



OPEN Development and validation of a machine learning-based nomogram for predicting prognosis in lung cancer patients with malignant pleural effusion

Xin Hu^{1,6}, Shiqiao Zhao^{2,6}, Yanlun Li³, Yiluo Heibi^{1,4}, Hang Wu¹ & Yongjie Jiang⁵✉

Malignant pleural effusion (MPE) is a common complication in patients with advanced lung cancer, significantly impacting their survival rates and quality of life. Effective tools for assessing the prognosis of these patients are urgently needed to enable early intervention. This study retrospectively analyzed patient data from the Affiliated Hospital of North Sichuan Medical College from 2013 to 2021, which served as the training cohort and internal testing cohort. Additionally, three external testing cohorts were introduced: data from Guang'an People's Hospital as cohort 1, data from Dazhou Central Hospital as cohort 2, and data from the Affiliated Hospital of North Sichuan Medical College from January 1, 2023, to December 31, 2023, constituting the temporal external testing cohort. Univariate logistic regression (LR) analysis of clinical variables ($P < 0.05$) was performed, followed by multivariate LR to identify independent predictors for inclusion in nine machine learning models: Decision Tree (DT), Random Forest (RF), Extreme Gradient Boosting (XGBoost), Elastic Net (Enet), Radial Support Vector Machine (rSVM), Multilayer Perceptron (MLP), LR, Light Gradient Boosting Machine (LightGBM), and K-Nearest Neighbors (KNN). The best-performing model was used to develop a nomogram for patient risk stratification. Three variables—treatment regimen, presence of pericardial effusion, and total pleural effusion volume—were identified as significant predictors in the study. The LR model demonstrated the best performance, achieving area under the curve (AUC) values of 0.885 in the training cohort, 0.954 in the internal testing cohort, and 0.920 in external testing cohort 1. To further validate the model's robustness, the nomogram developed from the LR model was evaluated in two additional validation cohorts: external testing cohort 2 and a temporal external testing cohort. The nomogram achieved AUCs of 0.962 in external testing cohort 2 and 0.949 in the temporal external testing cohort, demonstrating strong predictive accuracy. Calibration curves confirmed excellent model-reality concordance across all cohorts, and decision curve analysis (DCA) revealed superior clinical utility. The nomogram enabled individualized risk quantification and showed significant survival differences between high-risk/very high-risk groups and low-risk/medium-risk groups. This study evaluated nine machine learning models for prognostic prediction in lung cancer patients with MPE, finding that the LR-based model offered the best performance. A nomogram based on this model can effectively stratify patients for prognostic assessment and early intervention.

Keywords Lung cancer, Machine learning, MPE, Nomogram, Predictive model

Abbreviations

MPE Malignant pleural effusion

¹Department of Respiratory and Critical Care Medicine, Affiliated Hospital of North Sichuan Medical College, Nanchong, Sichuan, China. ²Department of Emergency Medicine, Affiliated Hospital of North Sichuan Medical College, Nanchong, Sichuan, China. ³Department of Orthopedics, Dazhou Central Hospital, Dazhou, Sichuan, China. ⁴Department of Respiratory and Critical Care Medicine, Guang'an People's Hospital, Guang'an, Sichuan, China. ⁵Department of Respiratory and Critical Care Medicine, Chengdu Third People's Hospital, 19 Yangshi Street, Qingyang District, Chengdu 610000, Sichuan, People's Republic of China. ⁶Xin Hu and Shiqiao Zhao contributed equally to this work. ✉email: 631153933@qq.com

DT	Decision tree
RF	Random forest
XGBoost	Extreme gradient boosting
Enet	Elastic net
rSVM	Radial support vector machine
MLP	Multilayer perceptron
LR	Logistic regression
LightGBM	Light gradient boosting machine
KNN	K-nearest neighbors
AUC	Area under the curve
DCA	Decision curve analysis
ROC	Receiver operating characteristic

Background

MPE can arise from primary malignant tumors of the pleura or from metastases of tumors originating in other parts of the body. The pathogenesis of MPE is multifactorial, involving direct invasion of the pleura by tumor cells, lymphatic obstruction, and increased vascular permeability^{1–3}. These pathophysiological changes lead to fluid accumulation in the pleural cavity, often accompanied by symptoms such as dyspnea and chest pain, significantly impairing patients' daily activities. As a result, the median survival for patients with MPE is typically limited to 3 to 12 months^{4–6}.

Various therapeutic approaches are available for MPE, including ultrasound-guided thoracentesis, indwelling pleural catheter placement, chemical pleurodesis, and intrapleural chemotherapy^{7,8}. However, these treatments are often associated with complications such as chest pain, pneumonia, empyema, pneumothorax, catheter-related infections, and bone marrow suppression^{7–10}. Moreover, there is no conclusive evidence indicating that any particular treatment offers the greatest survival benefit for patients.

Lung cancer is the most common cause of MPE, accounting for approximately one-third of all cases, followed by breast cancer (23%) and lymphoma (10%)¹¹. For lung cancer patients with MPE, accurate prognostic prediction is crucial for tailoring treatment strategies and improving quality of life. This study aims to leverage machine learning techniques to enhance the accuracy of prognostic predictions for these patients, providing clinicians with a reliable tool to optimize patient management and outcomes.

Methods

Patients

This retrospective study collected data from patients treated at the Affiliated Hospital of North Sichuan Medical College and Guang'an People's Hospital between January 1, 2013 and December 31, 2021. Patients from the Affiliated Hospital of North Sichuan Medical College were stratified into a training cohort and an internal testing cohort using a 7:3 ratio. The cohort from Guang'an People's Hospital during the same period served as external testing cohort 1. To enhance the robustness of model validation, we additionally included patients from Dazhou Central Hospital treated between January 1, 2013 and December 31, 2021 as external testing cohort 2, and patients treated at the Affiliated Hospital of North Sichuan Medical College between January 1, 2023 and December 31, 2023 as temporal external testing cohort. The study population included lung cancer patients diagnosed with MPE at initial diagnosis. Diagnostic criteria for MPE were: (a) histopathologically confirmed lung cancer (both non-small cell lung cancer and small cell lung cancer), and (b) MPE confirmed through cytological or pathological evaluation of pleural fluid or pleural tissue. MPE diagnosis was also considered if one of the following conditions was met: (1) histopathologically confirmed lung cancer with no alternative explanation for pleural effusion, or (2) diagnosed lung cancer with pleural effusion and imaging studies (CT, PET-CT, MRI, or ultrasound) revealing pleural nodules, masses, or irregular thickening indicative of pleural metastasis. Exclusion criteria included: (a) lung cancer patients without MPE at initial diagnosis, (b) patients with other types of malignant tumors, (c) individuals under 18 years of age, (d) cases with incomplete clinical information, and (e) patients lost to follow-up. This study was approved by the Ethics Committee of the Affiliated Hospital of North Sichuan Medical College (Approval Number: 2023ER423-1). All methods were performed in accordance with the relevant guidelines and regulations, including the Declaration of Helsinki. Informed consent was waived due to the retrospective nature of the study, where data was sourced from the hospital's electronic medical records and electronic health check databases. All data were anonymized to protect the privacy of participants. Details of the screening process and study flow are presented in Fig. 1.

Machine learning

In the training cohort, variables with a p -value < 0.05 from univariate LR were included in multivariate LR analysis to identify independent predictors. Variables with a p -value < 0.05 in the multivariate analysis were retained for further model development. Nine commonly used machine learning algorithms were applied to predict patient prognosis, including DT, RF, XGBoost, Enet, rSVM, MLP, LR, LightGBM, and KNN. Each model was trained using 5-fold cross-validation and validated on both internal and external testing cohorts. Model performance was assessed using data from the training, internal, and external testing cohorts.

Statistical analysis

Model performance was evaluated using Receiver Operating Characteristic (ROC) curves, with sensitivity, specificity, recall, and AUC calculated. Calibration curves were used to assess the agreement between predicted

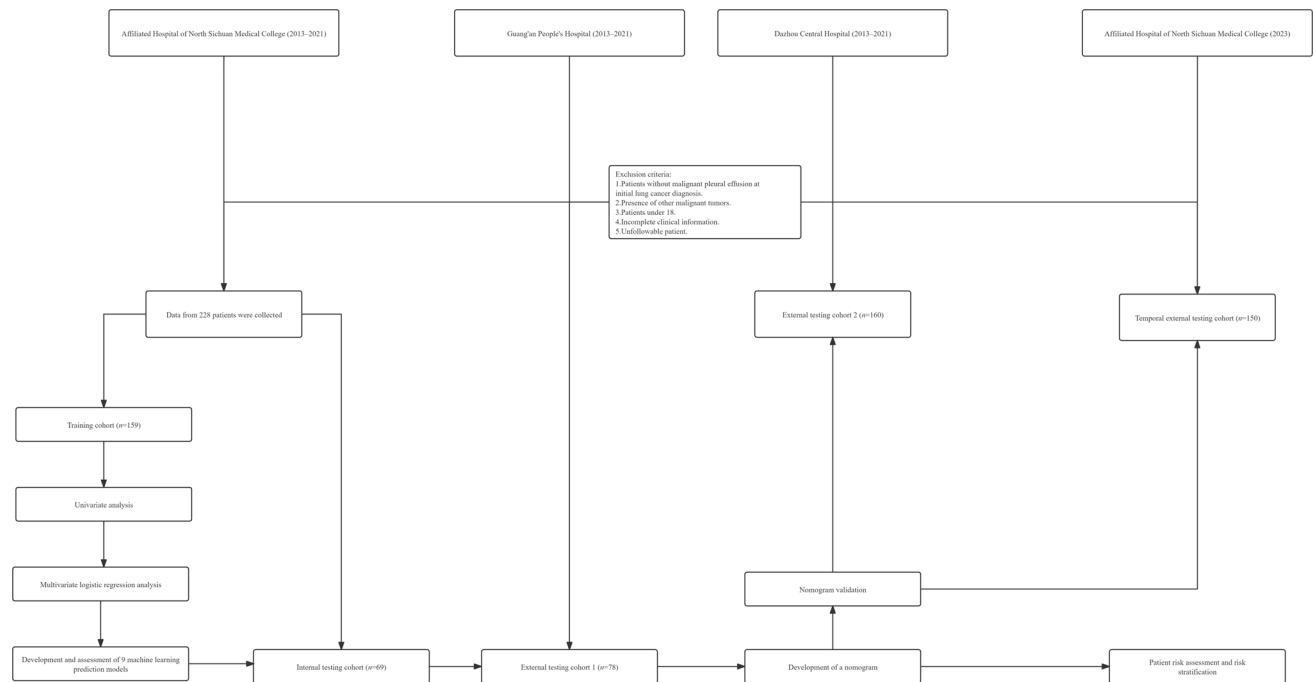


Fig. 1. Overall flowchart of the study.

and observed probabilities, reflecting the model's calibration. Clinical utility was assessed using DCA, which quantifies the net benefit of using the model compared to other diagnostic and prognostic strategies. The best-performing machine learning model was used to develop a nomogram for risk stratification. Risk scores were calculated for each patient, and patients were classified into four risk groups—low, medium, high, and very high—based on thresholds for highest sensitivity, highest specificity, and the combined sensitivity and specificity. Survival curves were generated to illustrate the prognoses for different risk groups. All statistical analyses were conducted using R software (version 4.4.1).

Results

Feature selection

Clinical data from 228 patients were collected at the Affiliated Hospital of North Sichuan Medical College, based on the inclusion and exclusion criteria. These data were divided into a training cohort and an internal testing cohort in a 7:3 ratio (Supplementary Table 1). An additional 79 patients from Guang'an People's Hospital were included as the external testing cohort 1. To further validate the model's generalizability, we established two additional validation cohorts: 160 patients from Dazhou Central Hospital constituted external testing cohort 2, and 150 patients treated at the Affiliated Hospital of North Sichuan Medical College between January 1, 2023 and December 31, 2023 were collected to establish a temporal external testing cohort.

In the training cohort, univariate LR analysis was conducted on all features (Supplementary Table 2). Features with a *p*-value less than 0.05 were identified, including the total volume of pleural effusion, appearance of pleural fluid, smoking history, location of pleural effusion, high-sensitivity C-reactive protein in pleural fluid, pulmonary disease history, pathological subtype of lung cancer, treatment regimen, presence of pericardial effusion, and white blood cell count. These features were further analyzed using multivariate LR, and those with a *p*-value < 0.05 were retained for model construction. The final selected features included treatment regimen, presence of pericardial effusion, and total volume of pleural effusion (Table 1). These three variables were then used in the development of the machine learning models.

Performance of 9 machine learning prediction models

9 machine learning models, including DT, RF, XGBoost, rSVM, MLP, Enet, LR, LightGBM, and KNN, demonstrated good predictive performance across the training cohort (Fig. 2A), internal testing cohort (Fig. 2B), and external testing cohort (Fig. 2C). The area under the curve (AUC) values for each model exceeded 0.80, indicating robust performance (Table 2).

Among the nine models, the LR model showed the best calibration curve consistency across all cohorts (Fig. 2D, E and F). DCA further indicated that the LR model provided greater benefits for patients (Fig. 2G, H and I). Therefore, considering the performance of the LR model in the training cohort, internal testing cohort, and external validation cohort, the LR model is deemed to be the optimal model.

Variable	Coefficient	OR (95% CI)	P
(Intercept)	-1.58	0.21 (0.02–1.65)	0.139
Total volume of pleural effusion	0.14	1.15 (1.01–1.32)	0.037
Appearance of pleural fluid (vs. Yellow)			
Red	-0.10	0.90 (0.31–2.59)	0.847
Other colors	2.04	7.68 (0.83–193.56)	0.119
Smoking: Yes (vs. No)	0.20	1.22 (0.41–3.61)	0.717
Location of pleural effusion (vs. Left side)			
Right side	0.25	1.28 (0.42–3.96)	0.662
Both sides	-0.12	0.88 (0.21–3.71)	0.867
High sensitivity C reactive protein in pleural fluid	0.02	1.02 (1.00–1.05)	0.073
Pulmonary disease: Yes (vs. No)	0.03	1.03 (0.32–3.29)	0.958
Pathological subtype of lung cancer (vs. Adenocarcinoma)			
Squamous cell carcinoma	-0.04	0.96 (0.21–4.93)	0.962
Small cell lung cancer	0.36	1.43 (0.30–7.37)	0.656
Treatment regimen (vs. Untreated)			
Targeted therapy	-3.46	0.03 (0.01–0.10)	<0.001
Other treatment	-2.47	0.08 (0.02–0.28)	<0.001
Presence of pericardial effusion: Yes (vs. No)	1.40	4.04 (1.16–16.21)	0.036
White blood cells	0.13	1.14 (0.96–1.39)	0.155

Table 1. Results of multivariate logistic regression analysis

Development of the nomogram and patient risk assessment

The nomogram, developed from the LR model identified as the optimal machine learning approach, enables individualized risk quantification (Fig. 3A). To validate its stability, we evaluated the nomogram in both an external testing cohort and a temporal external testing cohort. The nomogram demonstrated robust predictive accuracy across both validation sets (Fig. 3B), achieving AUCs of 0.962 (95% CI: 0.937–0.987) in the external cohort and 0.949 (95% CI: 0.918–0.981) in the temporal external cohort. Calibration curves confirmed strong model-reality concordance in both cohorts (Fig. 3C, D). Decision curve analysis revealed superior clinical utility, with net benefit thresholds spanning 0.01–0.95 (Fig. 3E) and 0.02–0.96 (Fig. 3F).

Through the nomogram's point allocation system, we calculated comprehensive risk scores for all patients. These continuous scores were stratified into four clinically actionable categories using thresholds derived from the training cohort's score distribution: low-risk (≤ 19 points), moderate-risk (19–88 points), high-risk (88–162 points), and very high-risk (≥ 162 points), as demonstrated in the training cohort's risk profile visualization (Fig. 3G).

In the low-risk group (≤ 19 points), consisting of 60 patients, one patient had a survival time of less than one year, resulting in a mortality rate of 1.67%. In the moderate-risk group (19–88 points), comprising 245 patients, 56 had a survival time of less than one year, with a mortality rate of 22.86%. The high-risk group (88–162 points), consisting of 201 patients, had 169 patients with a survival time of less than one year, resulting in a mortality rate of 84.08%. The very high-risk group (≥ 162 points), including 111 patients, had a 100% mortality rate, with all patients dying within one year.

Significant survival differences were observed among the risk groups (Fig. 3C), with median survival times of 27 months, 22 months, 8 months, and 2 months for the low-risk, moderate-risk, high-risk, and very high-risk groups, respectively.

Discussion

With advancements in medical technology and a deeper understanding of the pathogenesis of MPE, novel therapies and drugs are continually being developed. For example, research is exploring the use of tumor-derived microparticles to encapsulate chemotherapeutic drugs, allowing for targeted tumor cell destruction while simultaneously activating immune responses to enhance therapeutic efficacy^{12–14}. Additionally, newly identified immunotherapeutic targets, signaling pathways, immunomodulatory cytokines, and non-coding RNAs delivered via exosomes hold promise for advancing immunotherapy in MPE¹⁵. Despite the availability of various treatment options, no consensus has been reached on the most effective approach. Therefore, further investigation into the mechanisms of MPE formation, early prognostic assessment, and personalized treatment strategies is essential.

This study found that the treatment regimen (targeted therapy, other treatment, or no treatment), presence of pericardial effusion, and total volume of pleural effusion were significantly associated with patients' one-year survival rate. Targeted therapy, by specifically targeting molecular markers of tumors, can more precisely inhibit tumor growth and spread^{16–18}, thereby reducing the incidence of malignant pleural effusion. Due to its high efficacy and relatively low side effects, targeted therapy offers superior long-term outcomes compared to other treatment options, such as chemotherapy and radiotherapy¹⁹. In the cases examined in this study, cancer cells typically spread to the pleura and pericardium via the lymphatic system or bloodstream, leading

	Training cohort				Internal testing cohort				External testing cohort 1				F1 Score
	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	F1 Score	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	F1 Score	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)		
Models													
DT	0.838 (0.777–0.898)	0.744 (0.649–0.838)	0.870 (0.795–0.945)	0.797	0.890 (0.814–0.966)	0.667 (0.513–0.821)	0.939 (0.858–1.000)	0.774	0.873 (0.787–0.958)	0.893 (0.812–0.974)	0.870 (0.732–1.000)	0.917	
RF	0.889 (0.840–0.939)	0.817 (0.733–0.901)	0.844 (0.763–0.925)	0.832	0.939 (0.888–0.991)	0.778 (0.642–0.914)	0.939 (0.858–1.000)	0.848	0.909 (0.844–0.974)	0.893 (0.812–0.974)	0.783 (0.614–0.951)	0.901	
XGBoost	0.899 (0.851–0.947)	0.805 (0.719–0.891)	0.870 (0.795–0.945)	0.835	0.946 (0.898–0.995)	0.750 (0.609–0.891)	0.939 (0.858–1.000)	0.831	0.905 (0.837–0.973)	0.857 (0.765–0.949)	0.783 (0.614–0.951)	0.881	
Enet	0.883 (0.831–0.935)	0.841 (0.762–0.921)	0.844 (0.763–0.925)	0.847	0.959 (0.919–1.000)	0.861 (0.748–0.974)	0.939 (0.858–1.000)	0.899	0.920 (0.860–0.981)	0.893 (0.812–0.974)	0.783 (0.614–0.951)	0.901	
rSYM	0.884 (0.832–0.936)	0.841 (0.762–0.921)	0.818 (0.732–0.904)	0.836	0.961 (0.921–1.000)	0.861 (0.748–0.974)	0.939 (0.858–1.000)	0.899	0.923 (0.863–0.983)	0.893 (0.812–0.974)	0.783 (0.614–0.951)	0.901	
MLP	0.881 (0.829–0.934)	0.829 (0.748–0.911)	0.857 (0.779–0.935)	0.845	0.955 (0.912–0.998)	0.806 (0.676–0.935)	0.939 (0.858–1.000)	0.866	0.915 (0.853–0.977)	0.893 (0.812–0.974)	0.783 (0.614–0.951)	0.901	
LR	0.885 (0.834–0.937)	0.829 (0.748–0.911)	0.844 (0.763–0.925)	0.840	0.954 (0.909–1.000)	0.833 (0.712–0.955)	0.939 (0.858–1.000)	0.882	0.920 (0.859–0.980)	0.893 (0.812–0.974)	0.783 (0.614–0.951)	0.901	
LightGBM	0.894 (0.846–0.943)	0.829 (0.748–0.911)	0.844 (0.763–0.925)	0.840	0.956 (0.913–0.999)	0.833 (0.712–0.955)	0.939 (0.858–1.000)	0.882	0.903 (0.833–0.972)	0.893 (0.812–0.974)	0.783 (0.614–0.951)	0.901	
KNN	0.878 (0.827–0.930)	0.732 (0.636–0.828)	0.896 (0.828–0.964)	0.800	0.929 (0.869–0.989)	0.611 (0.452–0.770)	0.939 (0.858–1.000)	0.733	0.887 (0.806–0.968)	0.857 (0.765–0.949)	0.870 (0.732–1.000)	0.897	

Table 2. Performance of 9 machine learning prediction models across three cohorts.

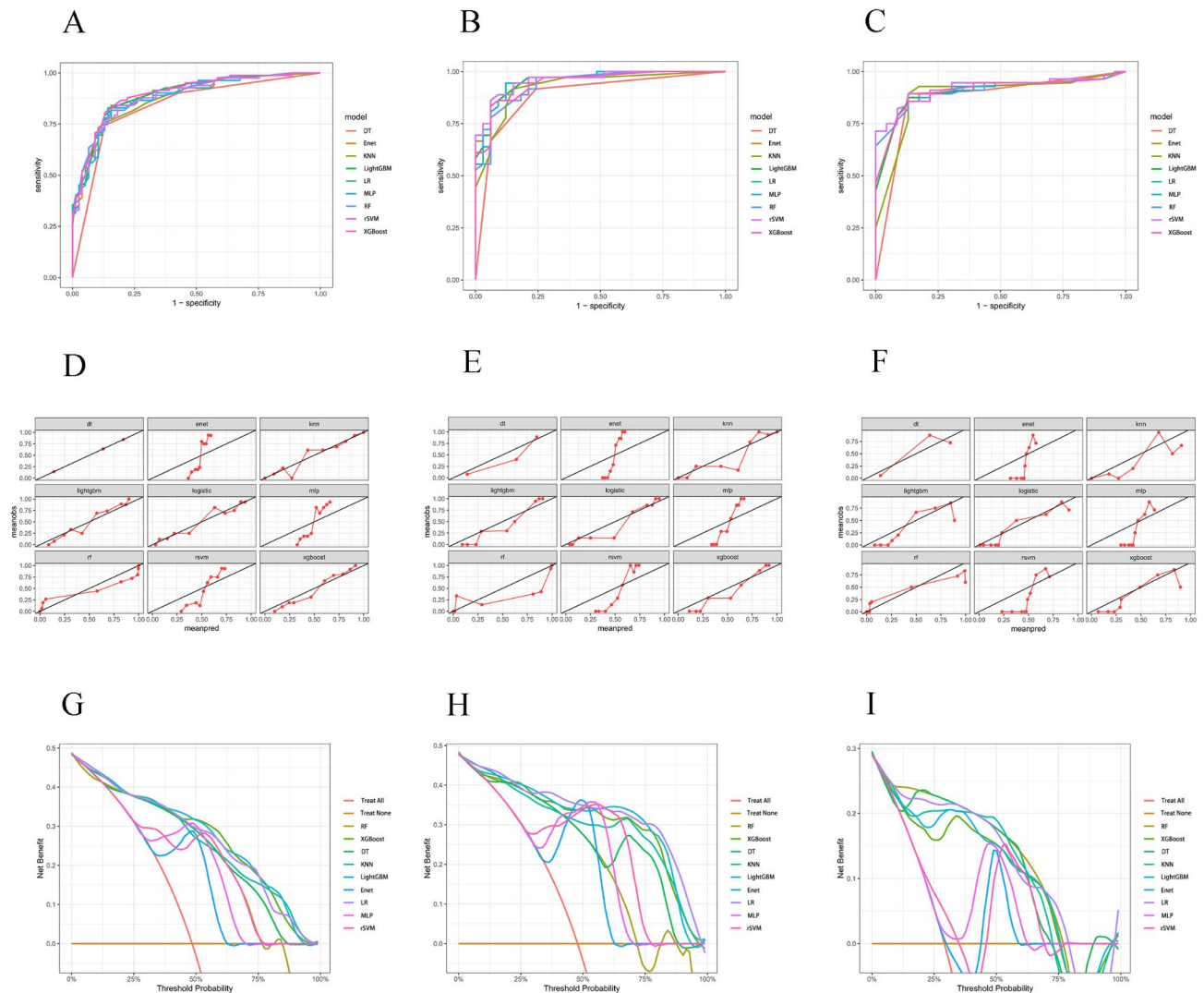
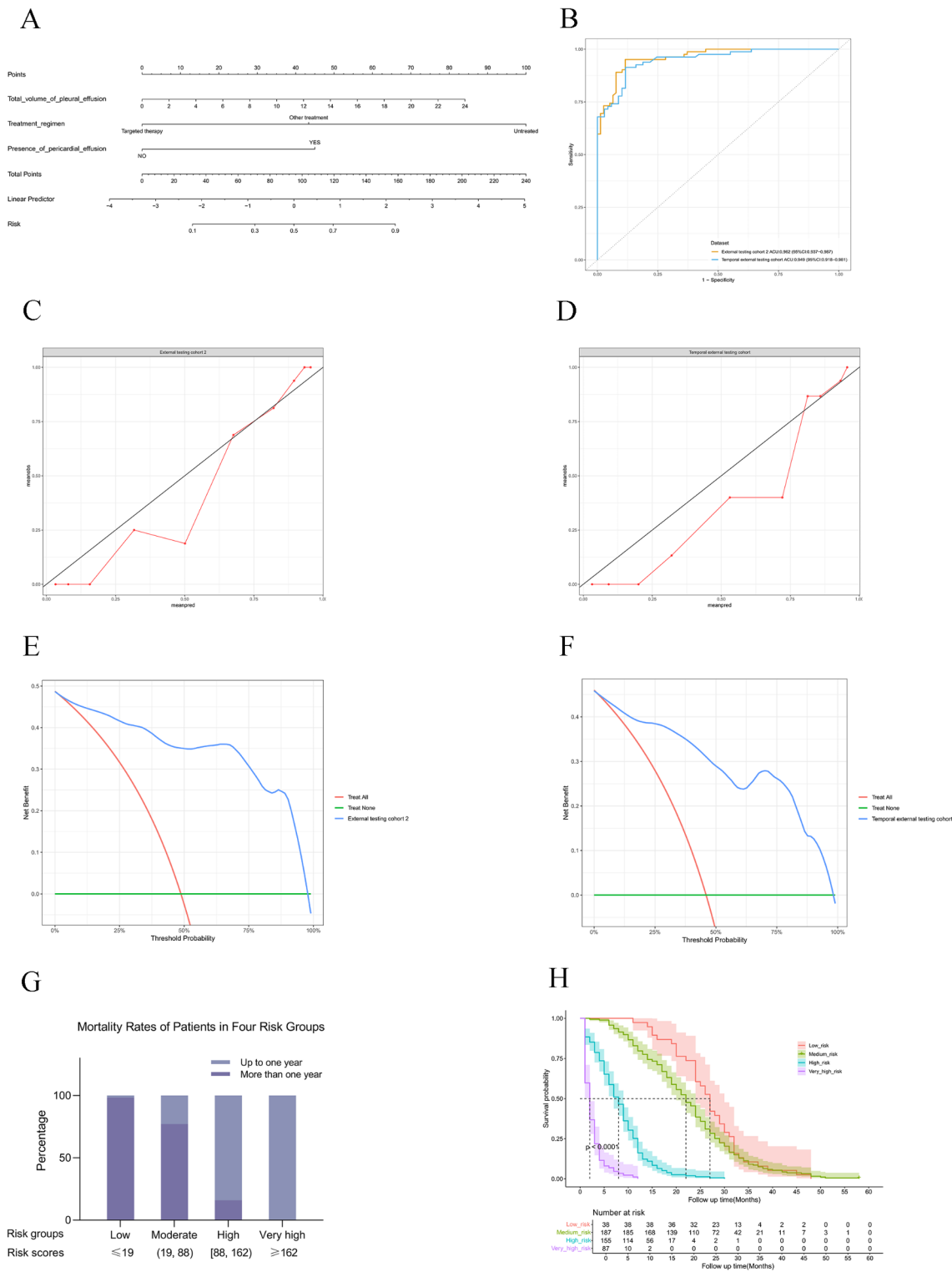


Fig. 2. Evaluation of nine machine learning models for predicting patient prognosis. (A)–(C). ROC curves of the machine learning models in the training cohort (A), internal testing cohort (B), and external testing cohort 1 (C). (D)–(F). Calibration curves of the machine learning models in the training cohort (D), internal testing cohort (E), and external testing cohort 1 (F). (G)–(I). DCA of the machine learning models in the training cohort (G), internal testing cohort (H), and external testing cohort 1 (I). Abbreviations: ROC, receiver operating characteristic; DCA, decision curve analysis.

to the formation of pleural and pericardial effusions^{20,21}. The coexistence of these conditions, along with larger volumes of malignant pleural effusion, often indicates advanced cancer progression and greater therapeutic challenges. Pericardial effusion directly affects cardiac function, potentially causing heart failure, hypotension, and circulatory collapse^{22–24}. Meanwhile, pleural effusion exerts a compressive effect on the lungs, impairing respiratory function, oxygen exchange, and overall cardiopulmonary health^{7,20}. As the volume of effusion increases, these compressive effects intensify. Large amounts of pleural effusion can further compress the heart, restricting its function⁶. Similarly, pericardial effusion can compress the lungs, limiting pulmonary expansion, reducing ventilation capacity, and exacerbating hypoxemia. Together, these factors contribute to poor patient prognosis.

This study identified three critical variables—treatment regimen, pericardial effusion, and total pleural effusion volume—for prognostic prediction in lung cancer patients with MPE. The LR model showed the best performance, with AUCs of 0.885 (training), 0.954 (internal testing), and 0.920 (external testing cohort 1). The nomogram achieved AUCs of 0.962 (external testing cohort 2) and 0.949 (temporal external validation cohort), with excellent calibration and clinical utility (net benefit thresholds: 0.01–0.95 and 0.02–0.96, respectively). The nomogram stratifies patients into four risk groups—low-risk (≤ 19 points), moderate-risk (19–88 points), high-risk (88–162 points), and very high-risk (≥ 162 points)—providing an effective tool for predicting one-year mortality and assessing overall survival prognosis.

The treatment of lung cancer with MPE involves both local therapy (symptom relief) and systemic therapy (control of primary tumors and metastases)²⁵. The goal of local therapy is primarily to alleviate dyspnea, reduce



◀ **Fig. 3.** Development, validation, and risk stratification of the nomogram. (A) A nomogram with three predictive variables developed from the LR model for predicting 1-year survival probability in patients. (B) ROC curves of the nomogram in external testing cohort 2 and temporal external testing cohort. (C)–(D). Calibration curves of the nomogram in external testing cohort 2 (C) and temporal external testing cohort (D). (E)–(F). DCA of the nomogram in external testing cohort 2 (E) and temporal external testing cohort (F). (G) Risk stratification based on nomogram scores: All patients were scored using the nomogram. Patients in the training cohort were classified into four risk groups—low, medium, high, and very high—using thresholds for highest sensitivity, highest specificity, and combined sensitivity and specificity. The proportions of two distinct survival outcome categories within each risk group are displayed. (H) Survival curves of patients across different risk groups (excluding those from the temporal external testing cohort, as follow-up data for this cohort, collected from January 1, 2023, to December 31, 2023, did not meet the minimum 2-year follow-up requirement). Abbreviations: LR, logistic regression; ROC, receiver operating characteristic; DCA, decision curve analysis.

effusion recurrence, and improve quality of life. The main local treatment modalities include therapeutic thoracentesis, pleurodesis, hyperthermic intrathoracic chemotherapy, intrapleural drug perfusion, and indwelling pleural catheter. These approaches, which focus solely on managing MPE without targeting the primary tumor, were categorized as “untreated” in our study. In contrast, systemic therapy aims to inhibit tumor progression and prolong patient survival, encompassing systemic chemotherapy, targeted therapy, immune checkpoint inhibitors, and anti-angiogenic therapy. Among these, our study found that targeted therapy provided the greatest benefit to patients compared to all other treatment modalities. Based on the nomogram developed in our study and the risk stratification of patients (low-risk, moderate-risk, high-risk, and very high-risk groups), specific treatment recommendations can be made. For patients in the low-risk or moderate-risk groups, regular systemic therapy is strongly recommended, as these groups achieved median survival times of 27 months and 22 months, respectively. For patients in the high-risk group, long-term systemic therapy may be considered based on their financial situation, although the potential benefit is relatively limited, with a median survival time of only 8 months. For patients in the very high-risk group, the focus should be on reducing MPE and improving quality of life, given their short median survival time of just 2 months.

Despite the strong predictive performance of the nomogram, this study has several limitations. First, the analysis was limited to patients’ laboratory and basic clinical data, lacking an evaluation of lung cancer imaging features. Incorporating tumor imaging characteristics might further enhance the nomogram’s predictive accuracy. Second, although external validation was performed, the external dataset was relatively small. Therefore, additional validation with larger and more diverse samples from various regions is required to enhance the generalizability and reliability of the model.

Conclusion

In this study, we comprehensively assessed the predictive performance of nine machine learning models for prognostication in lung cancer patients with MPE. Our analysis revealed that the LR model, based on key predictors such as treatment regimen, presence of pericardial effusion, and total pleural effusion volume, demonstrated the best performance in terms of accuracy, calibration, and clinical utility. The nomogram derived from this model enables effective risk stratification, offering clinicians a practical tool for early prognostic assessment and intervention.

Data availability

The datasets that were either generated or analyzed during the current study are made available by the corresponding author upon receipt of a reasonable request.

Received: 21 December 2024; Accepted: 10 March 2025

Published online: 21 March 2025

References

- Psallidas, I., Kalomenidis, I., Porcel, J. M., Robinson, B. W. & Stathopoulos, G. T. Malignant pleural effusion: From bench to bedside. *Eur. Respir. Rev.* **25** (140), 189–198 (2016).
- Penz, E., Watt, K. N., Hergott, C. A., Rahman, N. M. & Psallidas, I. Management of malignant pleural effusion: Challenges and solutions. *Cancer Manag. Res.* **9**, 229–241 (2017).
- Feller-Kopman, D. & Light, R. Pleural disease. *New. Engl. J. Med.* **378** (18), 1754 (2018).
- Clive, A. O. et al. Predicting survival in malignant pleural effusion: Development and validation of the LENT prognostic score. *Thorax* **69**(12), 1098–1104 (2014).
- Asciak, R. & Rahman, N. M. Malignant pleural effusion: From diagnostics to therapeutics. *Clin. Chest Med.* **39** (1), 181–193 (2018).
- Bibby, A. C. et al. ERS/EACTS statement on the management of malignant pleural effusions. *Eur. J. Cardio-Thorac.* **55** (1), 116–132 (2019).
- Bashour, S. I., Mankidy, B. J. & Lazarus, D. R. Update on the diagnosis and management of malignant pleural effusions. *Resp. Med.* **196**, 106802 (2022).
- Zhu, G. et al. Evaluation of efficacy and safety for Kanglaite injection in the control of the malignant pleural effusions via thoracic perfusion: A systematic review and Meta-Analysis of randomized controlled trials. *Front. Pharmacol.* **12**, 694129 (2021).
- Egan, A. M., McPhillips, D., Sarkar, S. & Breen, D. P. Malignant pleural effusion. *QJM-Int J. MED.* **107** (3), 179–184 (2014).
- Chalhoub, M., Saqib, A. & Castellano, M. Indwelling pleural catheters: Complications and management strategies. *J. Thorac. Dis.* **10** (7), 4659–4666 (2018).

11. DetterbeckFC, LewisSZ, Diekemper, R., Addrizzo-Harris, D. & Alberts, W. M. Executive summary: Diagnosis and management of lung cancer, 3rd ed: American college of chest physicians evidence-based clinical practice guidelines. *Chest* **143** (5 Suppl), 7S–37S (2013).
12. Ma, J., Zhang, H., Tang, K. & Huang, B. Tumor-derived microparticles in tumor immunology and immunotherapy. *Eur. J. Immunol.* **50** (11), 1653–1662 (2020).
13. Xu, P. et al. Chemotherapeutic tumor microparticles elicit a neutrophil response targeting malignant pleural effusions. *Cancer Immunol. Res.* **8** (9), 1193–1205 (2020).
14. Shao, J., Zaro, J. & Shen, Y. Advances in exosome-based drug delivery and tumor targeting: From tissue distribution to intracellular fate. *Int. J. Nanomed.* **15**, 9355–9371 (2020).
15. Ge, S. et al. Immune modulation in malignant pleural effusion: From microenvironment to therapeutic implications. *Cancer Cell. Int.* **24** (1), 105 (2024).
16. Alessandrini, L. et al. Cancer targeted therapy strategy: The pathologist's perspectives. *Curr. Cancer Drug Tar.* **18** (5), 410–420 (2018).
17. Jonna, S. & Subramaniam, D. S. Molecular diagnostics and targeted therapies in non-small cell lung cancer (NSCLC): An update. *Discov. Med.* **27** (148), 167–170 (2019).
18. Herrera-Juarez, M., Serrano-Gomez, C., Bote-de-Cabo, H. & Paz-Ares, L. Targeted therapy for lung cancer: Beyond EGFR and ALK. *Cancer-Am. Cancer Soc.* **129** (12), 1803–1820 (2023).
19. Zhou, Z. & Li, M. Targeted therapies for cancer. *BMC Med.* **20** (1), 90 (2022).
20. Yang, L. & Wang, Y. Malignant pleural effusion diagnosis and therapy. *Open. Life Sci.* **18** (1), 20220575 (2023).
21. Karpathiou, G. et al. Peoc'H M: pericardial and pleural metastases: Clinical, histologic, and molecular differences. *Ann. Thorac. Surg.* **106** (3), 872–879 (2018).
22. Pawlak, C. A. et al. Diagnosis of malignant pericarditis: A single centre experience. *Kardiol Pol.* **70** (11), 1147–1153 (2012).
23. Azarbal, A. & LeWinter, M. M. Pericardial effusion. *Cardiol. Clin.* **35** (4), 515–524 (2017).
24. Adler, Y. et al. Cardiac tamponade. *Nat. Rev. Dis. Primers.* **9** (1), 36 (2023).
25. Zhao, Y. et al. Current status of and progress in the treatment of malignant pleural effusion of lung cancer. *Front. Oncol.* **12**, 961440 (2022).

Acknowledgements

We sincerely thank Professor Li Jiang for her invaluable guidance and support throughout the development of this work. Her insightful suggestions and encouragement have greatly contributed to the completion of this manuscript.

Author contributions

Xin Hu, Shiqiao Zhao: Substantial contributions to conception and design, material preparation, data collection, data analysis, and drafting the article; Yanlun Li, Yiluo Heibi and Hang Wu: material preparation, data collection, reviewing the article; Yongjie Jiang: Substantial contributions to conception and design, data analysis, reviewing the article, and providing critical feedback. All authors read and approved the manuscript.

Funding

No funding was received for this study.

Declarations

Competing interests

The authors declare no competing interests.

Ethics approval and consent to participate

The study was approved by the Medical Ethics Committee of Affiliated Hospital of North Sichuan Medical College with file number 2023ER423-1. Informed consent was waived due to the retrospective nature of the study, where data was sourced from the hospital's electronic medical records and electronic health check databases.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-93842-4>.

Correspondence and requests for materials should be addressed to Y.J.

Reprints and permissions information is available at 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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

© The Author(s) 2025