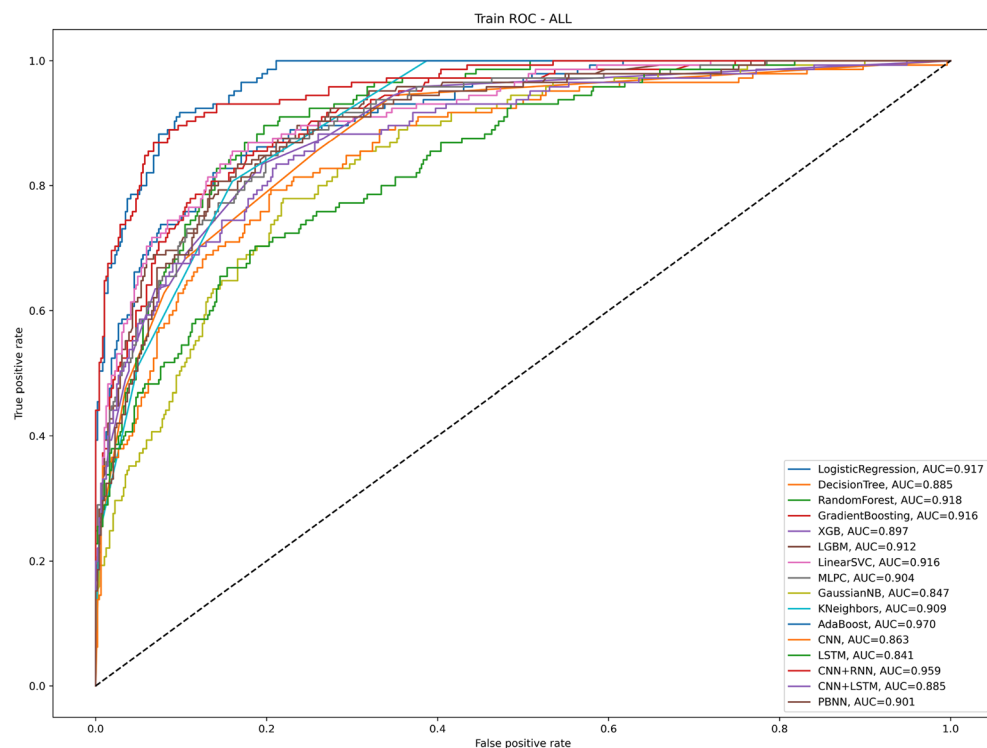


Figure 2. Variable importance of features included in machine learning algorithm.

Model name	AUC	Accuracy	Precision	Recall	F1	Mse	MCC	Specificity
Logistic regression	0.917	0.886	0.807	0.662	0.727	0.114	0.661	0.953
Decision tree classifier	0.885	0.853	0.802	0.476	0.597	0.147	0.541	0.965
Random forest classifier	0.918	0.821	0.944	0.234	0.376	0.179	0.418	0.996
Gradient boosting classifier	0.916	0.848	0.877	0.393	0.543	0.152	0.522	0.984
XGB classifier	0.897	0.858	0.789	0.517	0.625	0.142	0.515	0.977
LGBM classifier	0.912	0.858	0.867	0.448	0.591	0.142	0.556	0.980
LinearSVC	0.916	0.877	0.813	0.600	0.690	0.123	0.627	0.959
MLPC	0.904	0.864	0.847	0.497	0.626	0.136	0.579	0.973
gnb	0.847	0.806	0.568	0.634	0.599	0.194	0.473	0.857
knn	0.909	0.850	0.755	0.510	0.609	0.150	0.536	0.951
adab	0.970	0.915	0.864	0.745	0.800	0.085	0.749	0.965
CNN	0.863	0.831	0.694	0.469	0.56	0.169	0.502	0.941
LSTM	0.841	0.834	0.744	0.421	0.537	0.166	0.459	0.934
CNNRNN	0.959	0.912	0.845	0.752	0.796	0.088	0.821	0.973
CNNLSTM	0.885	0.863	0.790	0.545	0.645	0.137	0.589	0.930
PBNN	0.901	0.853	0.686	0.662	0.674	0.147	0.611	0.959

**Table 2.** Forecast results for training group. Abbreviate: Logistic Regression, Decision Tree, Random Forest, Gradient Boosting Decision Tree-Gradient Boosting, Extreme gradient boosting-XGB, light gradient boosting machine-LGBM, Linear Support Vector-LinearSVC, Multilayer Perceptron Classifier-MLPC, Gaussian naive Bayes-gnb, K-nearest neighbors-knn, AdaBoost-adab, Convolutional Neural Network-CNN, Long Short Term Memory-LSTM, Convolutional Neural Network + Recurrent Neural Networks-CNNRNN, Convolutional Neural Network + Long Short Term Memory-CNNLSTM and Pruning Bayesian neural network-PBNN; Matthews correlation coefficient-MCC.



**Figure 3.** Different AI algorithms predict the PPCs in the training group. Abbreviate: Logistic Regression, Decision Tree, Random Forest, Gradient Boosting Decision Tree-Gradient Boosting, Extreme gradient boosting-XGB, light gradient boosting machine-LGBM, Linear Support Vector-LinearSVC, Multilayer Perceptron Classifier-MLPC, Gaussian naive Bayes-gnb, K-nearest neighbors-knn, AdaBoost-adab, Convolutional Neural Network-CNN, Long Short Term Memory-LSTM, Convolutional Neural Network + Recurrent Neural Networks-CNNRNN, Convolutional Neural Network + Long Short Term Memory-CNNLSTM and Pruning Bayesian neural network-PBNN.

## Results of AI algorithms for predicting pulmonary complications after thoracoscopy in the test group

In terms of accuracy, the two best algorithms were LogisticRegression (0.831) and LGBM (0.827); in terms of precision, the two best algorithms were GradientBoosting (0.75) and LGBM (0.742); in terms of recall, the three best algorithms were gnb (0.581), LogisticRegression (0.532), and PBNN (0.516); in terms of F1 score, the two best algorithms were LogisticRegression (0.589) and PBNN (0.566); and in terms of AUC, the two best algorithms were LGBM (0.873) and PBNN (0.869); the algorithms with the highest MCC value were logistic regression and PBNN; and except for the CNNLSTM, CNNRNN, and gnb algorithms, the specificity values of other algorithms are all greater than 0.900 (Table 3 and Fig. 4).

Taken together, PBNN performed best among these AI algorithms in predicting post-thoracoscopic pulmonary complications.

## Discussion

With its high incidence, PPCs associated with pneumonectomy are a major contributor to prolonged hospitalization, increased postoperative mortality, and medical costs<sup>21,22</sup>. In spite of the use of perioperative pulmonary protective ventilation strategies and minimally invasive thoracoscopic techniques, the incidence of PPCs remains between 12 and 50%<sup>23</sup>. As a result, preventing PPCs is crucial to the prognosis of post-thoracoscopic patients. In this study, PBNN were found to outperform other AI algorithms in predicting PPCs.

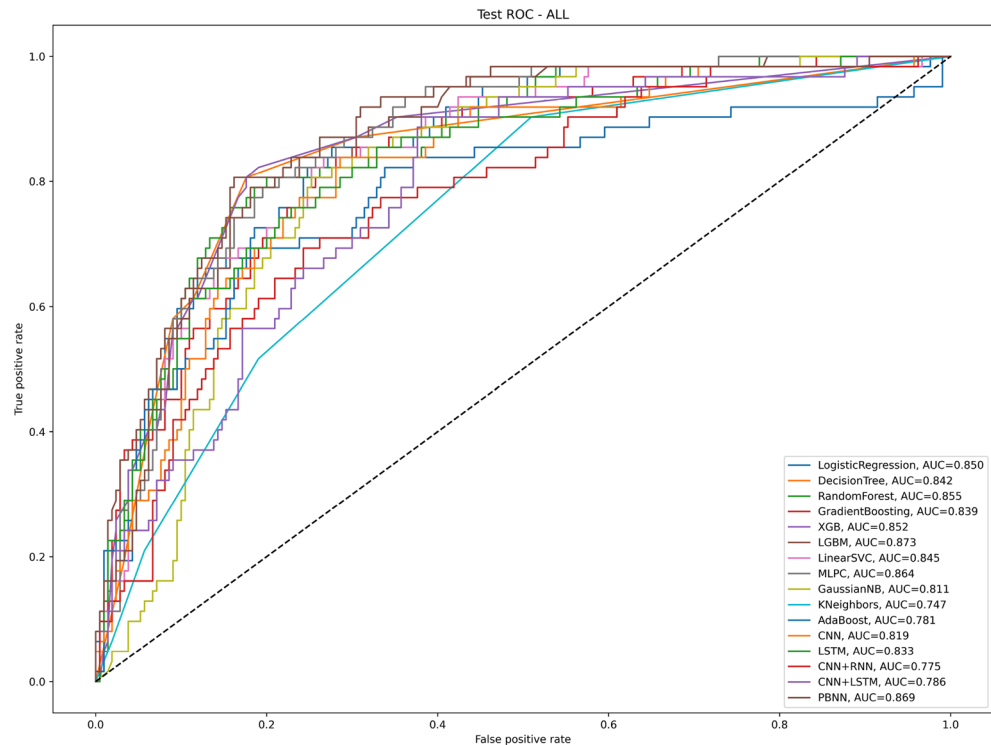
The weighted feature engineering constructed by the LGBM algorithm indicated that the main factors for developing pulmonary complications after thoracoscopy were single-lung ventilation duration, smoking history, surgery duration, ASA score, and blood glucose. The occurrence of PPCs has been shown to be closely related to preoperative interstitial pneumonia and smoking history<sup>24</sup>. A predictive risk model for PPCs can be constructed using age, smoking status, and postoperative 1-s forced expiratory volume<sup>25</sup>. PPCs are also associated with prolonged surgery times<sup>26</sup>. In multivariate analysis, the risk factors associated with increased prevalence of PPCs were ASA physical status  $\geq$  III and surgery duration  $>$  5 h<sup>27</sup>. Surgery duration, one-lung ventilation duration, and ASA score are significant predictors of PPCs after thoracic surgery<sup>28</sup>. The incidence of PPCs was 10.9% among the 6,063 patients who were analyzed, and factors such as advanced age, ASA score, and surgery duration  $\geq$  1 h were the main determinants of pulmonary complications<sup>29</sup>. PPCs have been reported to be significantly influenced by smoking, postoperative blood glucose, and ventilation duration in patients undergoing noncardiac surgery<sup>30</sup>. Additionally, diabetic patients have a higher risk of pulmonary complications during the perioperative period of coronary artery bypass surgery than do non-diabetics<sup>31</sup>. These conclusions are also supported by our findings.

This study does have its limitations. Firstly, it was a retrospective study conducted at a single center, and therefore it is subject to single-center bias. For internal validation, cross-validation were used; however, further multicenter and prospective studies are needed. Furthermore, this retrospective study did not include detailed information on intraoperative hemodynamic fluctuations, postoperative pain, or its treatment.

This study's results suggest that AI algorithms such as PBNN can be used to assess the possibility of pulmonary complications after thoracoscopy, and to identify high-risk groups prior to surgery. Moreover, PBNN's accuracy

Model name	AUC	Accuracy	Precision	Recall	F1	Mse	MCC	Specificity
Logistic regression	0.850	0.831	0.660	0.532	0.589	0.169	0.489	0.919
Decision tree classifier	0.842	0.816	0.658	0.403	0.500	0.184	0.413	0.938
Random forest classifier	0.855	0.779	0.625	0.081	0.143	0.221	0.165	0.986
Gradient boosting classifier	0.839	0.809	0.750	0.242	0.366	0.191	0.351	0.976
XGB classifier	0.852	0.809	0.625	0.403	0.490	0.191	0.435	0.971
LGBM classifier	0.873	0.827	0.742	0.371	0.495	0.173	0.439	0.962
LinearSVC	0.845	0.816	0.667	0.387	0.490	0.184	0.408	0.943
MLPC	0.864	0.801	0.618	0.339	0.438	0.199	0.351	0.938
gnb	0.811	0.787	0.529	0.581	0.554	0.213	0.415	0.848
knn	0.747	0.776	0.520	0.21	0.299	0.224	0.221	0.943
adab	0.781	0.816	0.630	0.468	0.537	0.184	0.433	0.919
CNN	0.819	0.794	0.571	0.387	0.462	0.206	0.334	0.914
LSTM	0.833	0.809	0.647	0.355	0.458	0.191	0.418	0.933
CNNRNN	0.775	0.783	0.528	0.452	0.487	0.217	0.345	0.857
CNNLSTM	0.786	0.787	0.556	0.323	0.408	0.213	0.308	0.871
PBNN	0.869	0.820	0.627	0.516	0.566	0.180	0.452	0.914

**Table 3.** Forecast results for testing group. Abbreviate: Logistic Regression, Decision Tree, Random Forest, Gradient Boosting Decision Tree-Gradient Boosting, Extreme gradient boosting-XGB, light gradient boosting machine-LGBM, Linear Support Vector-Linear SVC, Multilayer Perceptron Classifier-MLPC, Gaussian naive Bayes-gnb, K-nearest neighbors-knn, AdaBoost-adab, Convolutional Neural Network-CNN, Long Short Term Memory-LSTM, Convolutional Neural Network + Recurrent Neural Networks-CNNRNN, Convolutional Neural Network + Long Short Term Memory-CNNLSTM and Pruning Bayesian neural network-PBNN; Matthews correlation coefficient-MCC.



**Figure 4.** Different AI algorithms predict the PPCs in the test group. Abbreviate: Logistic Regression, Decision Tree, Random Forest, Gradient Boosting Decision Tree-Gradient Boosting, Extreme gradient boosting-XGB, light gradient boosting machine-LGBM, Linear Support Vector-LinearSVC, Multilayer Perceptron Classifier-MLPC, Gaussian naive Bayes-gnb, K-nearest neighbors-knn, AdaBoost-adab, Convolutional Neural Network-CNN, Long Short Term Memory-LSTM, Convolutional Neural Network + Recurrent Neural Networks-CNNRNN, Convolutional Neural Network + Long Short Term Memory-CNNLSTM and Pruning Bayesian neural network-PBNN.

rate is 82%, the AUC value is 0.869, and its recall rate and F1 value are both greater than 0.5. Therefore, AI algorithms should be able to facilitate early intervention which will reduce the likelihood of pulmonary complications, and facilitate the management of patients in the perioperative period.

### Data availability

The data is available from the BioStudies public database (<https://www.ebi.ac.uk/biostudies/europepmc/studies/S-EPMC8572520>).

Received: 21 March 2023; Accepted: 20 March 2024

Published online: 25 March 2024

### References

- Kim, S. J. *et al.* Changes in pulmonary function in lung cancer patients after video-assisted thoracic surgery. *Ann. Thorac. Surg.* **99**(1), 210–217 (2015).
- Bendixen, M. *et al.* Postoperative pain and quality of life after lobectomy via video-assisted thoracoscopic surgery or anterolateral thoracotomy for early stage lung cancer: A randomised controlled trial. *Lancet Oncol.* **17**(6), 836–844 (2016).
- Villamizar, N. R. *et al.* Thoracoscopic lobectomy is associated with lower morbidity compared with thoracotomy. *J. Thorac. Cardiovasc. Surg.* **138**(2), 419–425 (2009).
- Suzuki, K. *et al.* Comparison of pulmonary segmentectomy and lobectomy: Safety results of a randomized trial. *J. Thorac. Cardiovasc. Surg.* **158**(3), 895–907 (2019).
- Agostini, P. *et al.* Postoperative pulmonary complications following thoracic surgery: Are there any modifiable risk factors?. *Thorax* **65**(9), 815–818. <https://doi.org/10.1136/thx.2009.123083> (2010).
- Yan, T. *et al.* Prophylactic penehyclidine inhalation for prevention of postoperative pulmonary complications in high-risk patients: A double-blind randomized trial. *Anesthesiology* **136**(4), 551–566 (2022).
- Wang, B., Chen, Z., Zhao, R., Zhang, L. & Zhang, Y. Development and validation of a nomogram to predict postoperative pulmonary complications following thoracoscopic surgery. *PeerJ* **9**, e12366 (2021).
- Miskovic, A. & Lumb, A. B. Postoperative pulmonary complications. *Br. J. Anaesth.* **118**(3), 317–334 (2017).
- Canet, J. *et al.* Prediction of postoperative pulmonary complications in a population-based surgical cohort. *Anesthesiology* **113**(6), 1338–1350. <https://doi.org/10.1097/ALN.0b013e3181fc6e0a> (2010).
- Canet, J. & Gallart, L. Predicting postoperative pulmonary complications in the general population. *Curr. Opin. Anaesthesiol.* **26**, 107–115 (2013).
- Lugg, S. T. *et al.* Long-term impact of developing a postoperative pulmonary complication after lung surgery. *Thorax* **71**(2), 171–176. <https://doi.org/10.1136/thoraxjnl-2015-207697> (2016).

12. Güldner, A. *et al.* Intraoperative protective mechanical ventilation for prevention of postoperative pulmonary complications: A comprehensive review of the role of tidal volume, positive end-expiratory pressure, and lung recruitment maneuvers. *Anesthesiology* **123**(3), 692–713 (2015).
13. Mazo, V., Sabaté, S. & Canet, J. How to optimize and use predictive models for postoperative pulmonary complications. *Minerva Anesthesiol.* **82**(3), 332–342 (2016).
14. Im, Y. *et al.* Impact of interstitial lung abnormalities on postoperative pulmonary complications and survival of lung cancer. *Thorax* **78**(2), 183–190 (2023).
15. Ren, H. *et al.* Predicting acute onset of heart failure complicating acute coronary syndrome: An explainable machine learning approach. *Curr. Probl. Cardiol.* **48**(2), 101480 (2023).
16. Zhou, C. M., Wang, Y., Xue, Q., Yang, J. J. & Zhu, Y. Predicting difficult airway intubation in thyroid surgery using multiple machine learning and deep learning algorithms. *Front. Public Health* **10**, 937471. <https://doi.org/10.3389/fpubh.2022.937471> (2022).
17. Xue, Q. *et al.* Developing machine learning algorithms to predict pulmonary complications after emergency gastrointestinal surgery. *Front. Med. (Lausanne)* **8**, 655686. <https://doi.org/10.3389/fmed.2021.655686> (2021).
18. Xue, Q. *et al.* Predicting intraoperative bleeding in patients undergoing a hepatectomy using multiple machine learning and deep learning techniques. *J. Clin. Anesth.* **74**, 110444. <https://doi.org/10.1016/j.jclinane.2021.110444> (2021).
19. Jammer, I. *et al.* Standards for definitions and use of outcome measures for clinical effectiveness research in perioperative medicine: European Perioperative Clinical Outcome (EPCO) definitions: A statement from the ESA-ESICM joint taskforce on perioperative outcome measures. *Eur. J. Anaesthesiol.* **32**(2), 88–105 (2015).
20. Nijbroek, S. G., Schultz, M. J. & Hemmes, S. N. T. Prediction of postoperative pulmonary complications. *Curr. Opin. Anaesthesiol.* **32**(3), 443–451 (2019).
21. Arslantas, M. K. *et al.* Effect of the amount of intraoperative fluid administration on postoperative pulmonary complications following anatomic lung resections. *J. Thorac. Cardiovasc. Surg.* **149**(1), 314–321.e1. <https://doi.org/10.1016/j.jtcvs.2014.08.071> (2015).
22. Simonsen, D. F., Søgaard, M., Bozi, I., Horsburgh, C. R. & Thomsen, R. W. Risk factors for postoperative pneumonia after lung cancer surgery and impact of pneumonia on survival. *Respir. Med.* **109**(10), 1340–1346. <https://doi.org/10.1016/j.rmed.2015.07.008> (2015).
23. Kim, E. S. *et al.* Prevalence of and risk factors for postoperative pulmonary complications after lung cancer surgery in patients with early-stage COPD. *Int. J. Chron. Obstruct. Pulmon. Dis.* **11**, 1317–1326. <https://doi.org/10.2147/COPD.S105206> (2016).
24. Sekine, Y., Suzuki, H., Nakajima, T., Yasufuku, K. & Yoshida, S. Risk quantification for pulmonary complications after lung cancer surgery. *Surg. Today* **40**(11), 1027–1033. <https://doi.org/10.1007/s00595-009-4182-7> (2010).
25. Yepes-Temiño, M. J., Monedero, P. & Pérez-Valdivieso, J. R. Risk prediction model for respiratory complications after lung resection: An observational multicentre study. *Eur. J. Anaesthesiol.* **33**(5), 326–333. <https://doi.org/10.1097/EJA.0000000000000354> (2016).
26. Foster, C. A. *et al.* Development and validation of procedure-specific risk score for predicting postoperative pulmonary complication: A NSQIP analysis. *J. Am. Coll. Surg.* **229**(4), 355–365.e3. <https://doi.org/10.1016/j.jamcollsurg.2019.05.028> (2019).
27. BevilacquaFilho, C. T. *et al.* Risk factors for postoperative pulmonary complications and prolonged hospital stay in pulmonary resection patients: A retrospective study. *Braz. J. Anesthesiol.* **71**(4), 333–338. <https://doi.org/10.1016/j.bjane.2021.02.003> (2021).
28. Ülger, G. *et al.* The effectiveness of ARISCAT Risk Index, other scoring systems, and parameters in predicting pulmonary complications after thoracic surgery. *Medicine (Baltimore)* **101**(30), e29723. <https://doi.org/10.1097/MD.00000000000029723> (2022).
29. Neto, A. S. *et al.* The LAS VEGAS risk score for prediction of postoperative pulmonary complications: An observational study. *Eur. J. Anaesthesiol.* **35**(9), 691–701. <https://doi.org/10.1097/EJA.0000000000000845> (2018).
30. Jin, Y. *et al.* Incidence and risk factors of postoperative pulmonary complications in noncardiac Chinese patients: A multicenter observational study in university hospitals. *Biomed Res Int.* **2015**, 265165. <https://doi.org/10.1155/2015/265165> (2015).
31. Lauruschkat, A. H. *et al.* Diabetes mellitus as a risk factor for pulmonary complications after coronary bypass surgery. *J. Thorac. Cardiovasc. Surg.* **135**(5), 1047–1053. <https://doi.org/10.1016/j.jtcvs.2007.07.066> (2008).
32. Wang, B., Chen, Z., Zhao, R., Zhang, L. & Zhang, Y. Development and validation of a nomogram to predict postoperative pulmonary complications following thoracoscopic surgery. *PeerJ* **9**, e12366. <https://doi.org/10.7717/peerj.12366> (2021).

## Acknowledgements

We are also grateful to the BioStudies public database for including the original data<sup>32</sup>.

## Author contributions

C.M.Z., J.J.Y, Y.Z. wrote the main manuscript text. Q.X. and H.J.L. prepared Figs. 1, 2, 3 and 4. C.M.Z., J.J.Y, Y.Z. authors reviewed the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to C.-M.Z., J.-J.Y. or Y.Z.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024