Am J Transl Res 2026;18(2):1239-1248
www.ajtr.org /ISSN:1943-8141/AJTR0169439

*Original Article*

# A gradient boosting machine model for predicting prognosis in patients with acute respiratory distress syndrome

Yuji Liang[1], Yan Yang[2], Qixian Liang[1], Rucheng Liao[1], Ling Li[1], Qiuhua Yang[1]

[1]*Department of Critical Care Medicine, Qinzhou First People's Hospital, Qinzhou 535000, Guangxi, China;* [2]*Department of Internal Medicine, Qinzhou First People's Hospital, Qinzhou 535000, Guangxi, China*

Received October 20, 2025; Accepted January 14, 2026; Epub February 15, 2026; Published February 28, 2026

**Abstract:** Objectives: To develop a gradient boosting model for predicting the prognosis of patients with acute respiratory distress syndrome (ARDS), providing a data-driven reference for early identification of high-risk patients in clinical settings. Methods: This retrospective study analyzed the 28-day mortality in 307 ARDS patients treated at Qinzhou First People's Hospital between July 2023 and June 2024. Patients were divided into a mortality group (n=92) and a survival group (n=215) based on in-hospital death. Demographic characteristics, clinical variables, and biochemical parameters were collected. Univariate and multivariate logistic regression analyses were performed to identify independent predictors, which were subsequently used to construct a gradient boosting machine (GBM) model and a nomogram model. Model performance was evaluated with calibration curves and the area under the receiver operating characteristic (ROC) curve (AUC). Results: Logistic regression identified age, oxygenation index (OI), neutrophil-to-lymphocyte ratio (NLR), interleukin-8 (IL-8), and N-terminal pro-B-type natriuretic peptide (NT-proBNP) as independent prognostic factors for ARDS. In the GBM model, the relative importance of NT-proBNP, age, NLR, IL-8, and OI was ranked. The nomogram indicated that older age, lower OI, and higher levels of NLR, IL-8, and NT-proBNP were associated with poorer prognosis. The AUC values for the GBM model in the training and validation sets were 0.907 (95% CI: 0.866-0.947) and 0.887 (95% CI: 0.803-0.971), respectively, which surpassed the values of 0.866 (95% CI: 0.810-0.923) and 0.835 (95% CI: 0.733-0.937) for the Nomogram model. Conclusion: The 28-day mortality rate among ARDS patients was 29.97%, and was mainly associated with age, oxygenation index, NLR, IL-8, and NT-proBNP levels. A GBM model constructed using these factors showed good predictive performance, offering valuable data references for clinical identification of ARDS patients at a high-risk of poor prognosis.

**Keywords:** Acute respiratory distress syndrome, gradient ascender model, predictive performance, prognosis

## Introduction

Acute respiratory distress syndrome (ARDS) is a severe form of lung injury, which leads to increased vascular permeability and accumulation of protein-rich fluid in the alveoli [1, 2]. It is clinically manifested as progressive dyspnea and the failure of the body to respond to low oxygen levels [3]. Current therapeutic strategies for ARDS include lung-protective mechanical ventilation [4], supportive respiratory therapies, and extracorporeal membrane oxygenation (ECMO) in severe cases [5]. Despite advances in supportive care, the treatment efficacy remains limited, and the incidence and mortality of ARDS are still high [6]. Therefore, identifying reliable biological markers that can accurately reflect disease severity and predict patient prognosis is extremely important for improving clinical management and outcomes.

A prior study based on multivariate Cox regression analysis reported that the lactate-to-albumin ratio (LAR) is an independent predictor of 28-day mortality in patients with RDS; however, the reliance on a single indicator limits its generalizability and clinical applicability [7]. Other studies have developed mortality prediction models for ARDS using the Random Forest (RF) algorithm and identified platelet count and lactate as the strongest predictors of in-hospital mortality. However, these analyses relied on

https://doi.org/10.62347/NFOR8502

the Medical Information Mart for Intensive Care (MIMIC-III) and Telehealth Intensive Care Unit Collaborative Research databases (eICU-CRD), in which certain variables were not directly accessible, potentially introducing selection bias [8].

The Gradient Boosting Machine (GBM) algorithm is one of the most popular machine-learning method that integrates multiple weak classifiers to generate a robust predictive model. By iteratively optimizing the loss function, the GBM algorithm effectively reduces prediction errors and demonstrates good performance and interpretability in medical data analysis [9]. In this study, a GBM model incorporating multiple clinical and laboratory indicators was developed to predict the prognosis of patients with ARDS, aiming to facilitate the early identification of ARDS patients at high risk of poor prognosis.

## Materials and methods

### *General information*

Sample size estimation was performed based on the events-per-variable (EPV) principle. The literature reports that the in-hospital death rate of ARDS patients ranges from 32% to 51% [11]. Assuming a 28-day mortality rate of 40% in this study and the inclusion of 10 variables in the final GBM model, a sample size of 250 cases was required with an EPV of 10. Ultimately, a total of 307 patients were enrolled in this study. This study was approved by the Ethics Committee of Qinzhou First People's Hospital (No.2023083).

A retrospective analysis was conducted on data from 307 ARDS patients treated at Qinzhou First People's Hospital between July 2023 and June 2024. Patients were categorized into a mortality group (n=92) and a survival group (n=215) based on their 28-day mortality status. Additionally, an independent cohort of 143 ARDS patients treated at Qinzhou First People's Hospital from July 2024 to March 2025 were enrolled for external validation of both the GBM and nomogram models.

Inclusion criteria: (1) meets the Berlin diagnostic criteria for ARDS [10]; (2) age ≥18 years; (3) hospitalization time longer than 48 hours, and (4) informed consent is obtained from patients and families. Exclusion criteria: (1) Cardiogenic pulmonary edema; (2) Concurrent chronic pulmonary diseases such as pulmonary vascular disease or chronic obstructive pulmonary disease (COPD); (3) Pregnancy or concurrent multiple organ failure.

### *Information collection*

(1) Baseline demographic and clinical characteristics were collected, including sex, age, body mass index (BMI), history of underlying conditions (e.g., hypertension, diabetes, coronary heart disease).

(2) Clinical variables reflecting disease severity, acute physiological disturbance, and the treatment intensity were recorded, including symptom onset time, length of hospital stay, oxygenation index (OI), duration of mechanical ventilation, medication usage, and the Acute Physiology and Chronic Health Evaluation II (APACHE II) score.

(3) Biochemical indicators, including complete blood count, coagulation parameters, blood lactate (LAC), uric acid (UA), albumin (Alb), C-reactive protein (CRP), procalcitonin (PCT), inflammatory cytokines, and N-terminal pro-B-type natriuretic peptide (NT-proBNP) were collected to assess inflammatory stress, organ injury, and metabolic status of patients.

### *Statistical analysis*

Statistical analyses was performed using SPSS 27.0. Continuous variables with a normal distribution were presented as mean ± standard deviation (SD), and the differences between the groups were compared using the independent samples t-test. Count data were expressed as [n (%)], and comparisons were carried out using the chi-square test or Fisher's exact test, as appropriate. Significant factors affecting outcomes were identified through univariate and multivariate logistic regression analysis.

A GBM model was built with R 4.4.3. The receiver operating characteristic (ROC) curve was plotted, and the area under the curve (AUC) was calculated to evaluate the predictive performance of the GBM model for ARDS prognosis. Calibration curves were generated to evaluate model fit. A two-sided *P* value <0.05 was considered statistically significant.

**Table 1.** Univariate analysis of factors associated with patient prognosis

| | Total (n=307) | Death group (n=92) | Survival group (n=215) | $t/x^2$ | $P$ |
|---|---|---|---|---|---|
| Age (years) | 61.32±9.00 | 65.40±10.02 | 59.57±7.93 | 5.442 | <0.001 |
| Sex (Male/Female) | 218 (71.01)/89 (28.99) | 71 (77.17)/21 (22.83) | 147 (68.37)/68 (31.63) | 2.425 | 0.119 |
| BMI (kg/m$^2$) | 23.66±2.28 | 23.62±2.15 | 23.68±2.34 | -0.218 | 0.828 |
| Hypertension (Yes/No) | 148 (48.21)/159 (51.79) | 41 (44.57)/51 (55.43) | 107 (49.77)/108 (50.23) | 0.698 | 0.403 |
| Diabetes (Yes/No) | 128 (41.69)/179 (58.31) | 42 (45.65)/50 (54.35) | 86 (40.00)/129 (60.00) | 0.849 | 0.357 |
| Coronary Heart Disease (Yes/No) | 141 (45.93)/166 (54.07) | 36 (39.13)/56 (60.87) | 105 (48.84)/110 (51.16) | 2.444 | 0.118 |
| Medication Use (Yes/No) | 148 (48.21)/159 (51.79) | 46 (50.00)/46 (50.00) | 102 (47.44)/113 (52.56) | 0.169 | 0.681 |
| Smoking History (Yes/No) | 135 (43.97)/172 (56.03) | 49 (53.26)/43 (46.74) | 86 (40.00)/129 (60.00) | 4.599 | 0.032 |
| Alcohol Consumption (Yes/No) | 137 (44.63)/170 (55.37) | 38 (41.30)/54 (58.70) | 99 (46.05)/116 (53.95) | 0.586 | 0.444 |
| Length of Hospital Stay (d) | 39.48±9.87 | 39.24±10.33 | 39.58±9.69 | -0.278 | 0.781 |
| Onset Duration (h) | 63.31±11.22 | 64.47±10.98 | 62.81±11.31 | 1.183 | 0.238 |
| OI (mmHg) | 123.41±34.75 | 112.79±40.79 | 127.96±30.81 | -3.570 | <0.001 |
| APACHE II Score (points) | 20.36±6.02 | 24.17±6.18 | 18.73±5.16 | 7.962 | <0.001 |
| Duration of MV (h) | 187.34±24.59 | 187.41±24.51 | 187.30±24.68 | 0.036 | 0.971 |
| WBC (×10$^9$/L) | 14.71±4.11 | 13.87±5.17 | 15.06±3.52 | -2.341 | 0.020 |
| HGB (g/L) | 103.70±20.03 | 103.60±21.37 | 103.75±19.48 | -0.060 | 0.952 |
| PLT (×10$^9$/L) | 161.37±29.46 | 140.04±25.17 | 170.50±26.31 | -9.410 | <0.001 |
| NLR | 11.03±4.15 | 13.11±4.71 | 10.15±3.54 | 6.056 | <0.001 |
| RDW-S [(fl) | 46.46±6.99 | 46.11±6.31 | 46.61±7.27 | -0.578 | 0.564 |
| TG (mmol/L) | 1.93±0.51 | 1.97±0.66 | 1.91±0.44 | 0.864 | 0.389 |
| TC (mmol/L) | 3.32±0.96 | 3.30±1.02 | 3.32±0.93 | -0.200 | 0.842 |
| HDL-C (mmol/L) | 0.69±0.15 | 0.68±0.17 | 0.69±0.14 | -0.703 | 0.482 |
| LDL-C (mmol/L) | 1.78±0.45 | 1.79±0.51 | 1.78±0.42 | 0.150 | 0.881 |
| PT (s) | 14.96±3.12 | 14.90±3.38 | 14.99±3.01 | -0.229 | 0.819 |
| APTT (s) | 40.94±8.52 | 39.56±8.67 | 41.53±8.42 | -1.858 | 0.064 |
| TT (s) | 17.81±2.41 | 17.47±2.34 | 17.96±2.43 | -1.609 | 0.109 |
| FIB (g/L) | 4.22±1.15 | 4.33±1.38 | 4.17±1.03 | 1.143 | 0.254 |
| D-dimer (mg/L) | 1.13±0.34 | 1.38±0.35 | 1.03±0.28 | 9.325 | <0.001 |
| LAC (mmol/L) | 3.21±1.24 | 4.15±1.71 | 2.80±0.65 | 9.958 | <0.001 |
| UA (μmol/L) | 348.20±54.15 | 349.94±76.09 | 347.45±41.58 | 0.369 | 0.712 |
| Alb (g/L) | 28.02±6.24 | 28.06±6.26 | 28.00±6.25 | 0.068 | 0.946 |
| CRP (μg/L) | 67.70±11.57 | 68.36±14.90 | 67.42±9.84 | 0.650 | 0.516 |
| PCT (μg/L) | 13.34±14.35 | 19.72±24.62 | 10.61±3.42 | 5.322 | <0.001 |
| IL-6 (ng/L) | 112.32±23.96 | 128.13±25.76 | 105.55±19.63 | 8.377 | <0.001 |
| IL-8 (ng/L) | 121.36±29.10 | 135.06±28.11 | 115.49±27.56 | 5.664 | <0.001 |
| NT-proBNP (ng/L) | 3026.87±950.43 | 3755.83±926.49 | 2714.94±774.89 | 10.151 | <0.001 |

Note: APACHE II: Acute Physiology and Chronic Health Evaluation II; MV: Mechanical ventilation; OI: Oxygenation Index; WBC: White blood cell count; HGB: Hemoglobin; PLT: Platelet count; NLR: Neutrophil-to-lymphocyte ratio; RDW-S: Red cell distribution width; TG: Triglycerides; TC: Total cholesterol; HDL-C: High-density lipoprotein cholesterol; LDL-C: Low-density lipoprotein cholesterol; PT: Prothrombin time; APTT: Activated partial thromboplastin time; TT: Thrombin time; FIB: Fibrinogen; LAC: Lactate; UA: Uric acid; Alb: Albumin; CRP: C-Reactive protein; PCT: Procalcitonin; IL: Interleukin; NT-proBNP: N-terminal pro-B-Type natriuretic peptide.

## Results

### *Univariate analysis*

Results showed that patients in the mortality group were significantly older than those in the survival group (65.40±10.02 vs. 59.57 ±7.93) and had a significantly higher proportion of smoking history (53.26% vs. 40.00%). The mortality group also showed significantly higher APACHE II scores (24.17±6.18 vs. 18.73±5.16) and higher levels of NLR (13.11±4.71 vs. 10.15±3.54), D-dimer (1.38 ±0.35 vs. 1.03±0.28), LAC (4.15±1.71 vs. 2.80±0.65), PCT (19.72±24.62 vs. 10.61± 3.42), IL-6 (128.13±25.76 vs. 105.55±19.63), IL-8 (135.06±28.11 vs. 115.49±27.56), and NT-proBNP (3755.83±926.49 vs. 2714.94± 774.89) compared with the survival group (all $P$<0.05). In contrast, patients in the mortality group had significantly lower OI values (112.79±40.79 vs. 127.96±30.81), WBC counts (13.87±5.17 vs. 15.06±3.52), and lower PLT counts (140.04±25.17 vs. 170.50± 26.31) (P<0.05) than those in the survival group. No significant differences were observed in other indicators ($P$>0.05). The details are shown in **Table 1**.

**Table 2.** Logistic regression analysis

| | β | SE | Wald | *P* | *OR* | 95% CI |
|---|---|---|---|---|---|---|
| Age | -0.056 | 0.017 | 10.731 | <0.001 | 0.945 | 0.914-0.978 |
| Smoke | -0.232 | 0.297 | 0.609 | 0.435 | 0.793 | 0.443-1.419 |
| OI | 0.010 | 0.004 | 4.789 | 0.029 | 1.010 | 1.001-1.019 |
| APACHE II score | -0.016 | 0.027 | 0.349 | 0.555 | 0.984 | 0.933-1.038 |
| NLR | -0.079 | 0.04 | 3.872 | 0.049 | 0.924 | 0.854-1.000 |
| D-dimer | -0.761 | 0.464 | 2.688 | 0.101 | 0.467 | 0.188-1.160 |
| LAC | -0.178 | 0.173 | 1.058 | 0.304 | 0.837 | 0.597-1.175 |
| PCT | -0.022 | 0.04 | 0.296 | 0.586 | 0.979 | 0.905-1.058 |
| IL-6 | -0.012 | 0.007 | 3.431 | 0.064 | 0.988 | 0.975-1.001 |
| IL-8 | -0.016 | 0.005 | 9.075 | 0.003 | 0.984 | 0.973-0.994 |
| NT-proBNP | -0.001 | 0.000 | 10.74 | <0.001 | 0.999 | 0.999-1.000 |
| Constant | 11.112 | 1.883 | 34.808 | | | |

Note: OI: Oxygenation Index; APACHE II: Acute Physiology and Chronic Health Evaluation II; NLR: Neutrophil-to-lymphocyte ratio; LAC: Blood lactate; PCT: Procalcitonin; IL: Interleukin; NT-proBNP: N-terminal pro-B-type natriuretic peptide.

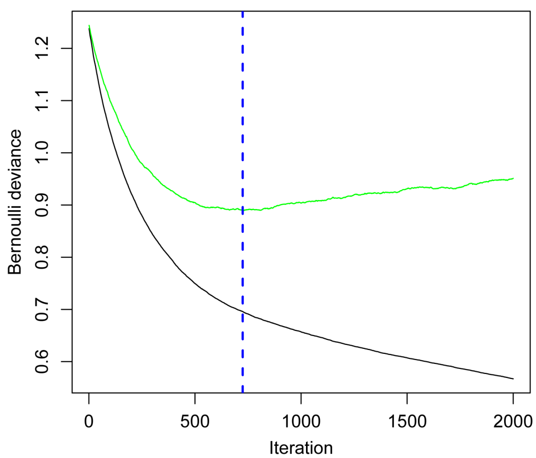


**Figure 1.** Iteration optimization of the GBM model. Note: GBM: gradient boosting machine.

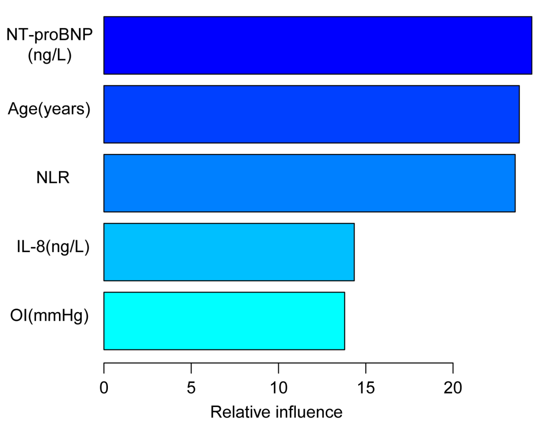


**Figure 2.** Relative importance of variables in the GBM model. Note: GBM: gradient boosting machine; NT-proBNP: N-terminal pro-B-type natriuretic peptide; NLR: Neutrophil-to-lymphocyte ratio; IL: Interleukin; OI: Oxygenation Index.

*Logistic regression analysis*

Patient outcome was used as the dependent variable (1=death, 0=survival). Variables with statistical significance in univariate analysis were entered as independent variables, including age, smoking history (1=yes, 0=no), oxygenation index, APACHE II score, WBC, PLT, NLR, D-dimer, LAC, PCT, IL-6, IL-8, and NT-proBNP. Multivariate logistic regression analysis identified age, OI, NLR, IL-8, and NT-proBNP as independent prognostic factors for ARDS patients ($P<0.05$). The logistic regression equation for predicting mortality: $P=11.112 - 0.056 \times \text{age} - 0.010 \times \text{OI} - 0.079 \times \text{NLR} - 0.016 \times \text{IL-8} - 0.001 \times \text{NT-proBNP}$ (**Table 2**).

*Development of a GBM model for predicting outcomes in ARDS patients*

Variables with $P<0.05$ in the logistic regression analysis were selected to construct the GBM model, with an optimal number of iterations of 725 (**Figure 1**). The importance of variables, ranked from highest to lowest, was NT-proBNP, age, NLR, IL-8, and OI (**Figure 2**). Predicted probabilities were obtained by summing the contributions of all trees and transforming them through the Sigmoid function. Calibration curves demonstrated good agreement between predicted and observed outcomes in both the training and testing sets, with the GBM prediction curves closely approximating the ideal reference line (**Figure 3**).

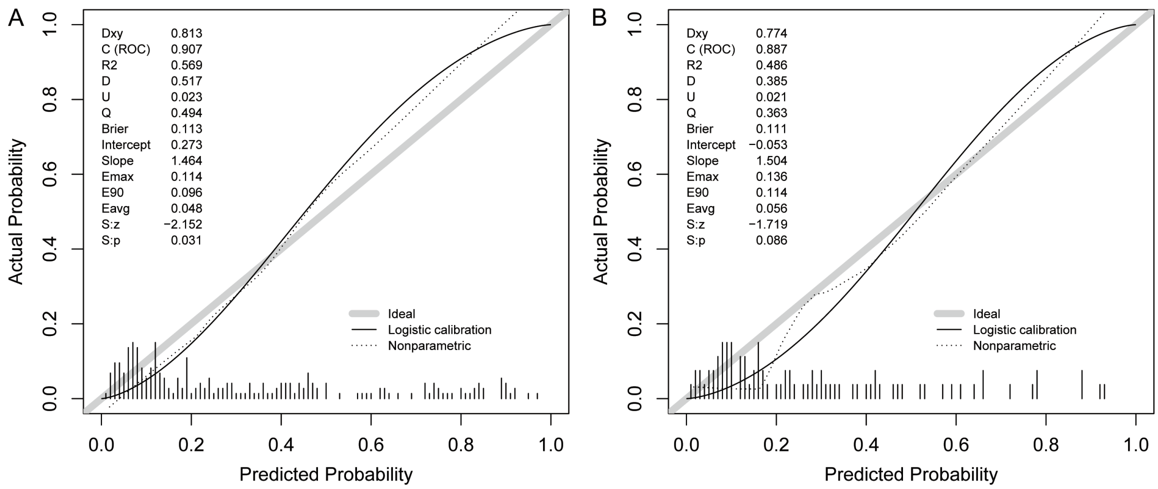


**Figure 3.** Calibration curves of the GBM model. A. Training set; B. Testing set. Note: GBM: gradient boosting machine.

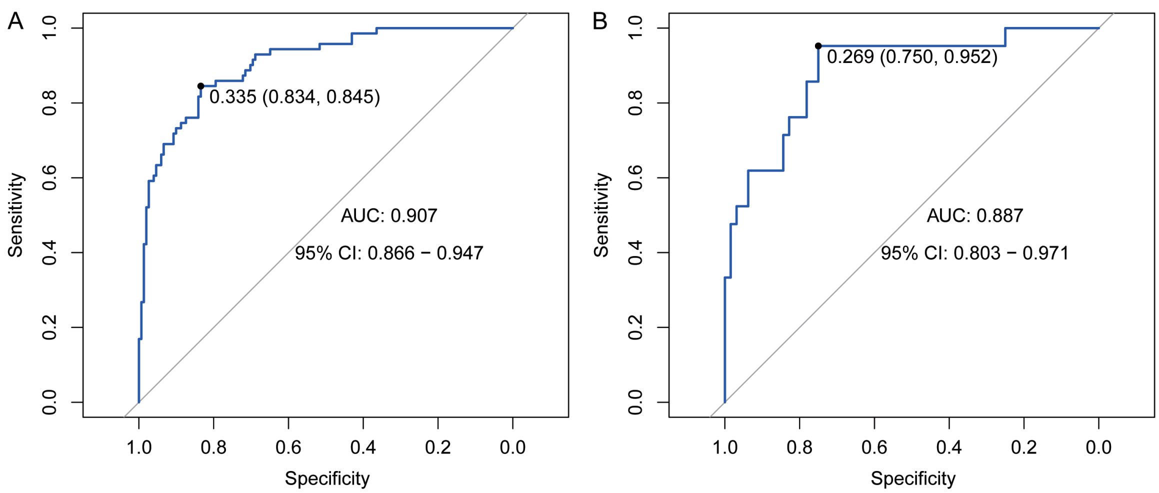


**Figure 4.** ROC curves for the GBM model. A. Training set; B. Testing set. Note: ROC: receiver operating characteristic; GBM: gradient boosting machine.

*ROC curve for GBM*

ROC curves for the developed GBM model in predicting the prognosis of ARDS patients in both training and testing sets were plotted (**Figure 4A**, **4B**). The AUC values in the training and testing sets were 0.907 (95% CI: 0.866-0.947) and 0.887 (95% CI: 0.803-0.971), respectively.

*Construction of a prognostic nomogram model for ARDS patients*

Using ARDS patient outcomes as the dependent variable (1=death, 0=survival), a nomogram model was constructed with age, oxygenation index, NLR, IL-8, and NT-proBNP as independent variables (**Figure 5**). The total nomogram score was calculated by summing the individual scores assigned to each predictor, and the corresponding probability of adverse outcome was estimated. Results indicated that advanced age, decreased OI, and elevated NLR, IL-8, and NT-proBNP levels were risk factors for poor prognosis in ARDS patients. Calibration curves for the training and testing sets demonstrated good agreement between predicted and observed outcomes, with mean errors of 0.036 and 0.033, respectively (**Figure 6A**, **6B**).

*ROC curve for nomogram*

The ROC curves for the nomogram model in predicting prognosis of ARDS patients in the testing and validation sets were plotted (**Figure**

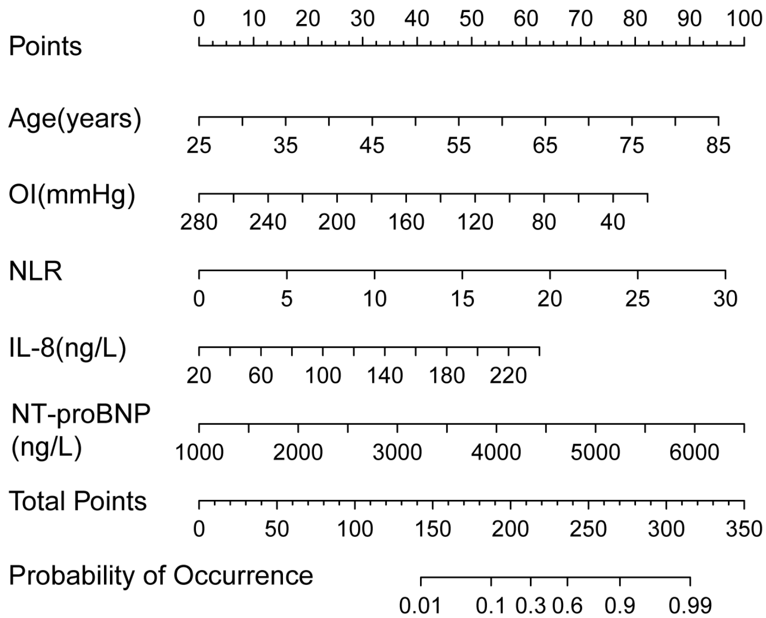


**Figure 5.** Nomogram model. Note: OI: Oxygenation Index; NLR: Neutrophil-to-lymphocyte ratio; IL: Interleukin; NT-proBNP: N-terminal pro-B-type natriuretic peptide.

**7A**, **7B**). The AUC values for the training and validation sets were 0.866 (95% CI: 0.810-0.923) and 0.835 (95% CI: 0.733-0.937), respectively.

*External validation of two models*

An independent cohort of 143 ARDS patients (**Table 3**) treated at Qinzhou First People's Hospital from July 2024 to March 2025 were enrolled for external validation of both the GBM and nomogram models (**Table 4** and **Figure 8**). The confusion matrix-based performance metrics for both models are shown in **Table 5**, where the GBM model outperformed the nomogram model across all indicators.

**Discussion**

ARDS is a type of rapidly progressive inflammatory lung injury characterized by dyspnea, tachypnea, and hypoxemia, and it represents a common life-threatening emergency in respiratory medicine [12, 13]. ARDS is associated with high morbidity and mortality ratesand remains a major challenge in critical care. The LUNG SAFE study, the largest international multicenter cohort on ARDS, reported that about 10.4% of ICU admissions is associated with ARDS, with mortality rates reaching up to 45%. This high mortality rate is closely related to the complex pathophysiological mechanisms of ARDS, which extend beyond lung injury to systemic inflammation, coagulation dysregulation, and other interconnected biological processes.

Current pharmacological treatments for ARDS primarily aim to reduce pulmonary edema and inflammation, support vasodilation, and facilitate the repair of epithelial, endothelial, and extracellular matrix structures. Despite these interventions, clinical outcomes remain unsatisfactory. Therefore, identifying risk factors for adverse clinical outcomes has become a research priority for reducing mortality in ARDS patients. Given that ARDS prognosis is affected by many interacting factors, the development of robust risk prediction models for early identification of high-risk individuals and prompt implementation of preventive measures may substantially improve patient outcomes [14, 15].

Previous studies have explored early prognostic markers for ARDS. For instance, dead-space ventilation indices have been shown to be independently associated with mortality in adults with ARDS [16]. However, such studies focused solely on a single factor, potentially introducing selection and outcome biases. In contrast, this study incorporated multidimensional indicators to predict ARDS outcomes, providing a more comprehensive prognostic assessment.

In this study, patients in the mortality group were significantly older than those in the survival group, consistent with previous research [17]. Advanced age is often associated with a higher burden of comorbidities, more severe lung injury, and greater extrapulmonary organ dysfunction. However, age alone is insufficient to differentiate ARDS from other pulmonary conditions [18]. Furthermore, physiological aging and age-related multisystem pathological

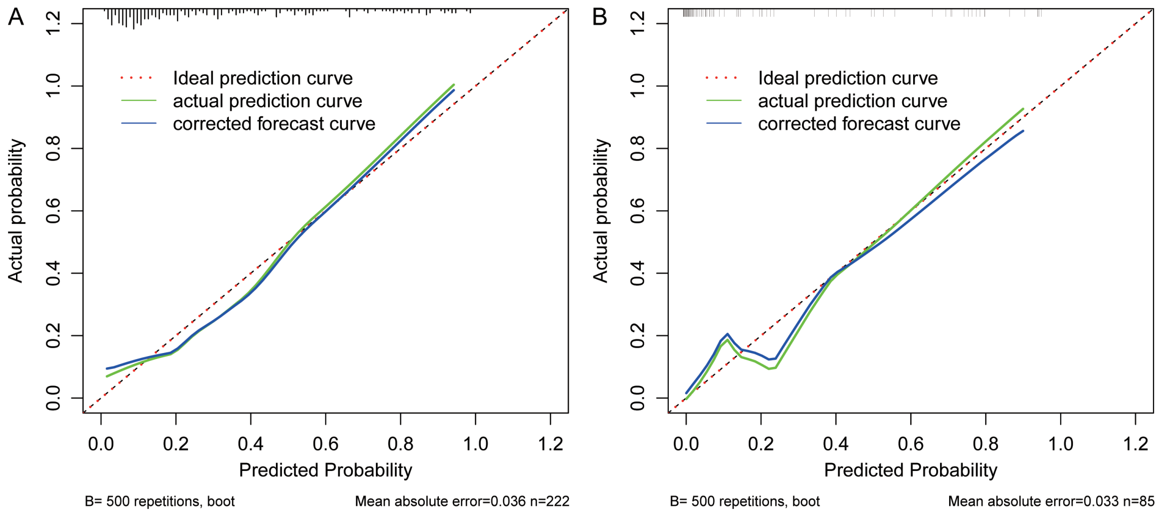


**Figure 6.** Calibration curves for the nomogram model. A. Training set; B. Testing set.

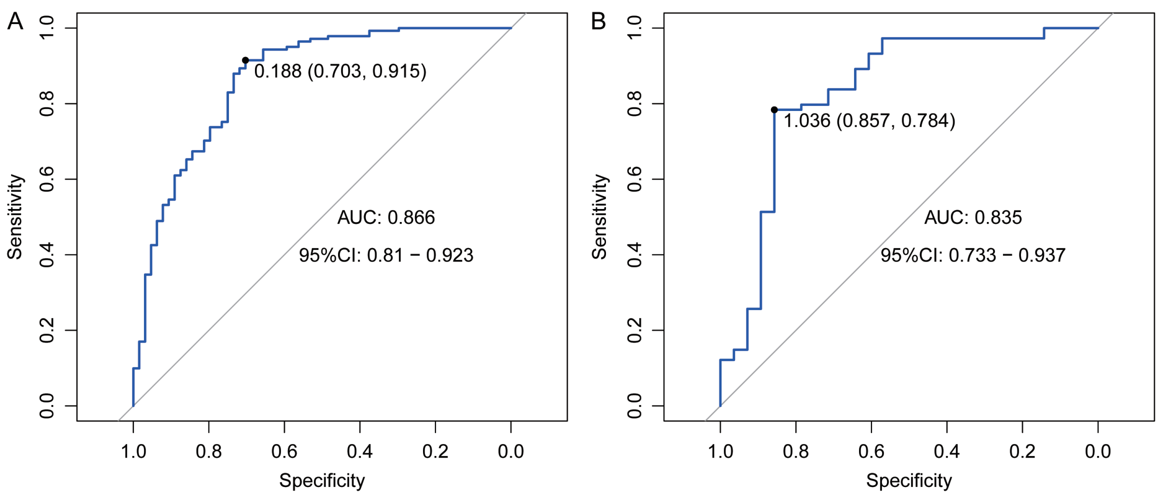


**Figure 7.** ROC curves for the nomogram model. A. Training set; B. Testing set. Note: ROC: receiver operating characteristic.

**Table 3.** External validation of patients' general characteristics

| | Death group (n=41) | Survival group (n=102) | $t/x^2$ | *P* |
|---|---|---|---|---|
| Age (years) | 65.85±9.13 | 58.97±7.77 | 4.553 | <0.001 |
| Sex (Male/Female) | 23.59±2.25 | 23.65±2.23 | -0.145 | 0.885 |
| BMI (kg/m$^2$) | 26 (63.41)/15 (36.59) | 54 (52.94)/48 (47.06) | 1.302 | 0.254 |
| Hypertension (Yes/No) | 18 (43.90)/23 (56.10) | 50 (49.02)/52 (50.98) | 0.307 | 0.579 |
| Diabetes (Yes/No) | 20 (48.78)/21 (51.22) | 47 (46.08)/55 (53.92) | 0.086 | 0.770 |
| Coronary Heart Disease (Yes/No) | 17 (41.46)/24 (58.54) | 55 (53.92)/47 (46.08) | 1.816 | 0.178 |
| Smoking History (Yes/No) | 26 (63.41)/15 (36.59) | 44 (43.14)/58 (56.86) | 4.812 | 0.028 |
| Alcohol Consumption (Yes/No) | 22 (53.66)/19 (46.34) | 41 (40.20)/61 (59.80) | 2.150 | 0.143 |

conditions may increase the risk of ARDS in the context of systemic illness [19]. Therefore, intensive monitoring in older patients should focus not only on ARDS development but also on concomitant organ dysfunction and systemic failure.

**Table 4.** External validation confusion matrix

| Model | Predicted | Reference | | Total |
|---|---|---|---|---|
| | | Death | Survival | |
| GBM | Death | 25 | 5 | 30 |
| | Survival | 16 | 97 | 113 |
| Total | | 41 | 102 | 143 |
| Nomogram | Death | 19 | 7 | 26 |
| | Survival | 22 | 95 | 117 |
| Total | | 41 | 102 | 143 |

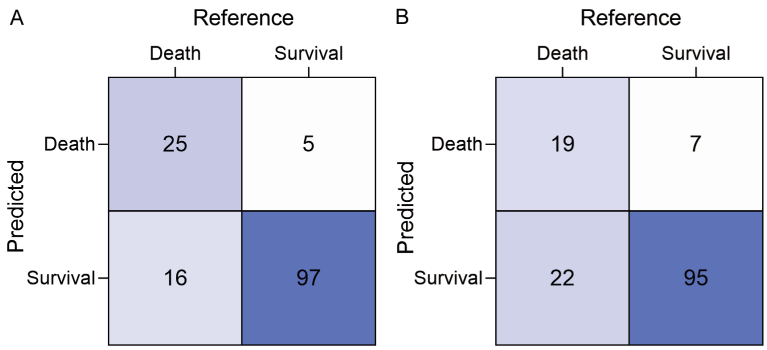


**Figure 8.** Confusion matrix. A. GBM model; B. Nomogram model. Note: GBM: gradient boosting machine.

In this study, patients in the mortality group had a significantly lower OI value than those in the survival group, which aligns with the finding in previous studies [20]. OI reflects the efficiency of pulmonary oxygen exchange and lung responsiveness to ventilatory support. An increase in OI indicates enhanced lung compliance and gas exchange, which is generally associated with better patient prognosis [21].

Compared with previous studies, this research additionally included NLR, IL-8, and NT-proBNP in the model. NLR, derived from routine blood tests, reflects systemic stress level and immune-inflammatory response [22]. Our findings demonstrated significantly higher NLR in the mortality group, indicating a more pronounced stress and inflammatory response. IL-8 is a key proinflammatory cytokine released during inflammatory responses [23]. In this study, patients in the mortality group exhibited a significantly higher IL-8 level compared to those in the survival group, indicating a more intense inflammatory cascade. Excessive inflammatory response can exacerbate pulmonary and extrapulmonary organ injury, resulting in poor prognosis. Brain natriuretic peptide (BNP) is primarily synthesized and secreted by ventricular cardiomyocytes in response to increased wall stress [24]. NT-proBNP, a more stable cleavage product with a longer half-life, is widely used as a surrogate marker of BNP activity [25]. Elevated NT-proBNP levels indicates increased cardiac load and impaired cardiac function [26]. In ARDS patients, reduced lung compliance and hypoxemia may increase pulmonary vascular resistance, contributing to pulmonary hypertension and right ventricular strain [27, 28]. These effects tend to worsen with disease severity [29]. Accordingly, elevated NT-proBNP level is associated with pulmonary hypertension and cardiac dysfunction in ARDS patients [30]. Although BNP gene transcripts have been detected in non-cardiac tissues, such as the central nervous system, thyroid, and lung [31], their contribution to circulating BNP levels is minimal compared with that of cardiomyocytes [32]. In this study, NT-proBNP levels in the mortality group were significantly higher than the survival group, suggesting more severe cardiopulmonary involvement and a poorer prognosis.

Both a nomogram model and a GBM model were developed for predicting outcomes in ARDS patients. These models incorporated factors known to influence patient prognosis. The GBM model identified NT-proBNP, age, NLR, IL-8, and OI as the most important predictors, ranked from most to least influential. The nomogram demonstrated that advanced age, reduced OI, elevated NLR, IL-8, and NT-proBNP levels were associated with poor prognosis in ARDS patients. Calibration curves indicated good agreement between predicted and observed outcomes for both models. Moreover, the GBM model demonstrated superior discriminative performance compared with the nomogram, supported by higher AUC values in both the training and validation sets (0.907 vs. 0.866 in training set; 0.887 vs. 0.835 in validation set).

**Table 5.** External validation metrics of confusion matrix for the two models (%)

| Model | Accuracy | Sensitivity | Specificity | Precision | Recall | F1 |
|---|---|---|---|---|---|---|
| GBM | 85.31 | 83.33 | 85.84 | 83.33 | 60.98 | 70.42 |
| Nomogram | 79.72 | 73.08 | 46.34 | 73.08 | 46.34 | 56.72 |

This study does have certain limitations. First, as a retrospective, single-center analysis, selection bias cannot be completely excluded, although strict inclusion and exclusion criteria were applied. Second, the available indicators were limited to routinely collect variables at our institution, which may have introduced residual confounding. Therefore, future prospective, multicenter studies using multidimensional indicators are necessary to validate and refine these predictive models for identifying ARDS patients at high risk of poor outcomes.

## Conclusion

The 28-day mortality rate in ARDS patients are primarily associated with age, oxygenation index, NLR, IL-8, and NT-proBNP. A GBM model constructed using these variables demonstrates strong predictive performance and may serve as a useful tool for early identification of ARDS patients at high risk of poor prognosis.

## Disclosure of conflict of interest

None.

**Address correspondence to:** Yuji Liang, Department of Critical Care Medicine, Qinzhou First People's Hospital, Qinzhou 535000, Guangxi, China. Tel: +86-13877757027; E-mail: liangyuji001@163.com

## References

[1] Poli S, Sciorio F, Piacentini G, Pietrobelli A, Pecoraro L and Pieropan S. Acute respiratory failure in autoimmune rheumatic diseases: a review. J Clin Med 2024; 13: 3008.
[2] Liu Z, Liu D, Wang Z, Zou Y, Wang H, Li X, Zheng D and Zhou G. Association between inflammatory biomarkers and acute respiratory distress syndrome or acute lung injury risk: a systematic review and meta-analysis. Wien Klin Wochenschr 2022; 134: 24-38.
[3] Wu L, Lei Q, Gao Z and Zhang W. Research progress on phenotypic classification of acute respiratory distress syndrome: a narrative review. Int J Gen Med 2022; 15: 8767-8774.
[4] Liaqat A, Mason M, Foster BJ, Kulkarni S, Barlas A, Farooq AM, Patak P, Liaqat H, Basso RG, Zaman MS and Pau D. Evidence-based mechanical ventilatory strategies in ARDS. J Clin Med 2022; 11: 319.
[5] Kelly-Hedrick M, Liu S, Hatfield J, Soto AL, Bartlett AM, Heo HJ, O'Callaghan E, Arulraja E, Kaplan S, Ohnuma T, Krishnamoorthy V, Colton K and Komisarow J. Management of traumatic brain injury and acute respiratory distress syndrome-What evidence exists? A scoping review. J Intensive Care Soc 2025; 26: 205-222.
[6] Yang J, Zhou X, Qiao X and Shi M. Friend or foe: the role of platelets in acute lung injury. Front Immunol 2025; 16: 1556923.
[7] Wang HX, Huang XH, Ma LQ, Yang ZJ, Wang HL, Xu B and Luo MQ. Association between lactate-to-albumin ratio and short-time mortality in patients with acute respiratory distress syndrome. J Clin Anesth 2024; 99: 111632.
[8] Huang B, Liang D, Zou R, Yu X, Dan G, Huang H, Liu H and Liu Y. Mortality prediction for patients with acute respiratory distress syndrome based on machine learning: a population-based study. Ann Transl Med 2021; 9: 794.
[9] Konstantinov AV and Utkin LV. Interpretable machine learning with an ensemble of gradient boosting machines. Knowledge-Based Systems 2021; 222: 106993.
[10] Ranieri VM, Rubenfeld GD, Thompson BT, Ferguson ND, Caldwell E, Fan E, Camporota L and Slutsky AS. Acute respiratory distress syndrome: the Berlin Definition. JAMA 2012; 307: 2526-2533.
[11] Bardají-Carrillo M, López-Herrero R, Aguilar G, Arroyo-Hernantes I, Gómez-Sánchez E, Camporota L, Villar J and Tamayo E. Epidemiological trends of mechanically ventilated acute respiratory distress syndrome in the twenty-first century: a nationwide, population-based retrospective study. J Intensive Care 2025; 13: 9.
[12] Liao L and Liao P. Bioinformatics analysis of the potential biomarkers for acute respiratory distress syndrome. Biosci Rep 2020; 40: BSR20192436.
[13] Yan Y, Bao J, Cai S, Zhong X, Geng B, Liang J, Deng Z, Chen Z, Qin Z, Hu H and Zeng Z. The effects of prolonged prone positioning on response and prognosis in patients with acute respiratory distress syndrome: a retrospective cohort study. J Intensive Care 2025; 13: 24.
[14] Wang Z, Xing L, Cui H, Fu G, Zhao H, Huang M, Zhao Y and Xu J. A Nomogram for predicting the mortality of patients with acute respiratory distress syndrome. J Healthc Eng 2022; 2022: 5940900.

[15] Huang D, He D, Gong L, Jiang W, Yao R and Liang Z. A nomogram for predicting mortality in patients with pneumonia-associated acute respiratory distress syndrome (ARDS). J Inflamm Res 2024; 17: 1549-1560.
[16] Jayasimhan D, Chieng J, Kolbe J and Sidebotham DA. Dead-space ventilation indices and mortality in acute respiratory distress syndrome: a systematic review and meta-analysis. Crit Care Med 2023; 51: 1363-1372.
[17] Chen W, Chen Q, Tang Z, Li Z, Chen W, Xiong X, Wen D and Zhang Z. Association of albumin-bilirubin (ALBI) grade with 28-day all-cause mortality in patients with acute respiratory distress syndrome: a retrospective analysis of the MIMIC-IV database. Mediators Inflamm 2025; 2025: 9930648.
[18] Lee JK, Kho BG, Yoon JY, Yoon CS, Na YO, Park HY, Kim TO, Kwon YS, Kim YI, Lim SC and Shin HJ. The clinical characteristics associated with the ratio between the main pulmonary artery and ascending aorta diameter in patients with acute exacerbation of chronic obstructive pulmonary disease. J Thorac Dis 2024; 16: 4924-4934.
[19] Papathanasiou IV, Fradelos EC, Malli F, Stefanidis I, Zintzaras E and Doxani C. A systematic review of observational studies assessing the impact of oxidative stress in cognitive decline. Wiad Lek 2021; 74: 1995-2003.
[20] Cai X, Li J, Qin P, An P, Yang H, Zuo M and Wang J. Prognostic prediction using the clinical data and ultrasomics-based model in acute respiratory distress syndrome (ARDS) combined with acute kidney injury (AKI). Int J Clin Pract 2022; 2022: 4822337.
[21] Gardner MM, Kirschen MP, Wong HR, McKeone DJ, Scott Halstead E, Thompson JM, Himebauch AS, Topjian AA and Yehya N. Biomarkers associated with mortality in pediatric patients with cardiac arrest and acute respiratory distress syndrome. Resuscitation 2022; 170: 184-193.
[22] Wang Q, Xie T, Gao R, Long X, Wei J, Ye L, Jiang J, Kang L, Wang J, Jun C and Lyu J. Neutrophil-to-lymphocyte ratio is a powerful predictor of adult patients with acute respiratory distress syndrome who might benefit from corticosteroid therapy. Am J Transl Res 2021; 13: 11556-11570.
[23] Cornejo RA, Arellano DH, Ruiz-Rudolph P, Guiñez DV, Morais CCA, Gajardo AIJ, Lazo MT, Brito RE, Cerda MA, González SJ, Rojas VA, Diaz GA, López LDM, Medel JN, Soto DI, Bruhn AR, Amato MBP and Estuardo NR. Inflammatory biomarkers and pendelluft magnitude in ards patients transitioning from controlled to partial support ventilation. Sci Rep 2022; 12: 20233.
[24] Li Y, Wu J and Lei R. Effects of mind-body exercise in chronic cardiopulmonary dyspnoea patients-a network meta-analysis of randomized controlled trials. Front Cardiovasc Med 2025; 12: 1546996.
[25] Ni Q, Li C and Lin H. Prognostic value of N-terminal probrain natriuretic peptide for patients with acute respiratory distress syndrome: a systematic review and meta-analysis. Biomed Res Int 2020; 2020: 3472615.
[26] Selçuk M, Keskin M, Çınar T, Günay N, Doğan S, Çiçek V, Kılıç Ş, Asal S, Yavuz S, Keser N and Orhan AL. Prognostic significance of N-terminal Pro-BNP in patients with COVID-19 pneumonia without previous history of heart failure. J Cardiovasc Thorac Res 2021; 13: 141-145.
[27] Baka M, Bagka D, Tsolaki V, Zakynthinos GE, Diakaki C, Mantzarlis K and Makris D. Hemodynamic and respiratory changes following prone position in acute respiratory distress syndrome patients: a clinical study. J Clin Med 2023; 12: 760.
[28] Grotberg JC, Reynolds D and Kraft BD. Management of severe acute respiratory distress syndrome: a primer. Crit Care 2023; 27: 289.
[29] Jayasimhan D, Foster S, Chang CL and Hancox RJ. Cardiac biomarkers in acute respiratory distress syndrome: a systematic review and meta-analysis. J Intensive Care 2021; 9: 36.
[30] Wang M, Zhao F, Sun L, Liang Y, Yan W, Sun X, Zhou Q and He B. High-flow nasal cannula versus noninvasive ventilation in AECOPD patients with respiratory acidosis: a retrospective propensity score-matched study. Can Respir J 2023; 2023: 6377441.
[31] Abassi Z, Hamo-Giladi DB, Kinaneh S and Heyman SN. The endocrine basis of the cardio-renal axis: new perspectives regarding corin. Physiol Rep 2024; 12: e16105.
[32] Fujishima S. Guideline-based management of acute respiratory failure and acute respiratory distress syndrome. J Intensive Care 2023; 11: 10.