Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Lung Ultrasound-Based Model for Early Recognition of Severe Mycoplasma pneumoniae Pneumonia in Children

Jiaojiao Gao

gaojiaojiao_bfh@163.com

Beijing Friendship Hospital

Zhengyang Han

Beijing Friendship Hospital

Jiale Gao

Beijing Friendship Hospital

Research Article

Keywords: Mycoplasma pneumoniae pneumonia, severe pneumonia, machine learning, stacking ensemble, lung ultrasound

Posted Date: November 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11362727/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Lung Ultrasound-Based Model for Early Recognition of Severe Mycoplasma pneumoniae Pneumonia in Children

Short running title: Early Recognition Model for SMPP in Children

Authorship: Jiaojiao Gao[1#*]; Zhengyang Han[2#]; Jiale Gao[3]

1.Department of Pediatrics, Beijing Friendship Hospital, Capital Medical University, Beijing, China.

2.Clinical Epidemiology & EBM Unit, Beijing Friendship Hospital, Capital Medical University, Beijing, China.

3.Department of General Surgery, Beijing Friendship Hospital, Capital Medical University; State Key Lab of Digestive Health, National Clinical Research Center for Digestive Diseases, Beijing, China.

# These authors contributed equally.

*** Corresponding author details:**

Jiaojiao Gao

Department of Pediatrics, Beijing Friendship Hospital, Capital Medical University

No. 95 Yong'an Road, Xi'cheng District, Beijing 100050, China

Telephone: +86 15210634717

Fax: Not available

Email: gaojiaojiao_bfh@163.com

Abstract

Background

To develop and prospectively validate an interpretable machine-learning model integrating lung ultrasound and routinely available clinical data for early recognition of severe Mycoplasma pneumoniae pneumonia (SMPP) in children.

Methods

This single-center cohort study included children aged 1 month to 14 years hospitalized with confirmed Mycoplasma pneumoniae pneumonia (MPP) at Beijing Friendship Hospital, Capital Medical University, from January 2018 to June 2025. A retrospective cohort of 1,696 children was randomly stratified 7:3 into a training set (n=1,187) and an internal validation set (n=509); an independent prospective cohort of 233 children was used for validation. Clinical, laboratory, and lung ultrasound variables available within 24 hours of admission were considered. Integrated feature ranking and recursive feature elimination selected predictors. Ten individual machine-learning models and stacking ensembles were evaluated using discrimination, calibration, decision-curve analysis, and SHapley Additive exPlanations.

Results

The retrospective cohort comprised 1,256 children with MPP and 440 with SMPP; the prospective cohort comprised 117 and 116 children, respectively. Twelve predictors were selected, including shred sign, lung ultrasound score, fever duration, age, inflammatory indices, and routine biochemical variables. The RF+XGBoost stacking model achieved the best internal-validation performance (AUC, 0.934; 95% CI, 0.920-0.949; accuracy, 0.848; sensitivity, 0.802; specificity, 0.884; F1 score, 0.821). In prospective validation, its AUC was 0.841 (95% CI, 0.786-0.887),

exceeding that of RF (AUC, 0.743; 95% CI, 0.678-0.802; McNemar P=0.0031). At the prespecified threshold of 0.497, accuracy, sensitivity, specificity, and F1 score were 0.755, 0.776, 0.768, and 0.759, respectively. The prospective Brier score was 0.167; the significant Hosmer-Lemeshow test indicated calibration deviation requiring recalibration before broad deployment.

Conclusions

An interpretable 12-feature RF+XGBoost stacking model combining lung ultrasound with routine clinical data showed strong discrimination for early SMPP recognition in prospective validation. Further multicenter validation and prospective recalibration are required before broad clinical deployment.

Trial registration:

The study was retrospectively registered at ClinicalTrials.gov (NCT07333833).

Keywords:

Mycoplasma pneumoniae pneumonia; severe pneumonia; machine learning; stacking ensemble; lung ultrasound

Background

Mycoplasma pneumoniae is an important cause of pediatric community-acquired pneumonia(1-3), and SMPP is associated with a prolonged disease course, heterogeneous clinical manifestations, increased risk of pulmonary complications, and extra-pulmonary organ involvement(4). In recent years, the incidence of SMPP has shown an upward trend, potentially driven by widespread transmission, recurrent infection, antimicrobial resistance, and mixed infections, posing an increasing threat to child health(5, 6).

The early identification of SMPP is challenging because its clinical manifestations and radiologic patterns often overlap with those of viral or mixed respiratory infections, and disease trajectories can differ substantially across children at presentation(7, 8). In routine practice, severity assessment still relies heavily on clinician judgment, supported by imaging and non-specific laboratory tests. However, commonly used inflammatory markers such as white blood cell (WBC), C-reactive protein (CRP), and procalcitonin (PCT) may be influenced by age, timing of sampling, prior antibiotic exposure, and co-infections, and therefore provide limited discriminatory value for SMPP risk stratification(9).

In addition, access to timely high-quality imaging can vary, and repeated radiography raises concerns regarding radiation exposure in children. Lung ultrasound (LUS) offers a rapid, bedside, radiation-free tool that can be repeated dynamically to evaluate lung involvement (e.g., consolidation, interstitial syndrome, and pleural effusion) and may better capture early changes in disease severity(3, 10). Nevertheless, these findings are not routinely integrated into standardized early warning approaches for SMPP. These limitations underscore the need for an objective, quantitative, and clinically practical early severity-recognition framework

that combines routine clinical and laboratory data with ultrasound-derived information to improve early SMPP identification.

Machine learning (ML), which leverages routinely collected electronic medical record data, provides a practical approach to integrate heterogeneous variables, thereby improving early risk prediction for SMPP. Moreover, ensemble learning, particularly stacking, which combines multiple complementary base learners through a metamodel, can exploit the strengths of different algorithms, reduce model variance and overfitting, and often yield more robust and generalizable performance than any single model, especially in clinically noisy, high-dimensional datasets(11, 12).

This study aimed to develop and prospectively validate an interpretable machine learning–based model using clinical variables available within 24 hour of hospital admission, facilitating early severity recognition and individualized management of pediatric SMPP.

Methods

Study design and participants

We conducted a single-center cohort study at Beijing Friendship Hospital, Capital Medical University, including children aged 1 month to 14 years hospitalized with confirmed MPP between January 2018 and June 2025. The study was approved by Ethics Committee of Beijing Friendship Hospital, Capital Medical University (2025-P2-501-01) and registered at ClinicalTrials.gov (NCT07333833). Retrospective data collection (January 2018–May 2024) was granted a waiver of informed consent, prospective validation (June 2024–June 2025) required written informed consent from parents/guardians.

Eligibility criteria were: (1) age 1 month–14 years, (2) clinical and radiographic evidence of community-acquired pneumonia, (3) laboratory-confirmed MP infection, and (4) hospital stay ≥72 hours. We excluded children with major pre-existing cardiopulmonary, metabolic, or immunodeficiency disorders; co-infections; or missing critical early data (eg, no imaging within 24 hours of admission).

Outcome definition

Patients were classified as having mild MPP (MPP group) or severe MPP (SMPP group) according to the guideline-based criteria(13, 14). SMPP was defined as a confirmed MPP meeting the criteria for severe community-acquired pneumonia. Any of the following conditions qualified as SMPP: poor general condition, refusal to feed or signs of dehydration, altered consciousness, markedly increased respiratory rate, cyanosis, respiratory distress, oxygen saturation ≤ 0.92, multilobar involvement or ≥2/3 of a single lobe involved, pleural effusion, or extrapulmonary complications.

Data collection

We extracted the earliest available clinical, laboratory, and imaging data within 24 h of admission. Clinical variables: disease course information and vital signs. Laboratory variables: complete blood count, biochemistry, coagulation, CRP, PCT, and derived indices(15), including systemic immune inflammation index (SII) = platelets× neutrophils/lymphocytes, pan-immune inflammation value (PIV) = (neutrophils×platelets×monocytes) / lymphocytes, neutrophil-to-lymphocyte ratio (NLR), inflammation-related ratio (IRR) =

(neutrophils×monocytes)/(lymphocytes×eosinophils), coagulation-related ratio (CRR) = platelets/platelet distribution width, anemia-related ratio (ARR) = hemoglobin×mean corpuscular hemoglobin concentration/RDW-CV. Lymphocyte-to-monocyte ratio (LMR), platelet-to-lymphocyte ratio (PLR), inflammatory burden index (IBI) = CRP× neutrophil/lymphocyte ratio, neutrophil-to-platelet ratio (NPR), and CRP-to-lymphocyte ratio (CLR). Imaging variables included lung ultrasonography, chest radiography, and CT.

Lung Ultrasound Examination

LUS was performed using a LOGIQ E9 ultrasound system equipped with a 5-14 MHz linear probe. All examinations followed a standardized 12-zone scanning protocol. Each hemithorax was divided into anterior, lateral, and posterior regions, and each region was further divided into upper and lower zones. The probe was placed longitudinally and transversely along the intercostal spaces to evaluate pleural-line abnormalities, B-lines, subpleural consolidation, and shred sign (pleural effusion sign was excluded from the final predictor set to avoid incorporation bias, as pleural effusion is part of the SMPP outcome definition). The shred sign was defined as an irregular, shredded interface between consolidated and aerated lung tissue, indicative of severe alveolar infiltration at the consolidation border. Each LUS examination was performed by an experienced pediatric sonographer, and the report was subsequently reviewed by a senior sonographer.

Data preprocessing and missing data

Continuous variables were maintained on the original scale. Text variables were encoded into

numeric or categorical formats. The binary outcome is encoded as SMPP = 1 and MPP = 0. After expert review and univariable screening, 34 candidate variables entered iterative collinearity assessment. Two rounds of variance inflation factor screening (VIF >15 followed by VIF >10) retained 18 variables. Expert review then reinstated five clinically important variables, yielding 23 candidates for integrated ranking. Missing values were imputed according to variable type, and all preprocessing parameters were derived within the model-development data to reduce information leakage.

Feature selection

The 23 candidate variables were ranked by their average importance rank across ten algorithms (LR, SVM, RF, GBDT, KNN, GNB, DT, BPNN, LightGBM, and XGBoost). Recursive feature elimination (RFE) then evaluated subsets of increasing size. The smallest subset with an internal-validation AUC within 99% of the maximum AUC was selected. The final predictor set comprised 12 variables: shred sign, CLR, LUS, SII, LMR, NLR, ALT, age, AST/ALT ratio, fever duration, CRP, and IBI (Figure 3A).

Model development and validation

The 1,696 retrospective cases were randomly stratified in a 7:3 ratio into a training set (n=1,187) and an internal validation set (n=509). The prospective cohort (n=233) was retained for independent prospective validation and was not used for model development or feature selection.

Ten individual ML models (LR, SVM, RF, GBDT, KNN, GNB, DT, BPNN, LightGBM, and

XGBoost) and candidate stacking ensembles were evaluated. Performance was assessed using AUC, accuracy, sensitivity, specificity, PPV, NPV, and F1 score. The Youden index identified the prespecified classification threshold. Weighted Borda ranking integrated AUC, sensitivity, and model complexity to identify the most balanced model.

The final RF+XGBoost stacking model was compared with leading individual models in the internal and prospective validation cohorts. Paired discrimination was compared with the DeLong test where applicable; prospective classification performance at the prespecified threshold was summarized by a confusion matrix and compared using McNemar's test.

Calibration was assessed using calibration curves, the Brier score, and the Hosmer-Lemeshow test. Clinical utility was assessed using decision curve analysis. SHapley Additive exPlanations (SHAP) were used to describe global feature importance, feature-level effects, and individual predictions.

Clinical application

Decision curve analysis quantified net benefit across threshold probabilities. Using the prespecified threshold of 0.497, prospective patients were stratified into low- and high-risk groups for SMPP. A web-based calculator was implemented to support bedside research use.

Statistical analysis

Continuous variables are presented as mean ± SD or median [IQR] and categorical variables are presented as n (%). Group comparisons were performed using the Mann-Whitney U test or χ^2 test. Analyses were performed using SPSS 29.0 and Python 3.12.

Results

Study population

A total of 2,016 children were screened for the retrospective cohort. After exclusion for co-infection with other pathogens (n=203), hospital stay <3 days (n=85), cardiopulmonary comorbidity (n=8), or missing key data within 24 h of admission (n=24), 1,696 cases were included (MPP, n=1,256; SMPP, n=440). The prospective cohort included 233 children (MPP, n=117; SMPP, n=116). The study workflow, participant selection, and cohort allocation are shown in Figures 1 and 2.

Baseline characteristics of the retrospective and prospective cohorts are summarized in Table 1. In the retrospective cohort, children with SMPP had higher inflammatory-marker levels and more pronounced lung ultrasonographic abnormalities than those with MPP.

Feature selection

RFE informed by integrated ranking selected 12 predictors for SMPP classification: shred sign, CLR, LUS, SII, LMR, NLR, ALT, age, AST/ALT ratio, fever duration, CRP, and IBI (Figure 3A).

Model development and evaluation

In the internal validation set, RF was the highest-performing individual model (AUC, 0.906; 95% CI, 0.877-0.935), followed by XGBoost (AUC, 0.896; 95% CI, 0.864-0.928) and LightGBM (AUC, 0.893; 95% CI, 0.866-0.919) (Figure 3B; Table S3).

Among stacking ensembles, RF+XGBoost achieved the highest internal-validation AUC

(0.934; 95% CI, 0.920-0.949), with accuracy 0.848, sensitivity 0.802, specificity 0.884, and F1 score 0.821. It ranked first by weighted Borda score (2.80), and outperformed RF, XGBoost, LightGBM, GBDT, and SVM in pairwise AUC comparisons (all $P<0.01$) (Figure 3C-D; Tables S4-S6).

Prospective validation

In the independent prospective cohort (n=233), the RF+XGBoost stacking model achieved an AUC of 0.841 (95% CI, 0.786-0.887), exceeding that of RF (AUC, 0.743; 95% CI, 0.678-0.802). McNemar's test indicated a significant difference in classification performance ($P=0.0031$) (Figure 4A; Table S7).

At the prespecified threshold of 0.497, the stacking model yielded accuracy 0.755, sensitivity 0.776, specificity 0.768, and F1 score 0.759. The confusion matrix showed 86 true negatives, 31 false positives, 26 false negatives, and 90 true positives (Figure 4B-C).

Decision curve analysis showed greater net benefit for the stacking model than for RF, treat-all, and treat-none strategies across a broad range of clinically relevant threshold probabilities (Figure 4D). The prospective Brier score was 0.167; the Hosmer-Lemeshow test was significant ($\chi^2=21.324$, $P=0.006$), indicating calibration deviation and the need for prospective recalibration before broad clinical deployment. The 1,000-bootstrap bias-corrected calibration curve had a mean absolute error of 0.091 (Figure 4E).

Model interpretability

SHAP analysis of the final 12-feature stacking model identified shred sign as the dominant

contributor, followed by LUS, fever duration, age, ALT, SII, and CLR. The SHAP summary and dependence plots showed the direction and magnitude of each predictor's contribution to the estimated SMPP probability (Figure 5A-C; Figure S1).

Risk stratification and web tool

Using the 0.497 threshold, the prospective high-risk group (n=130) had an SMPP event rate of 73.1% (95/130), compared with 20.4% (21/103) in the low-risk group (risk ratio, 3.58; 95% CI, 2.41-5.32; P<0.001). The web-based calculator was implemented to facilitate bedside research use (Figure 1; Figure S2) (URL: https://children-smpp-recoginition.online/).

Discussion

Based on a single-center cohort, this study developed an interpretable RF+XGBoost stacking model incorporating clinical, laboratory, and lung ultrasound variables for early SMPP recognition. The model achieved an internal-validation AUC of 0.934 and a prospective-validation AUC of 0.841, using 12 predictors led by shred sign and LUS. This tool is intended as a cross-sectional decision-support aid at admission rather than a longitudinal prognostic instrument.

Although MPP is generally considered a self-limiting disease, its incidence has shown an upward trend in recent years(1). Approximately 12% of MPP patients progress to SMPP, often accompanied by pulmonary consolidation, pleural effusion, necrotizing pneumonia, and extrapulmonary complications involving the cardiac and nervous systems, seriously threatening patient prognosis and significantly increasing the medical burden(16). Therefore,

261 constructing a precise prediction model is crucial for guiding clinical decisions such as
262 whether to initiate early corticosteroid therapy or bronchoscopic intervention(17). Compared
263 with previous studies, this study made important improvements in predictor selection and
264 model construction methods. A severe pneumonia prediction model using only routine
265 biomarker testing achieved similarly high predictive efficacy (AUC, 0.839). However, their
266 study population covered all types of pneumonia patients rather than specifically targeting
267 MPP, potentially diluting specific feature weights(18). Previous dynamic nomogram studies
268 on SMPP prediction, while achieving high AUCs, were mostly based on linear models such as
269 logistic regression, and relied on CT or X-ray examinations to assess lung lesions(10, 19). In
270 contrast, this study introduced ensemble machine-learning algorithms capable of capturing
271 nonlinear relationships, significantly improving the model's fitting ability. More critically, we
272 incorporated radiation-free, bedside-repeatable LUS indices into the model, replacing
273 radiation-risky CT scans, which have high application value in the pediatric population.
274 Furthermore, another study based on interpretable machine learning showed good
275 performance, but its core predictors relied on non-routine laboratory tests, such as
276 S100A8/A9, limiting its widespread adoption in primary hospitals(20). Conversely, this study
277 utilized only complete blood count, CRP, and bedside ultrasound data, greatly enhancing the
278 clinical accessibility and practicality of the model.
279 The final model emphasized both local pulmonary involvement and systemic inflammation.
280 This study identified CRP and NLR as important predictors reflecting the intense
281 inflammatory cascade and immune imbalance in children with SMPP. The progression to
282 severe MPP is often accompanied by a "cytokine storm" and excessive cell-mediated immune

lung injury (15, 21). A decreased LMR further suggests a decline in host immune defense capabilities and abnormal activation of the monocyte-macrophage system (22). These inflammation-related indices are positively correlated with pneumonia severity. Notably, lung ultrasound features, particularly the LUS score and the shred sign, had extremely high weights in this model. The shred sign typically indicates irregularity at the edge of lung consolidation and severe infiltration of the surrounding lung tissue, serving as a direct reflection of lung lesion severity(23). Combining these imaging features with systemic inflammation indices allows the model to comprehensively assess the condition from two dimensions: "local anatomical changes" and "systemic immune status," thereby outperforming single-dimension predictors.

Regarding the choice of imaging examination, this study did not include chest CT and X-ray results in the final model but instead chose lung ultrasound. Although CT is the "gold standard" for diagnosing lung lesions and can more clearly display necrosis and air bronchograms, its high radiation dose and sedation requirements limit its routine application in children. While X-ray examination is widespread, it has low sensitivity for identifying early consolidation and small amounts of pleural effusion. In recent years, lung ultrasound has gradually become an important tool for pediatric pneumonia assessment owing to its radiation-free nature, real-time capability, and bedside accessibility. Multiple studies have indicated that the sensitivity of LUS in diagnosing pneumonia is significantly better than that of X-rays and even comparable to that of CT(23, 24). LUS can quantify the severity of alveolar filling, interstitial edema, and inflammation in real time by identifying specific signs such as diffuse B-lines, large-area consolidation, and dynamic air bronchograms. The results

of this study confirm that a model including LUS features can achieve extremely high prediction accuracy, indicating that in the early risk stratification of SMPP, LUS has the potential to serve as an alternative to CT or as a routine screening tool, ensuring diagnostic efficacy while maximizing patient safety(25).

This study has limitations. It was conducted at a single tertiary center, and broader multicenter external validation is needed. Calibration drift was observed in the prospective cohort, suggesting that periodic recalibration may be required when epidemiology or case mix shifts. To facilitate safe real-world deployment, we recommend that adopting institutions establish a prospective monitoring pipeline to track the model's positive predictive value and calibration slope on a rolling basis. When a meaningful decline in calibration performance is detected, local recalibration using recent institutional data should be performed. Specifically, an annual pre-season recalibration, conducted before the peak incidence season of Mycoplasma pneumoniae, is recommended to maintain model accuracy across temporal and geographic shifts in pathogen virulence and host immunity. In addition, LUS is operator-dependent, and standardized training and quality control are important for implementation. Formal inter-observer reliability assessment (e.g., ICC or Kappa) should be incorporated in future prospective studies.

Conclusions

In conclusion, this study successfully developed and validated a stacking ensemble machine learning model incorporating lung ultrasound features and routine biomarkers for the early recognition of SMPP in children. The model demonstrated strong discrimination and

clinically useful sensitivity in prospective validation, offering a practical, radiation-free approach to support early risk stratification at admission.

Declarations

Ethics approval and consent to participate

The study was approved by the Ethics Committee of Beijing Friendship Hospital, Capital Medical University (2025-P2-501-01). For the retrospective data collection conducted from January 2018 to May 2024, the requirement for informed consent was waived by the Ethics Committee. For the prospective validation conducted from June 2024 to June 2025, written informed consent to participate in the study was obtained from the parents or legal guardians of all participants.

Consent for publication: Not applicable

Availability of data and materials: The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests: The authors declare that they have no competing interests.

Funding：All phases of this study were supported by the seed project no. YYZZ202416 from the Beijing Friendship Hospital, Capital Medical University.The funders had no role in the design and conduct of the study.

Authors' contributions: Jiaojiao Gao conceptualized and designed the study, supervised the study conduct and data collection, drafted the initial manuscript, and critically reviewed and revised the manuscript. Zhengyang Han developed the methodology, performed model development and validation, conducted the statistical analyses, and critically reviewed and revised the manuscript.

Jiale Gao curated the data and critically reviewed and revised the manuscript. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Acknowledgements: We appreciate the support of Junlei Xu for her writing assistance in this study.

References

1. Diaz H, Hersh L, Olson J, Shah SS, Hall M, Edens C. Mycoplasma pneumoniae Infections in Hospitalized Children - United States, 2018-2024. MMWR Morb Mortal Wkly Rep. 2025;74(23):394-400.

2. Xu M, Li Y, Shi Y, Liu H, Tong X, Ma L, et al. Molecular epidemiology of Mycoplasma pneumoniae pneumonia in children, Wuhan, 2020-2022. BMC Microbiol. 2024;24(1):23.

3. Reyes F, Conway A, Serrano C, Derde LPG, Dickson RP, Martin-Loeches I. Community-acquired pneumonia. Lancet. 2025;406(10517):2371-88.

4. Waites B, Xiao L, Liu Y, Balish MF, Atkinson TP. Mycoplasma pneumoniae from the Respiratory Tract and Beyond. Clin Microbiol Rev. 2017;30(3):747-809.

5. Kant R, Kumar N, Malik S, Everett D, Saluja D, Launey T, et al. Critical insights from recent outbreaks of Mycoplasma pneumoniae: decoding the challenges and effective interventions strategies. Int J Infect Dis. 2024;147:107200.

6. Global spatiotemporal dynamics of Mycoplasma pneumoniae re-emergence after COVID-19 pandemic restrictions: an epidemiological and transmission modelling study. Lancet Microbe. 2025;6(4):101019.

7. Yu A, Ran L, Sun X, Feng T. Significance of respiratory virus coinfection in children with Mycoplasma pneumoniae pneumonia. BMC Pulm Med. 2024;24(1):585.

8. Yan C, Xue GH, Zhao HQ, Feng YL, Cui JH, Yuan J. Current status of Mycoplasma pneumoniae infection in China. World J Pediatr. 2024;20(1):1-4.

9. Wang LP, Hu ZH, Jiang JS, Jin J. Serum inflammatory markers in children with Mycoplasma pneumoniae pneumonia and their predictive value for mycoplasma severity. World J Clin Cases. 2024;12(22):4940-6.

10. Li L, Guo R, Zou Y, Wang X, Wang Y, Zhang S, et al. Construction and Validation of a Nomogram Model to Predict the Severity of Mycoplasma pneumoniae Pneumonia in Children. J Inflamm Res. 2024;17:1183-91.

11. Zhao S, Wang J, Jin C, Zhang X, Xue C, Zhou R, et al. Stacking Ensemble Learning-Based [(18)F]FDG PET Radiomics for Outcome Prediction in Diffuse Large B-Cell Lymphoma. J Nucl Med. 2023;64(10):1603-9.

12. Fu Y, Liang X, Yang X, Li L, Meng L, Wei Y, et al. Stacking model framework reveals clinical biochemical data and dietary behavior features associated with type 2 diabetes: A retrospective cohort study. APL Bioeng. 2024;8(4):046111.

13. Association SGoRtSoPCM. Evidence-based guideline for the diagnosis and treatment of Mycoplasma pneumoniae pneumonia in children (2023). Pediatr Investig. 2025;9(1):1-11.

14. Association SGoRtSoPCM. [Guidelines for the management of community-acquired pneumonia in children (2024 revision)]. Zhonghua Er Ke Za Zhi. 2024;62(10):920-30.

15. Yang Z, Wang L, Hong B, He Z, Zhang Q, Shen T, et al. Inflammatory Burden Index as a Predictor of In-Hospital Mortality in Patients With Severe Fever With Thrombocytopenia Syndrome. J Med Virol. 2025;97(2):e70225.

16. Kutty PK, Jain S, Taylor TH, Bramley AM, Diaz MH, Ampofo K, et al. Mycoplasma pneumoniae Among Children Hospitalized With Community-acquired Pneumonia. Clin Infect Dis. 2019;68(1):5-12.

17. Gao L, Sun Y. Laboratory diagnosis and treatment of Mycoplasma pneumoniae infection in children: a review. Ann Med. 2024;56(1):2386636.

18. Cui W, Zhang X, Chen Y, Anthony D, Zhang G, Sheng Q, et al. Development and validation of an interpretable machine learning model using routine laboratory biomarkers to stratify severe pneumonia risk in young children. J Adv Res. 2025.

19. Zhang X, Sun R, Jia W, Li P, Song C. A new dynamic nomogram for predicting the risk of severe Mycoplasma pneumoniae pneumonia in children. Sci Rep. 2024;14(1):8260.

20. Xie S, Wu M, Shang Y, Tuo W, Wang J, Cai Q, et al. Development and validation of an early diagnosis model for severe mycoplasma pneumonia in children based on interpretable machine learning. Respir Res. 2025;26(1):182.

21. Li D, Gu H, Chen L, Wu R, Jiang Y, Huang X, et al. Neutrophil-to-lymphocyte ratio as a predictor of poor outcomes of Mycoplasma pneumoniae pneumonia. Front Immunol. 2023;14:1302702.

22. Enersen CC, Egelund GB, Petersen PT, Andersen S, Ravn P, Rohde G, et al. The ratio of neutrophil-to-lymphocyte and platelet-to-lymphocyte and association with mortality in community-acquired pneumonia: a derivation-validation cohort study. Infection. 2023;51(5):1339-47.

23. Padrao EMH, Caldeira Antonio B, Gardner TA, Miyawaki IA, Gomes C, Riceto Loyola Junior JE, et al. Lung Ultrasound Findings and Algorithms to Detect Pneumonia: A Systematic Review and Diagnostic Testing Meta-Analysis. Crit Care Med. 2025;53(11):e2271-e81.

24. Orso D, Fonda F, Bove T. Lung ultrasound vs chest radiography for pneumonia diagnosis in critically ill patients in the intensive care unit: a systematic review and Bayesian network diagnostic meta-analysis. EClinicalMedicine. 2025;89:103553.

25. Na F, Wang Y, Tian C, Zhang H, Jing Y, Ke X, et al. Diagnostic accuracy of various imaging modalities for children with pneumonia: a systematic review and network meta-analysis. Eur Respir Rev. 2025;34(178).

429 **Table 1: Baseline Characteristics of Retrospective and Prospective Cohorts**

		Retrospective cohort (N=1696)			**Prospective cohort (N=233)**		
Variable	**Categories**	**MPP group (N=1256)**	**SMPP group (N=440)**	**P-value**	**MPP group (N=117)**	**SMPP group (N=116)**	**P-value**
Basic characteristic							
Gender[a]				0.897			0.649
	Female	633 (50.4%)	224 (50.9%)		65 (55.6%)	60 (51.7%)	
	Male	623 (49.6%)	216 (49.1%)		52 (44.4%)	56 (48.3%)	
Age (months)[c]		75 (44-108)	91.5 (67-117.25)	**<0.001**	82 (47-116)	88 (55-121)	0.179
Age (years)[b]		6 (3-9)	8 (5-10)	**<0.001**	7 (5-8)	7 (5-8)	0.305
Clinical symptom							
Paroxysm duration[b]		7 (5-9)	7 (5.840-9)	0.185	7 (5-9)	7 (5-8)	0.289
Cough duration[b]		6 (4-8)	6 (5-8)	0.070	6 (4-7)	6 (4.75-7)	0.657
Fever duration[b]		5 (4-7)	6 (5-8)	**<0.001**	5 (3-7)	6 (5-8)	**<0.001**
Peak body		39.25 (39-39.7)	39.45 (39-39.93)	**<0.001**	39 (38.7-39.4)	39.2 (38.9-39.8)	**0.009**

		Retrospective cohort (N=1696)			Prospective cohort (N=233)		
Variable	**Categories**	**MPP group (N=1256)**	**SMPP group (N=440)**	**P-value**	**MPP group (N=117)**	**SMPP group (N=116)**	**P-value**
temperature[b]							
Physical sign							
Dyspnea[a]				**<0.001**			1.000
	no	1181 (94%)	432 (98.2%)		116 (99.1%)	116 (100.0%)	
	yes	75 (6%)	8 (1.8%)		1 (0.9%)	0 (0.0%)	
Retraction sign[a]				0.344			1.000
	no	1242 (98.9%)	438 (99.5%)		117 (100.0%)	116 (100.0%)	
	yes	14 (1.1%)	2 (0.5%)		0	0	
Medical history							
Allergic history[a]				0.250			0.859
	no	1063 (84.6%)	383 (87%)		102 (87.2%)	103 (88.8%)	
	yes	193 (15.4%)	57 (13%)		15 (12.8%)	13 (11.2%)	
Allergic disease				0.614			0.704

		Retrospective cohort (N=1696)			Prospective cohort (N=233)		
Variable	**Categories**	**MPP group (N=1256)**	**SMPP group (N=440)**	**P-value**	**MPP group (N=117)**	**SMPP group (N=116)**	**P-value**
history[a]							
	no	1148 (91.4%)	398 (90.5%)		99 (84.6%)	95 (81.9%)	
	yes	108 (8.6%)	42 (9.5%)		18 (15.4%)	21 (18.1%)	
Birth history[a]				0.374			0.243
	not first birth	250 (19.9%)	97 (22.0%)		33 (28.2%)	42 (36.2%)	
	first birth	1006 (80.1%)	343 (78.0%)		84 (71.8%)	74 (63.8%)	
Delivery history[a]				**0.012**			0.796
	vaginal	673 (53.6%)	267 (60.7%)		85 (72.6%)	87 (75.0%)	
	Cesarean resection	583 (46.4%)	173 (39.3%)		32 (27.4%)	29 (25.0%)	
Premature birth[a]				1			0.286
	no	1206 (96.0%)	423 (96.1%)		111 (94.9%)	114 (98.3%)	
	yes	50 (4.0%)	17 (3.9%)		6 (5.1%)	2 (1.7%)	

		Retrospective cohort (N=1696)			Prospective cohort (N=233)		
Variable	**Categories**	**MPP group (N=1256)**	**SMPP group (N=440)**	**P-value**	**MPP group (N=117)**	**SMPP group (N=116)**	**P-value**
Feeding history[a]				**0.038**			0.434
	artificial	128 (10.2%)	27 (6.1%)		5 (4.3%)	4 (3.4%)	
	breast	787 (62.7%)	285 (64.8%)		81 (69.2%)	89 (76.7%)	
	mixed	341 (27.1%)	128 (29.1%)		31 (26.5%)	23 (19.8%)	
Lung Ultrasonography							
Air bronchogram sign[a]				**<0.001**			**0.011**
	no	704 (56.1%)	161 (36.6%)		53 (45.3%)	33 (28.4%)	
	yes	552 (43.9%)	279 (63.4%)		64 (54.7%)	83 (71.6%)	
Shred sign[a]				**<0.001**			0.847
	no	915 (72.9%)	161 (36.6%)		104 (88.9%)	105 (90.5%)	
	yes	341 (27.1%)	279 (63.4%)		13 (11.1%)	11 (9.5%)	
LUS[b]		14 (8-18)	15.5 (10-20)	**<0.001**	12 (7-17)	15 (10-18)	**0.001**
Blood test							

		Retrospective cohort (N=1696)			Prospective cohort (N=233)		
Variable	**Categories**	**MPP group (N=1256)**	**SMPP group (N=440)**	**P-value**	**MPP group (N=117)**	**SMPP group (N=116)**	**P-value**
Mean corpuscular hemoglobin concentration (MCHC)[b]		332 (325.75-339)	332 (325.75-339)	0.672	335.00 (325.00-341.00)	335.00 (327.00-341.00)	0.464
Red Cell Distribution Width-Coefficient of Variation (RDW-CV)[b]		12.4 (12-13)	12.3 (11.9-12.8)	**<0.001**	12.30 (11.90-12.70)	12.60 (12.00-12.90)	**0.043**
Red cell distribution width-standard deviation (RDW-SD)[b]		36 (0-38.4)	34.75 (0-38.2)	**0.003**	0.00 (0.00-39.50)	0.00 (0.00-38.62)	0.624
Mean platelet volume (MPV)[b]		8.95 (8.3-9.7)	9.1 (8.4-9.8)	**0.044**	8.90 (8.20-9.50)	8.85 (8.38-9.62)	0.266
Platelet distribution width		15.7 (13.2-16.3)	15.8 (15.40-16.2)	**0.031**	15.80 (15.60-16.10)	15.80 (15.38-16.10)	0.467

		Retrospective cohort (N=1696)			Prospective cohort (N=233)		
Variable	**Categories**	**MPP group (N=1256)**	**SMPP group (N=440)**	**P-value**	**MPP group (N=117)**	**SMPP group (N=116)**	**P-value**
(PDW)[b]							
Alanine aminotransferase (ALT)[b]		12 (10-16)	13 (10-19)	**<0.001**	12.00 (10.00-15.00)	12.00 (9.75-16.00)	0.980
Aspartate transaminase (AST)[b]		27.85 (23-35)	28.3 (23.7-34.74)	0.079	27.80 (22.60-31.50)	28.00 (22.88-31.30)	0.860
AST/ALT[b]		2.28 (1.77-2.77)	2.12 (1.60-2.64)	**0.002**	2.31 (1.66-2.86)	2.24 (1.64-2.93)	0.865
Crea[b]		33.75 (26.8-40.1)	39 (32.47-44.9)	**<0.001**	37.60 (33.30-45.40)	36.65 (32.60-44.05)	0.486
Urea[b]		2.79 (2.32-3.29)	2.8 (2.39-3.46)	0.160	3.19 (2.85-3.86)	3.20 (2.68-3.73)	0.279
K+[b]		4.30 (4.09-4.53)	4.28 (4.01-4.49)	**0.039**	4.29 (4.14-4.54)	4.27 (4.01-4.41)	0.060
Cl-[b]		103.12 (102-105)	103 (101.86-105)	**0.023**	104.00 (102.00-106.00)	104.00 (103.00-105.00)	0.380
Na+[b]		138.55 (137.14-140)	138.20 (136.9-139.6)	**0.005**	138.30(136.50-139.50)	138.10 (136.13-139.50)	0.477

		Retrospective cohort (N=1696)			Prospective cohort (N=233)		
Variable	**Categories**	**MPP group (N=1256)**	**SMPP group (N=440)**	**P-value**	**MPP group (N=117)**	**SMPP group (N=116)**	**P-value**
Ca[c]		2.34 (2.26-2.43)	2.35 (2.25-2.44)	0.402	2.42 (2.31-2.53)	2.37 (2.25-2.49)	**0.005**
Lactate dehydrogenase (LDH)[b]		293 (254.75-341)	313.39 (267-370)	**<0.001**	289.00 (250.00-308.00)	289.00 (267.75-331.25)	0.196
Glucose[b]		4.91 (4.57-5.2)	4.96 (4.64-5.34)	**0.006**	4.77 (4.45-4.96)	4.91 (4.65-5.21)	0.005
CO_2[b/c]		21.2 (19.8-22.6)	22 (20.5-23.3)	**<0.001**	21.42 ± 2.01	21.85 ± 2.04	0.106
Creatine Kinase (CK)[b]		81 (56-118)	84 (57-128.25)	0.176	71.00 (53.00-99.00)	77.00 (54.75-105.25)	0.368
Anion Gap[b]		16.4 (14.7-18.3)	16.9 (15.6-18.5)	**<0.001**	16.80 (15.90-18.50)	16.50 (15.53-17.80)	0.197
Cardiac Troponin I (cTnI)[b]		0.00 (0.00-0.00)	0.00 (0.00-0.00)	**<0.001**	0.00 (0.00-0.00)	0.00 (0.00-0.00)	0.317
C-reactive protein (CRP)[b]		11.9 (4.78-24)	15.38 (7.07-33.01)	**<0.001**	6.99 (3.01-14.68)	13.16 (5.90-21.48)	**<0.001**
SII[b]		475.95 (288.56-785.	580.85 (403.16-971.	**<0.001**	435.87 (286.13-668.47)	525.00 (347.90-881.	**0.010**

		Retrospective cohort (N=1696)			Prospective cohort (N=233)		
Variable	**Categories**	**MPP group (N=1256)**	**SMPP group (N=440)**	**P-value**	**MPP group (N=117)**	**SMPP group (N=116)**	**P-value**
		03)	69)			25)	
PIV[b]		260.69 (134.26-456.95)	301.18 (172.7-559.81)	**<0.001**	226.41 (123.52-426.68)	294.27 (154.32-506.11)	**0.043**
NLR[b]		1.82 (1.07-2.76)	2.24 (1.52-3.24)	**<0.001**	1.59 (0.97-2.52)	1.94 (1.31-2.84)	**0.010**
IRR[b]		8.82 (3.71-26.16)	11.56 (4.59-29.51)	**0.020**	6.95 (3.57-16.45)	8.50 (4.03-18.62)	0.291
CRR[b]		18.67 (13.9-25.8)	17.99 (14.03-24.62)	0.258	18.76 (14.20-23.31)	19.61 (15.11-26.85)	0.189
ARR[b]		3446.8 (3164.13-3723.63)	3488 (3208.42-3751.53)	0.081	3529.77 (3300.00-3733.03)	3407.25 (3193.30-3702.88)	0.030
LMR[b]		4.59 (3.35-6.52)	3.9 (2.94-5.37)	**<0.001**	4.81 (3.30-6.14)	4.14 (3.26-5.62)	0.055
PLR[b]		108.08 (79.41-149.69)	138.2 (101.72-189.54)	**<0.001**	119.47 (85.71-150.86)	134.51 (101.38-174.47)	**0.013**
IBI[b]		20.72 (5.87-56.04)	34.79 (11.89-91.78)	**<0.001**	10.66 (3.47-31.61)	22.91 (8.34-55.00)	**<0.001**
NPR[b]		0.02	0.02	0.921	0.01 (0.01-0.02)	0.01	0.389

		Retrospective cohort (N=1696)			Prospective cohort (N=233)		
Variable	**Categories**	**MPP group (N=1256)**	**SMPP group (N=440)**	**P-value**	**MPP group (N=117)**	**SMPP group (N=116)**	**P-value**
		(0.01-0.02)	(0.01-0.02)			(0.01-0.02)	
CLR[b]		4.8 (1.54-10.83)	7.26 (2.97-19.35)	**<0.001**	3.31 (1.34-8.15)	5.97 (2.06-11.41)	**<0.001**

[a]Indicates the results of the chi-square test. [b]Indicates the results of the non-parametric test. [c]Indicates t-test results.

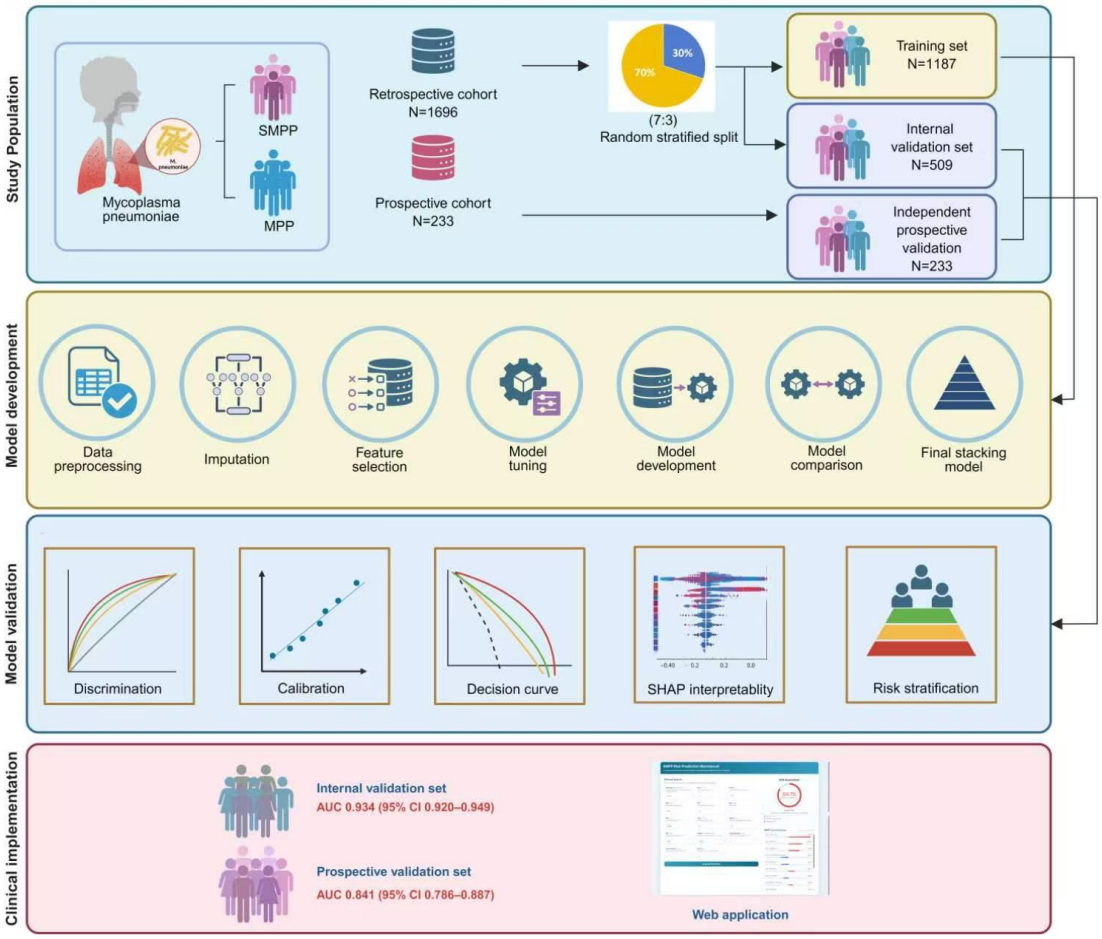


Figure 1. Study design, model development, validation, and clinical implementation. The retrospective cohort (n=1,696) was randomly split 7:3 into a training set (n=1,187) and internal validation set (n=509). The independent prospective cohort comprised 233 children.

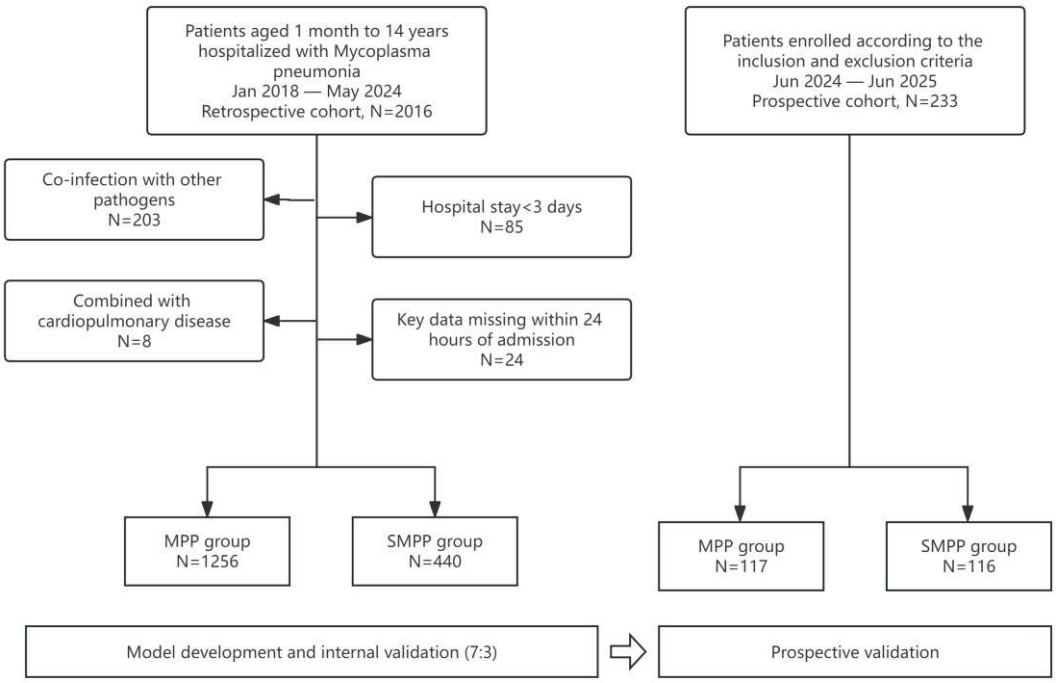


Figure 2. Participant selection. Of 2,016 retrospectively screened children, 1,696 were included after exclusions and were classified as MPP (n=1,256) or SMPP (n=440). The prospective cohort included 233 children (MPP, n=117; SMPP, n=116).

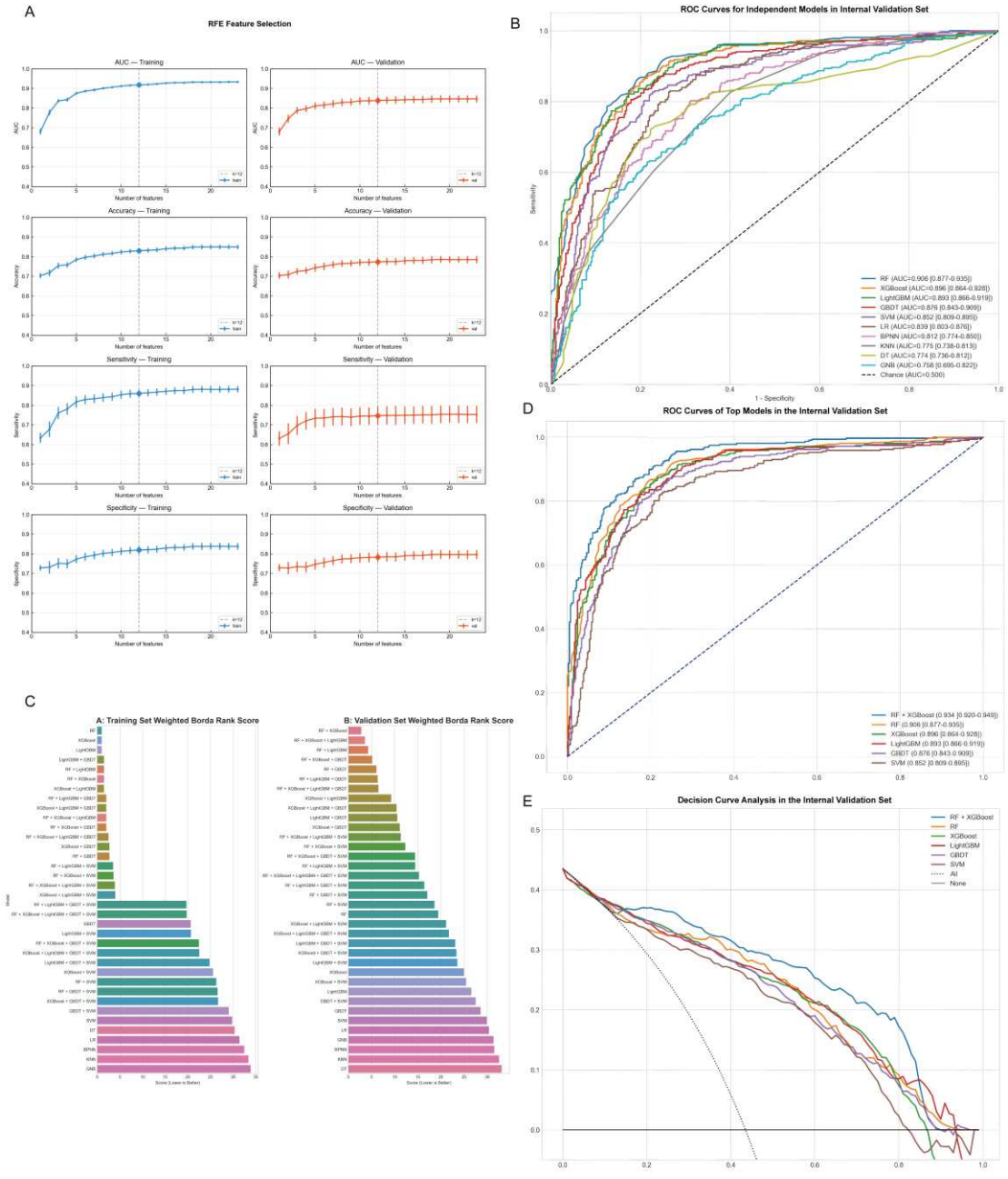


Figure 3. Feature selection and model evaluation in the internal validation set. (A) RFE performance curves for AUC, accuracy, sensitivity, and specificity; the selected subset contained 12 predictors. (B) ROC curves for 10 independent models. (C) Weighted Borda rankings in the training and internal validation sets. (D) ROC curves for the final RF+XGBoost stacking model and leading independent models. (E) Decision curve analysis of the leading models in the internal validation set. AUC, area under the curve; LR, logistic regression; SVM, support vector machine; RF, random forest; GBDT, gradient boosting

decision tree; KNN, k-nearest neighbors; GNB, Gaussian naive Bayes; DT, decision tree; BPNN, multilayer perceptron.

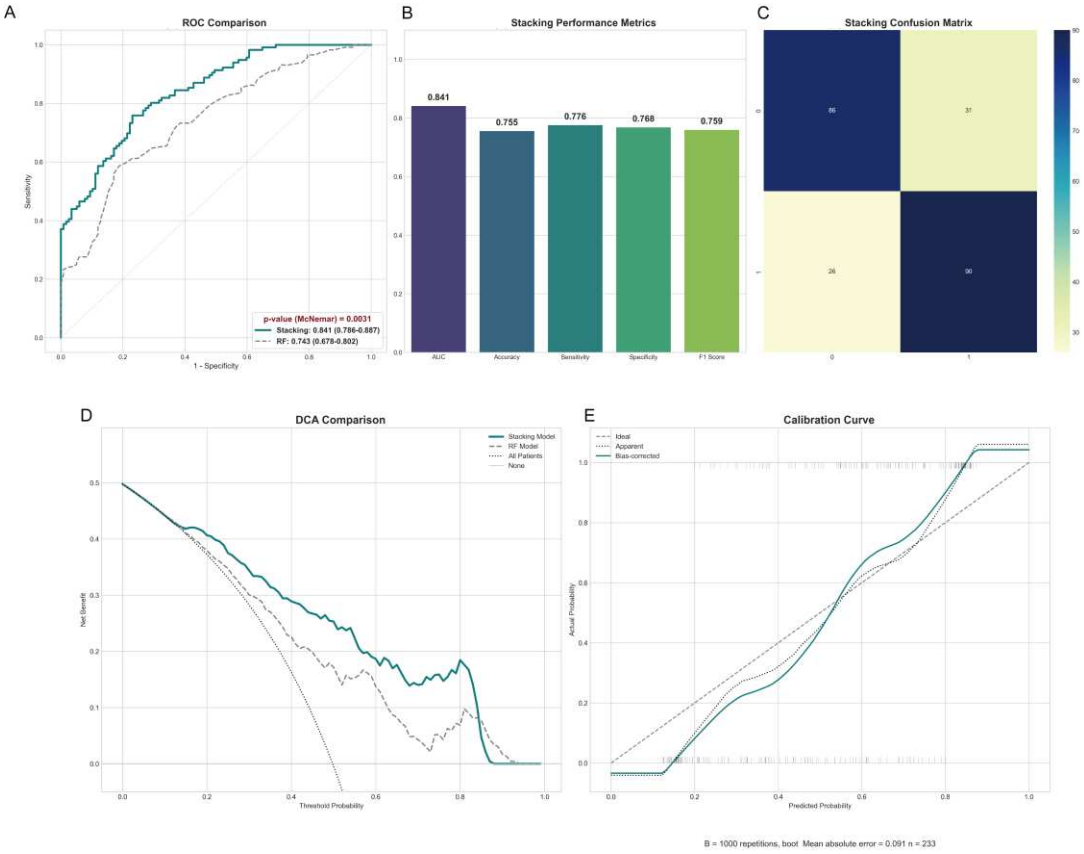


Figure 4. Prospective validation of the final RF+XGBoost stacking model. (A) ROC comparison of the stacking model (AUC, 0.841; 95% CI, 0.786-0.887) and RF (AUC, 0.743; 95% CI, 0.678-0.802); McNemar P=0.0031. (B) Stacking-model performance metrics. (C) Confusion matrix: true negatives=86, false positives=31, false negatives=26, and true positives=90. (D) Decision curve analysis. (E) Bias-corrected calibration curve based on 1,000 bootstrap resamples (mean absolute error=0.091).

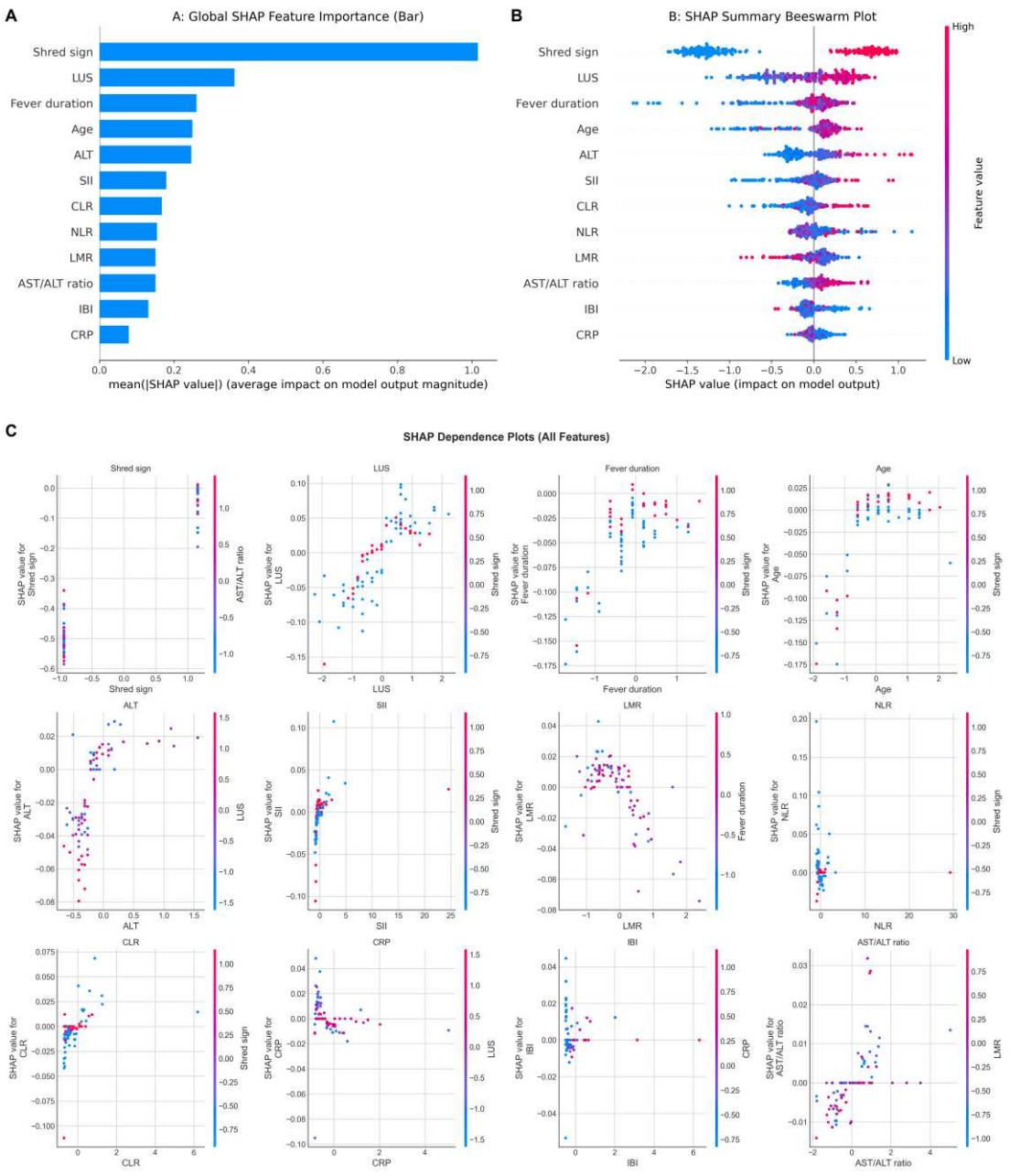


Figure 5. SHAP-based interpretability analysis of the 12-feature final stacking model. (A) Global feature importance based on mean absolute SHAP values. (B) SHAP beeswarm summary plot. (C) SHAP dependence plots for all predictors. Shred sign was the most influential predictor, followed by LUS, fever duration, age, ALT, SII, and CLR. SHAP, SHapley Additive exPlanations; LUS, lung ultrasound score; CRP, C-reactive protein; SII, systemic immune-inflammation index; CLR, CRP-to-lymphocyte ratio; LMR, lymphocyte-to-monocyte ratio; NLR, neutrophil-to-lymphocyte ratio; IBI, inflammatory

469 burden index.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementaryfigure.pdf
- supplementarytable.docx