

Construction and Validation of a Predictive Model for Chronic Obstructive Pulmonary Disease with Pulmonary Hypertension Based on Lasso Regression

Qiaoyun Zhou, Guangxi Zhu, Peng Zhang, Baolong Qi

Department of Respiratory Medicine, Anhui NO.2 Provincial People's Hospital, Hefei, Anhui, People's Republic of China

Correspondence: Qiaoyun Zhou, Department of Respiratory Medicine, Anhui NO.2 Provincial People's Hospital, Hefei, Anhui, 230041, People's Republic of China, Tel +86 152 5659 1836, Email Qiao_10205266@126.com

Purpose: Chronic obstructive pulmonary disease (COPD) is often complicated by pulmonary hypertension (PH), leading to symptom exacerbation, increased mortality, and poor prognosis. Early identification of high-risk patients is crucial for timely intervention. This study aims to construct and validate a machine learning-based predictive model for COPD patients with PH.

Patients and Methods: A retrospective analysis was conducted on 537 COPD patients admitted to Anhui NO.2 Provincial People's Hospital from December 2023 to November 2025. The patients were randomly divided into a model construction group (n=376) and an internal validation group (n=161) at a ratio of 7:3. Clinical and laboratory data were collected. LASSO regression, a method that effectively selects the most relevant predictors by shrinking less important variables to zero, was used to screen predictive variables. This was followed by the construction of a nomogram model through multivariate logistic regression. Model performance was evaluated using the area under the receiver operating characteristic curve (AUC), calibration curve, and clinical impact curve.

Results: LASSO regression identified six independent predictors: Duration of COPD, Long-Term Domiciliary Oxygen Therapy, systemic inflammatory response index, red blood cell distribution width, arterial oxygen partial pressure, and B-type natriuretic peptide. The nomogram model demonstrated excellent discriminative power, with an AUC of 0.812 (95% CI: 0.762–0.863) in the model group and 0.826 (95% CI: 0.754–0.897) in the validation group. Calibration curves and clinical impact curves showed high consistency between predicted and observed values.

Conclusion: Our study constructed and internally validated a nomogram model based on LASSO regression, incorporating six easily accessible clinical variables, which can be used to predict the risk of PH in COPD patients. The model facilitates early clinical identification of high-risk patients, risk stratification, and implementation of individualized management.

Keywords: chronic obstructive pulmonary disease, pulmonary hypertension, LASSO regression, predictive model, nomogram

Introduction

Chronic obstructive pulmonary disease (COPD) is a chronic respiratory disorder characterized by a persistent decline in lung function and extensive lung damage. It accounts for approximately 55% of all respiratory diseases and causes over 3 million deaths annually, making it the third leading cause of death globally, representing 6% of total global mortality.^{1,2} Although there is no cure, various comprehensive treatment approaches can effectively control symptoms, improve patients' quality of life, reduce acute exacerbations, and delay disease progression. COPD leads to impaired pulmonary ventilation and decreased oxygen uptake capacity, resulting in increased pulmonary vascular resistance and a gradual elevation of pulmonary artery pressure.³ In patients with COPD, the expression of inflammatory cytokines, particularly the pro-inflammatory cytokines IL-1 β , IL-6, IL-8, and TNF- α , is increased. These factors play a key role in the development of pulmonary hypertension by promoting pulmonary vascular remodeling, endothelial dysfunction, and smooth muscle cell proliferation.⁴ In the early stage of the disease, the right ventricle compensates for increased

pulmonary artery pressure through compensatory hypertrophy. However, as the disease progresses, this compensatory mechanism gradually fails, and the symptoms of pulmonary hypertension and cor pulmonale (pulmonary heart disease) gradually emerge, ultimately leading to heart failure and death. According to data from the 2022 China Cardiovascular Health and Disease Annual Report, there are currently 5 million patients with cor pulmonale in China.⁵ Clinically, 80% of chronic cor pulmonale is caused by COPD.⁶ As COPD progresses to PH, symptoms worsen, mortality rates increase, and the risk of acute exacerbations rises. COPD patients with PH have a poor long-term prognosis, with a median survival of only 2–5 years after diagnosis.⁷ Studies report a prevalence of PH in COPD patients ranging from 30% to 70%.⁸ COPD patients with PH often present with more complex and severe symptoms. In addition to the symptoms of COPD itself, such as cough, sputum production, shortness of breath, and dyspnea, manifestations of right heart failure may also occur, including lower limb edema, jugular vein distension, and hepatomegaly. These conditions significantly limit the patient's daily activities and substantially reduce their quality of life. Therefore, in clinical practice, early identification of factors contributing to pulmonary hypertension (PH) and timely intervention are of positive significance for the prevention and management of COPD.

Subjects and Methods

Study Subjects and Inclusion/Exclusion Criteria

This retrospective study continuously enrolled COPD patients hospitalized in the Department of Respiratory Medicine at Anhui NO.2 Provincial People's Hospital from December 2023 to November 2025. Inclusion criteria: Post-bronchodilator FEV₁/FVC < 0.70, with chest CT findings indicating emphysema or airway abnormalities, meeting the diagnostic criteria for COPD according to the Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2023 report.⁹ Exclusion criteria: incomplete clinical data; comorbidities including tumors, bronchial asthma, interstitial lung disease, deep vein thrombosis, severe heart failure, severe hepatic or renal dysfunction, significant sequelae of cerebrovascular disease, or cognitive impairment; other diseases causing pulmonary hypertension (PH), such as systemic lupus erythematosus, idiopathic pulmonary hypertension (IPH), familial pulmonary hypertension, thrombotic pulmonary hypertension, left heart disease-induced pulmonary hypertension, or high-altitude pulmonary hypertension; and use of medications that may elevate pulmonary artery pressure. The study was conducted in accordance with the Declaration of Helsinki and Good Clinical Practice guidelines. The protocol was approved by the Medical Ethics Committee of Anhui NO.2 Provincial People's Hospital (Approval No: (R) 2025-152). Due to its retrospective nature, informed consent was waived. The committee also approved the confidentiality measures, and we confirm that all patient data were anonymized and maintained with strict confidentiality prior to analysis.

Data Collection

Patient data collection included: (1) General information: age, gender, Duration of COPD, body mass index (BMI), and medical history (including smoking, alcohol use, diabetes, stroke, hypertension, coronary artery disease, and long-term domiciliary oxygen therapy (LTDO2)). (2) Laboratory parameters: complete blood count (neutrophil count, lymphocyte count, monocyte count, hematocrit (HCT), hemoglobin (Hb), red blood cell distribution width (RDW), platelet count (PLT)), arterial blood gas analysis (pH, arterial oxygen partial pressure (PaO₂), arterial carbon dioxide partial pressure (PaCO₂), bicarbonate (HCO₃⁻), lactate (Lac)), D-dimer, fibrinogen (FIB), fibrin degradation products (FDP), B-type natriuretic peptide (BNP), procalcitonin (PCT), albumin (Alb), blood urea nitrogen (BUN), creatinine (Crea), uric acid (UA), C-reactive protein (CRP), low-density lipoprotein cholesterol (LDL-c), high-density lipoprotein cholesterol (HDL-c). Systemic Inflammatory Response Index (SIRI) = neutrophil count × monocyte count / lymphocyte count. (3) Imaging data: cardiac ultrasound.

Diagnostic Criteria for PH

Due to the invasiveness of right heart catheterization for pulmonary artery pressure measurement, its implementation in non-surgical patients is challenging. Therefore, the ESC/ERS guidelines recommend Doppler echocardiography as the preferred non-invasive diagnostic tool for suspected PH in COPD patients. Accordingly, this study referenced the

diagnostic criteria outlined in the 2022 European Society of Cardiology and Respiratory Society (ESC/ERS) Guidelines for the Diagnosis and Treatment of Pulmonary Hypertension¹⁰. Transthoracic echocardiography was performed to measure the peak flow velocity of tricuspid regurgitation at rest, and pulmonary systolic pressure was calculated using the revised Bernoulli equation. A peak flow velocity of tricuspid regurgitation >2.8 m/s was considered indicative of pulmonary systolic pressure >40 mmHg. (Equipment: Philips EPIQ 7C, probe frequency 2–5 MHz). The examination was performed independently by one attending physician with ≥ 5 years of experience in cardiac ultrasound, who was blinded to the patients' clinical, laboratory, and blood gas analysis results. The average value of 3–5 clear cardiac cycles was taken as the final measurement. Inter-observer agreement was assessed by having another physician independently perform the measurements.

Statistical Analysis

Data analysis was performed using SPSS (version 26.0) and R (version 4.5.2). Continuous variables were expressed as mean $\pm$ standard deviation (SD) when following a normal distribution. Between-group comparisons were conducted using the independent samples *t*-test. For non-normally distributed continuous variables, median with interquartile range (IQR) was used, and Mann–Whitney *U*-test was employed for inter-group comparisons. Categorical variables were presented as frequency and percentage, with differences between groups analyzed using the chi-square test. A *P*-value <0.05 was considered statistically significant.

Using the R package “caret”, the enrolled patients were randomly divided into a model group ($n=376$) and an internal validation group ($n=161$) in a 7:3 ratio for model development and internal validation. Lasso regression analysis was performed using the R package “glmnet”, with 10-fold cross-validation employed to determine the optimal λ and screen the optimal feature subset. Subsequently, multivariate logistic regression analysis was conducted. The “rms” package in R was utilized to construct nomogram, and the “pROC” software package was employed to generate receiver operating characteristic (ROC) curves, with the area under the curve (AUC) calculated to assess predictive accuracy. The “Resource Selection” package was used for Hosmer-Lemeshow test calibration to evaluate model stability. To ensure robustness, 10-fold cross-validation with the “caret” package was performed for internal validation. Further, the “ggDCA” package in R was employed to analyze clinical impact curves (CIC) to evaluate the clinical utility of the nomogram. The predictive capacity of the model was internally assessed using ROC curves, AUC, Hosmer-Lemeshow goodness-of-fit tests, calibration curves, and CICs to evaluate discrimination, calibration, and clinical benefit utility, with validation performed using the validation set data.

Results

Case Screening Results

A total of 1232 COPD patients hospitalized at Anhui NO.2 Provincial People's Hospital from December 2023 to November 2025 were enrolled. Exclusion criteria included 102 cases of severe heart failure, 69 cases of severe hepatic or renal dysfunction, 58 cases with significant sequelae of cerebrovascular disease, 78 cases of cognitive impairment, 69 cases of malignancy, 231 cases of bronchial asthma, 28 cases of interstitial lung disease, 35 cases of deep vein thrombosis, and 25 cases of other diseases leading to PH. Ultimately, 537 patients were included, with 376 in the model group (283 Non-PH Group and 93 PH Group) and 161 in the validation group (Figure 1).

Comparison of Demographic Characteristics and Clinical Data Between the Model Group and the Internal Validation Group

No significant differences were observed between the two groups in terms of age, gender, Duration of COPD, BMI, medical history (smoking, alcohol use, diabetes, stroke, hypertension, Coronary Artery Disease, LTDO2), SIRI, HCT, Hb, RDW, PLT, pH, PaO₂, PaCO₂, HCO₃⁻, Lac, D-dimer, FIB, FDP, BNP, PCT, Alb, BUN, Crea, UA, CRP, LDL-C and HDL-C ($p > 0.05$) (Table 1).

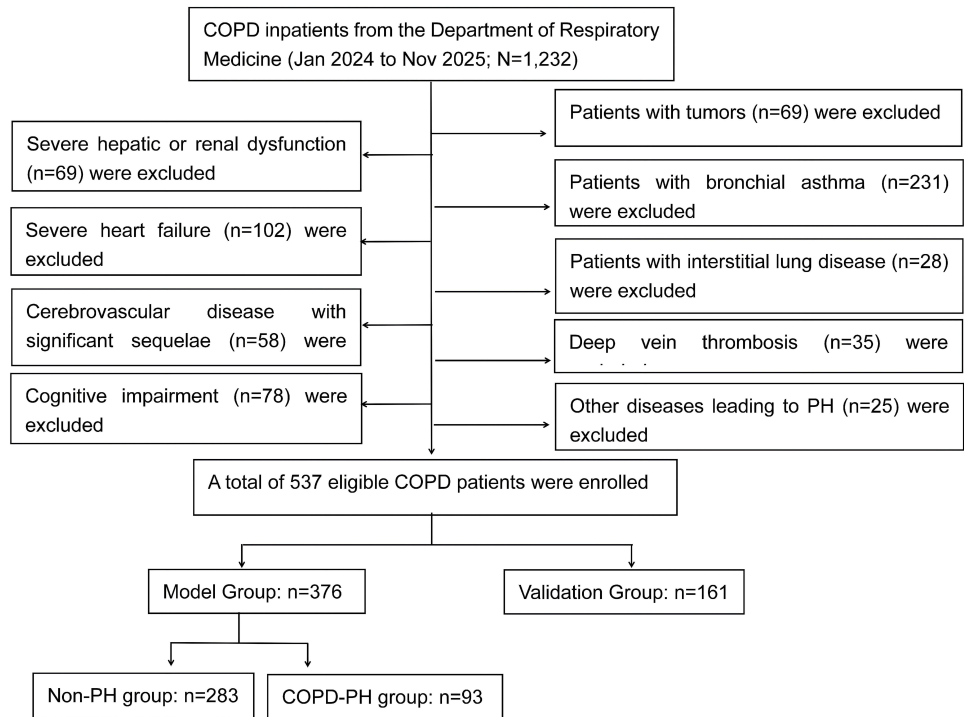


Figure 1 Enrollment Flowchart.

Comparison of Demographic Characteristics and Clinical Data Between the Non-PH Subgroup Within the Model Group and the PH Subgroup

No significant differences were observed between the two groups in terms of age, gender, BMI, medical history (smoking, alcohol use, diabetes, stroke, hypertension, Coronary Artery Disease), HCT, Hb, PLT, pH, PaCO₂, HCO₃⁻, Lac, D-dimer, FIB, FDP, PCT, Alb, BUN, Crea, UA, CRP, LDL-C, and HDL-C ($p > 0.05$). However, significant differences were observed between the two groups in terms of COPD duration, LTDO₂, SIRI, RDW, PaO₂, and BNP ($p < 0.05$) (Table 2).

Table 1 Comparison of Population Characteristics and Clinical Data in the Model Group and Internal Validation Group

Variable	Total (n=537)	Model Group (n=376)	Internal Validation Group (n=161)	t- χ^2 /Z value	Pvalue
Age (years, median [IQR])	71.00 (68.00–75.00)	71.000 (68.00–75.00)	72.00 (68.00–75.00)	0.504	0.614
Male (n,%)	431 (80.26%)	299 (79.52%)	132 (81.99%)	0.291	0.589
Duration of COPD (years; M, IQR)	14 (8.00–19.00)	14.00 (8.75–19.00)	13.00 (8.00–18.00)	1.269	0.176
BMI, median [IQR]	25.33 (22.03–28.55)	25.30 (21.54–28.73)	25.40 (22.76–28.19)	0.790	0.321
History					
Smoking (n,%)	418 (77.84%)	292 (77.66%)	126 (78.26%)	0.002	0.968
Alcohol Use (n,%)	146 (27.19%)	98 (26.06%)	48 (29.81%)	0.623	0.430
Diabetes (n,%)	109 (20.30%)	71 (18.88%)	38 (23.60%)	1.274	0.259
Stroke (n,%)	85 (15.86%)	55 (14.67%)	30 (18.63%)	1.074	0.300
Hypertension (n,%)	192 (35.75%)	127 (33.78%)	65 (40.37%)	1.858	0.173
Coronary Artery Disease (n,%)	148 (27.56%)	96 (25.53%)	52 (32.30%)	2.257	0.133
LTDO ₂ (n,%)	123 (22.91%)	83 (22.07%)	40 (24.85%)	0.346	0.557

(Continued)

Table 1 (Continued).

Variable	Total (n=537)	Model Group (n=376)	Internal Validation Group (n=161)	t, χ^2/Z value	Pvalue
Laboratory testing					
SIRI (median [IQR])	2.32 (1.63–3.41)	2.32 (1.59–3.65)	2.320 (1.82–3.08)	0.206	0.836
HCT (% median [IQR])	0.42 (0.37–0.46)	0.41 (0.37–0.46)	0.42 (0.38–0.47)	1.202	0.229
Hb (g/L, median [IQR])	126.00 (115.00–137.00)	126.00 (115.00–138.00)	126.00 (117.00–137.00)	0.312	0.755
RDW (% median [IQR])	14.07 (13.12–14.77)	14.10 (13.10–14.78)	14.01 (13.15–14.77)	0.126	0.900
PLT ($10^9/L$, $\bar{x} \pm s$)	186.80 $\pm$ 36.99	186.61 $\pm$ 37.30	187.24 $\pm$ 36.37	0.180	0.857
pH (% median [IQR])	7.40 (7.37–7.43)	7.40 (7.37–7.43)	7.40 (7.38–7.43)	1.318	0.186
PaO ₂ (mmHg, median [IQR])	84.00 (78.00–89.00)	84.00 (77.000–89.00)	84.00 (78.000–90.00)	1.259	0.208
PaCO ₂ (mmHg, median [IQR])	42.00 (38.00–45.00)	41.50 (38.00–45.00)	43.00 (38.00–46.00)	1.035	0.300
HCO ₃ ⁻ (mmol/L, median [IQR])	24.00 (22.00–25.000)	24.00 (22.00–26.00)	23.00 (22.00–25.00)	0.974	0.327
Lac (mmol/L, median [IQR])	1.77 (1.54–2.01)	1.77 (1.53–1.99)	1.77 (1.58–2.06)	1.449	0.147
D-Dimer (mg/L, median [IQR])	0.44 (0.35–0.54)	0.44 (0.35–0.53)	0.42 (0.34–0.55)	0.294	0.769
FIB (g/L, $\bar{x} \pm s$)	2.60 $\pm$ 0.35	2.60 $\pm$ 0.34	2.60 $\pm$ 0.36	0.071	0.943
FDP (ug/mL, $\bar{x} \pm s$)	3.89 $\pm$ 0.92	3.92 $\pm$ 0.90	3.80 $\pm$ 0.95	1.469	0.142
BNP (ng/L, median [IQR])	162.53 (139.48–187.65)	160.92 (139.34–187.98)	163.51 (140.97–187.00)	0.165	0.869
PCT (ng/mL, median [IQR])	0.11 (0.09–0.13)	0.11 (0.09–0.13)	0.11 (0.09–0.13)	0.674	0.498
Alb (g/L, median [IQR])	41.30 (37.00–45.90)	41.35 (36.90–45.73)	41.30 (37.30–46.20)	0.187	0.852
BUN (mmol/L, median [IQR])	7.49 (5.80–9.35)	7.36 (5.77–9.24)	7.65 (5.96–9.68)	1.140	0.254
Crea (μ mol/L, median [IQR])	70.00 (54.00–88.00)	70.50 (54.00–88.00)	69.00 (54.00–89.00)	0.304	0.761
UA (μ mol/L, $\bar{x} \pm s$)	296.80 $\pm$ 54.021	297.63 $\pm$ 54.331	294.88 $\pm$ 53.41	0.539	0.590
CRP (mg/L, $\bar{x} \pm s$)	15.32 $\pm$ 6.11	15.46 $\pm$ 6.04	15.00 $\pm$ 6.28	0.789	0.430
LDL-C (mmol/L, $\bar{x} \pm s$)	2.81 $\pm$ 0.31	2.80 $\pm$ 0.316	2.85 $\pm$ 0.30	1.700	0.090
HDL-C (mmol/L, $\bar{x} \pm s$)	1.38 $\pm$ 0.46	1.37 $\pm$ 0.46	1.39 $\pm$ 0.46	0.437	0.662

Abbreviations: BMI, Body Mass Index; LTDO₂, Long-Term Domiciliary Oxygen Therapy; SIRI, Systemic Inflammatory Response Index; HCT, Hematocrit; Hb, Hemoglobin; RDW, Red Blood Cell Distribution Width; PLT, Platelet; FIB, Fibrinogen; FDP, Fibrinogen Degradation Products; BNP, B-type Natriuretic Peptide; PCT, Procalcitonin; Alb, Albumin; BUN, Blood Urea Nitrogen; Crea, Creatinine; UA, Uric Acid; CRP, C-reactive Protein; LDL-C, Low-Density Lipoprotein Cholesterol; HDL-C, High-Density Lipoprotein Cholesterol.

Table 2 Comparison of Population Characteristics and Clinical Data in the Model Groups

Variable	Total (n=376)	Non-PH Group (n=283)	PH Group (n=93)	t, χ^2/Z value	Pvalue
Age (years, median [IQR])	71.00 (68.00–75.00)	71.00 (68.00–75.00)	72.00 (68.00–76.00)	0.268	0.789
Male (n,%)	299 (79.52%)	227 (80.21%)	72 (77.42%)	0.186	0.667
Duration of COPD (years; M, IQR)	14.00 (8.75–19.00)	13.00 (7.00–18.00)	16.00 (13.00–20.00)	4.064	<0.001
BMI, median [IQR]	25.30 (21.54–28.73)	25.33 (21.45–28.74)	24.54 (21.77–28.65)	0.082	0.935
History					
Smoking (n,%)	292 (77.66%)	219 (77.39%)	73 (78.49%)	0.006	0.937
Alcohol Use (n,%)	98 (26.06%)	76 (26.86%)	22 (23.66%)	0.224	0.636
Diabetes (n,%)	71 (18.88%)	50 (17.67%)	21 (22.58%)	0.806	0.369
Stroke (n,%)	55 (14.63%)	39 (13.78%)	16 (17.20%)	0.411	0.521
Hypertension (n,%)	127 (33.78%)	97 (34.28%)	30 (32.26%)	0.053	0.818
Coronary Artery Disease (n,%)	96 (25.53%)	69 (24.38%)	27 (29.03%)	0.570	0.45
LTDO ₂ (n,%)	83 (22.07%)	70 (24.74%)	13 (13.98%)	4.104	0.043
Laboratory testing					
SIRI (median [IQR])	2.32 (1.59–3.65)	2.11 (1.44–3.39)	2.73 (2.14–4.12)	4.383	<0.001
HCT (% median [IQR])	0.41 (0.37–0.46)	0.41 (0.37–0.45)	0.42 (0.38–0.46)	1.051	0.292
Hb (g/L, median [IQR])	126.00 (115.00–138.00)	128.00 (116.00–138.00)	125.00 (113.00–135.00)	1.232	0.218
RDW (% median [IQR])	14.10 (13.10–14.78)	13.90 (12.90–14.67)	14.56 (13.82–15.38)	4.451	<0.001
PLT ($10^9/L$, $\bar{x} \pm s$)	186.61 $\pm$ 37.30	186.92 $\pm$ 35.52	185.70 $\pm$ 42.47	0.272	0.785

(Continued)

Table 2 (Continued).

Variable	Total (n=376)	Non-PH Group (n=283)	PH Group (n=93)	t, χ^2 /Z value	Pvalue
pH (% median [IQR])	7.40 (7.37–7.43)	7.40 (7.37–7.43)	7.39 (7.36–7.43)	1.456	0.144
PaO ₂ (mmHg, median [IQR])	84.00 (77.00–89.00)	85.00 (78.00–90.50)	78.00 (74.00–84.00)	6.053	<0.001
PaCO ₂ (mmHg, median [IQR])	41.50 (38.00–45.00)	41.00 (38.00–46.00)	42.00 (38.00–45.00)	0.745	0.455
HCO ₃ ⁻ (mmol/L, median [IQR])	24.00 (22.00–26.00)	24.00 (22.00–26.00)	23.00 (21.00–25.00)	1.772	0.075
Lac (mmol/L, median [IQR])	1.77 (1.53–1.99)	1.79 (1.54–2.00)	1.66 (1.48–1.99)	1.132	0.258
D-Dimer (mg/L, median [IQR])	0.44 (0.35–0.53)	0.45 (0.35–0.54)	0.43 (0.35–0.50)	0.845	0.398
FIB (g/L, $\bar{x} \pm s$)	2.59 (2.38–2.84)	2.60 (2.38–2.84)	2.55 (2.37–2.88)	0.160	0.873
FDP (ug/mL, $\bar{x} \pm s$)	3.92 $\pm$ 0.90	3.92 $\pm$ 0.86	3.92 $\pm$ 1.00	0.029	0.977
BNP (ng/L, median [IQR])	160.92 (139.34–187.98)	157.44 (136.58–184.45)	174.74 (153.27–194.32)	3.554	<0.001
PCT (ng/mL, median [IQR])	0.11 (0.09–0.13)	0.11 (0.09–0.13)	0.11 (0.10–0.13)	0.225	0.821
Alb (g/L, median [IQR])	41.35 (36.90–45.73)	42.00 (36.90–45.65)	40.80 (37.10–46.20)	0.465	0.642
BUN (mmol/L, median [IQR])	7.36 (5.77–9.24)	7.33 (5.68–9.34)	7.50 (6.01–8.89)	0.007	0.995
Crea (μ mol/L, median [IQR])	70.50 (54.00–88.00)	70.00 (53.50–88.00)	72.00 (56.00–88.00)	0.403	0.687
UA (μ mol/L, $\bar{x} \pm s$)	297.63 $\pm$ 54.33	298.28 $\pm$ 56.10	295.63 $\pm$ 48.79	0.407	0.684
CRP (mg/L, $\bar{x} \pm s$)	15.46 $\pm$ 6.04	15.51 $\pm$ 6.18	15.31 $\pm$ 5.62	0.277	0.782
LDL-C (mmol/L, $\bar{x} \pm s$)	2.80 $\pm$ 0.32	2.81 $\pm$ 0.32	2.76 $\pm$ 0.30	1.214	0.225
HDL-C (mmol/L, $\bar{x} \pm s$)	1.37 $\pm$ 0.46	1.37 $\pm$ 0.47	1.37 $\pm$ 0.42	0.012	0.991

Abbreviations: BMI, Body Mass Index; LTDO₂, Long-Term Domiciliary Oxygen Therapy; SIRI, Systemic Inflammatory Response Index; HCT, Hematocrit; Hb, Hemoglobin; RDW, Red Blood Cell Distribution Width; PLT, Platelet; FIB, Fibrinogen; FDP, Fibrinogen Degradation Products; BNP, B-type Natriuretic Peptide; PCT, Procalcitonin; Alb, Albumin; BUN, Blood Urea Nitrogen; Crea, Creatinine; UA, Uric Acid; CRP, C-reactive Protein; LDL-C, Low-Density Lipoprotein Cholesterol; HDL-C, High-Density Lipoprotein Cholesterol.

Construction of the Prediction Model

Lasso Regression Analysis Variable Screening

Using the presence or absence of concurrent PH in the model group as the dependent variable, all independent variables were included in LASSO regression analysis to identify the most representative predictors of PH. The coefficient path diagram shows that as the penalty coefficient λ varies, the independent variable coefficients are progressively compressed to zero (Figure 2A). The cross-validation error plot presents the mean squared error (MSE) corresponding to different λ values, where the left dashed line represents the λ value corresponding to the minimum MSE (NULL), and the right dashed line represents the λ value one standard error higher than the minimum MSE (0.045), ie, the optimal penalty coefficient λ_{1se} of the model. Under this optimal coefficient, the model screened out 6 variables with non-zero coefficients as the optimal feature subset, specifically including: COPD disease duration, home oxygen therapy, SIRI, RDW, arterial oxygen partial pressure, and BNP (Figure 2B).

Establishment of the Visualized Nomogram Model

The six variables screened by Lasso regression were incorporated into a multivariate logistic regression analysis. The results demonstrated that Duration of COPD (OR = 1.085, 95% CI: 1.043–1.129; $P < 0.001$), SIRI (OR = 1.214, 95% CI: 1.043–1.413; $P = 0.012$), RDW (OR = 1.696, 95% CI: 1.317–2.184; $P < 0.001$), and BNP (OR = 1.011, 95% CI: 1.003–1.018; $P = 0.005$) were risk factors for PH, whereas LTDO₂ (OR = 0.461, 95% CI: 0.219–0.972; $P = 0.041$) and PaO₂ (OR = 0.912, 95% CI: 0.880–0.946; $P < 0.001$) were protective factors for PH (Table 3). A nomogram model was constructed based on these six predictors. Each independent predictor was scored according to the top scale (Points). The total score of the patient was the sum of all risk factor scores. The value on the risk axis corresponding to the total score represented the probability of COPD patients developing PH. A higher total score indicated a greater risk of PH in COPD patients (Figure 3).

Evaluation and Validation of Model Predictive Power

The ROC curve showed that the area under the curve (AUC) for the Nomogram model was 0.812 (95% CI: 0.762–0.863), whereas the AUC for the validation set was 0.826 (95% CI: 0.754–0.897). The Youden's index for this

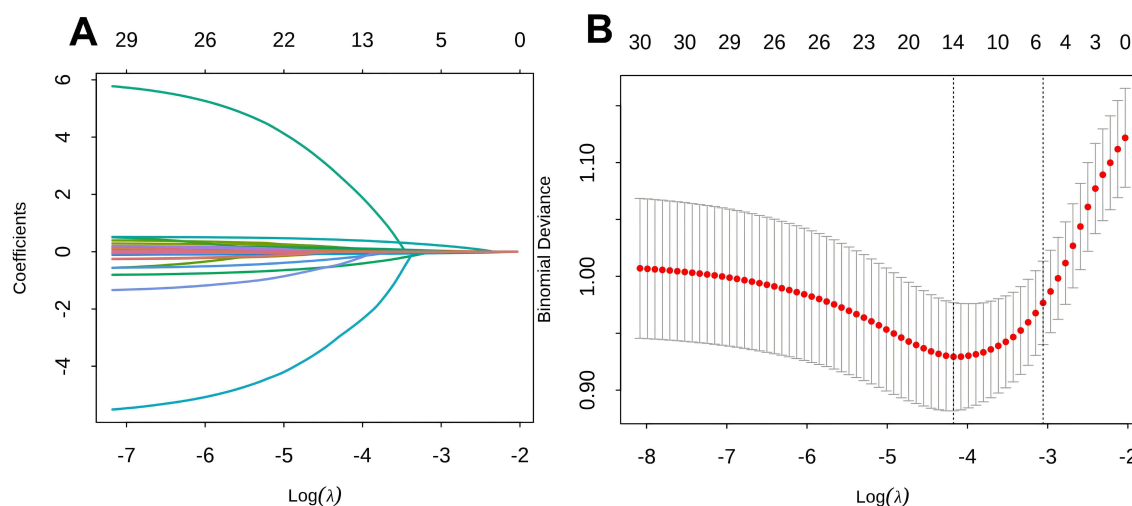


Figure 2 Variable Selection via LASSO Regression: (A) Distribution of LASSO Regression Coefficients for Candidate Predictor Variables. (B) Selection of the Optimal Penalty Coefficient in the LASSO Regression.

predictive model was 0.526, with sensitivities and specificities of 0.699 and 0.827, respectively (Figure 4A and B). The Hosmer-Lemeshow revealed a p-value of 0.276 for the model group and 0.162 for the validation group ($p > 0.05$), indicating a good fit of the model. The calibration curve demonstrated high agreement between the actual curve and the bias-corrected curve, with a cutoff value of 0 and a slope of 1, indicating strong consistency between the model and observed values and good calibration. The calibration curve for the validation set also showed excellent fit (Figure 5A and B). The clinical impact curve (CIC) was used to evaluate the clinical utility of the columnar predictive model. The X-axis represents the high-risk threshold, and the Y-axis represents the number of PH per 1000 individuals. The red line indicates the number of PH predicted by the model, while the blue line represents the actual number of PH cases. When the risk threshold was 0.6, the predicted number was close to the actual number, with low rates of false positives and false negatives, suggesting good predictive performance and high clinical utility of the model (Figure 6A and B).

Discussion

We identified six key predictors from a large number of clinical indicators through Lasso regression: Duration of COPD, LTDO2, PaO2, RDW, SIRI, and BNP, and constructed a comprehensive Nomogram predictive model. All indicators are commonly used in clinical practice and are significantly beneficial for COPD treatment. The developed Nomogram predictive model demonstrates strong clinical applicability. It assists clinicians and community physicians in rapidly identifying high-risk COPD patients at risk of developing PH, supporting treatment selection, prognosis management, and rehabilitation planning decisions. Additionally, it facilitates discussions between physicians and patients/families regarding disease risk prognosis and treatment options.

Table 3 Multivariate Regression Analysis of PH

Variable	OR	Coefficient	95% CI.low	95% CI.upp	p.value
Duration of COPD	1.085	0.081	1.043	1.129	< 0.001
LTDO2	0.461	-0.774	0.219	0.972	0.041
SIRI	1.214	0.194	1.043	1.413	0.012
RDW	1.696	0.528	1.317	2.184	< 0.001
PaO2	0.912	-0.092	0.880	0.946	< 0.001
BNP	1.011	0.011	1.003	1.018	0.005

Abbreviations: LTDO2, Long-Term Domiciliary Oxygen Therapy; SIRI, Systemic Inflammatory Response Index; RDW, Red Blood Cell Distribution Width; BNP, B-type Natriuretic Peptide.

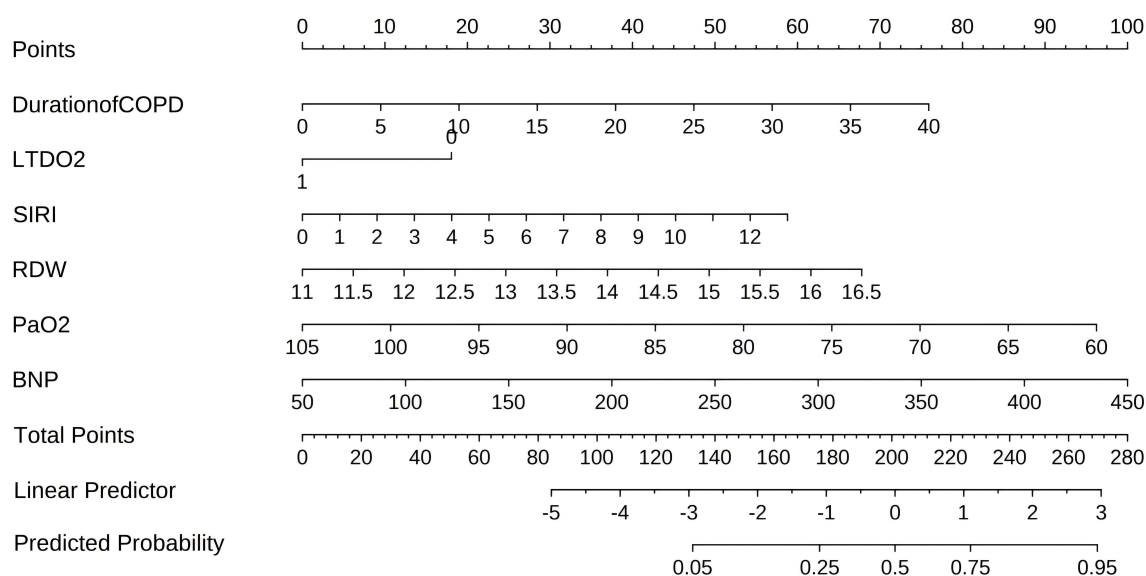


Figure 3 Nomogram predicting pulmonary hypertension in patients with chronic obstructive pulmonary disease.

Abbreviations: LTDO2, Long-Term Domiciliary Oxygen Therapy; SIRI, Systemic Inflammatory Response Index; RDW, Red Blood Cell Distribution Width; BNP, B-type Natriuretic Peptide.

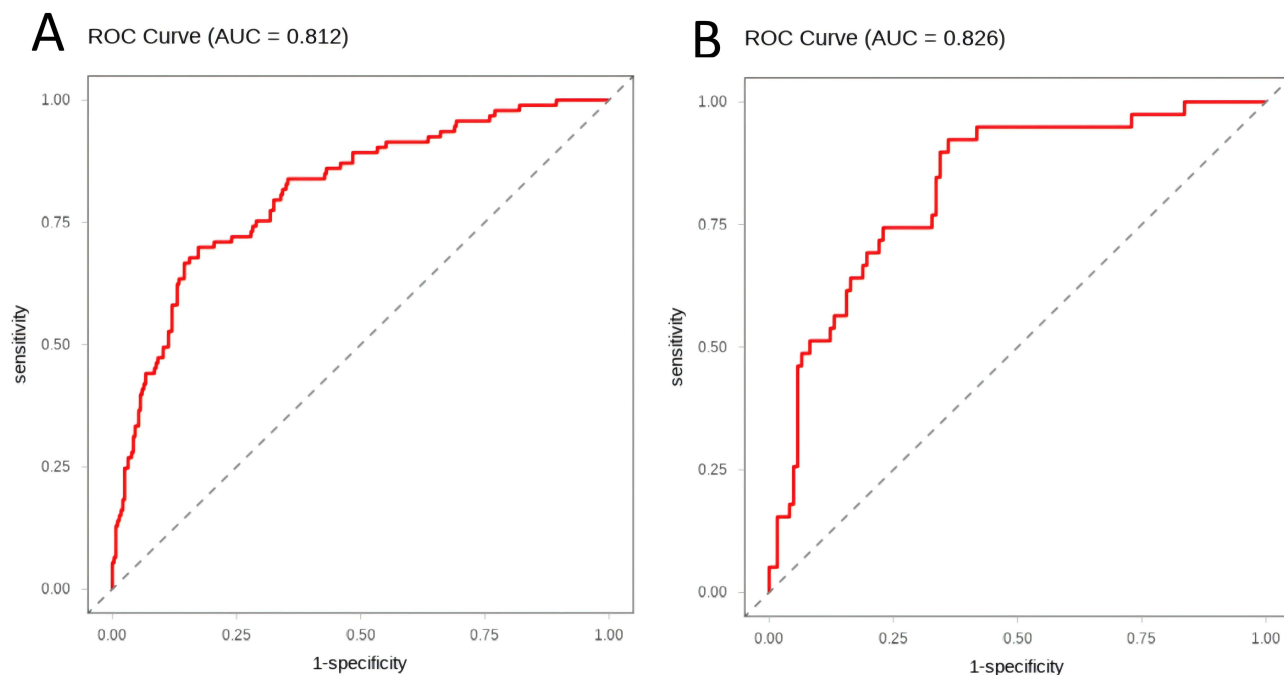


Figure 4 Receiver operating characteristic (ROC) curve: (A) Model group; (B) Validation group.

In patients with COPD, the duration of the disease is closely associated with the development and progression of pulmonary hypertension (PH). As the course of COPD prolongs, pulmonary ventilation and gas exchange function progressively decline. Chronic hypoxemia can induce pulmonary vasoconstriction and remodeling, increase pulmonary circulatory resistance, and subsequently trigger PH.^{11,12} From a pathophysiological perspective, hypoxemia promotes the activation of key transcription factors and signaling pathways, thereby exacerbating inflammatory responses and

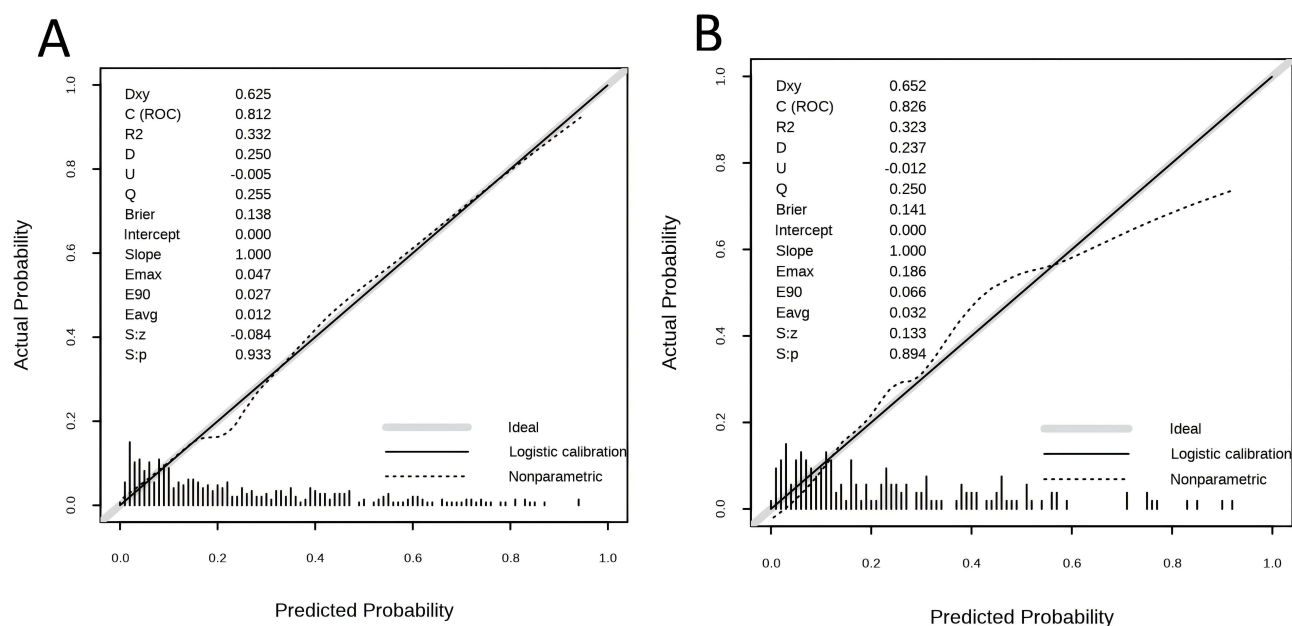


Figure 5 Calibration curves: (A) Model group; (B) Validation group.

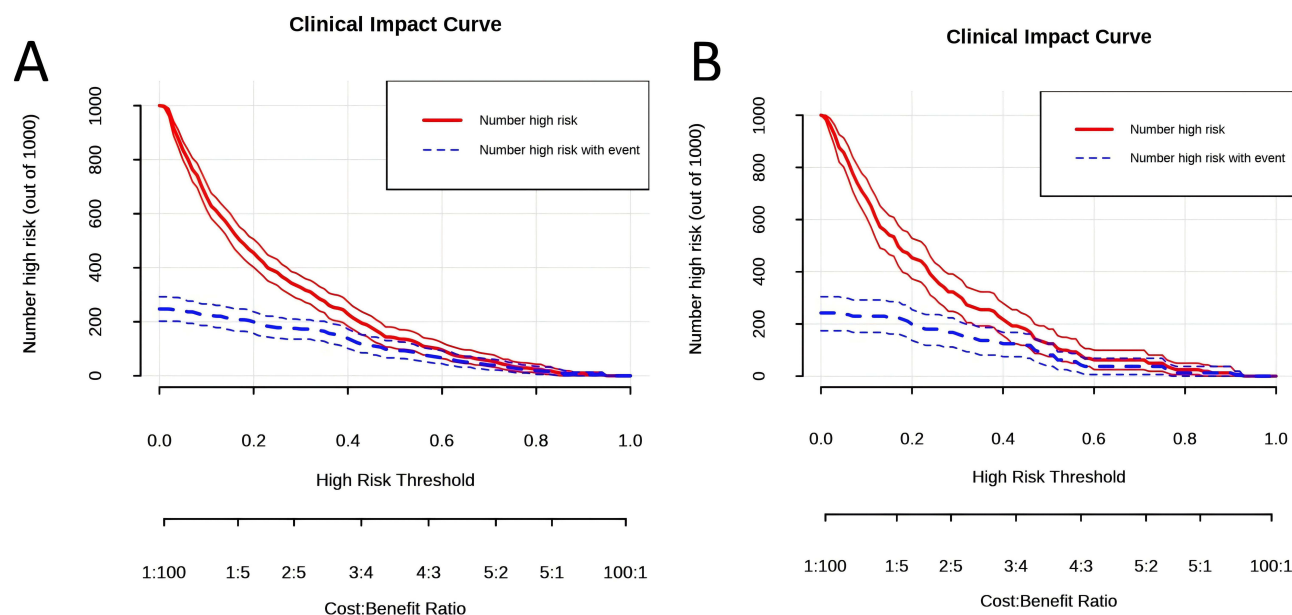


Figure 6 Clinical impact curves (CIC): (A) Model group; (B) Validation group.

neutrophil infiltration, leading to pulmonary vascular endothelial cell damage and disruption of normal vascular structure and function.¹³ Meanwhile, the reduction of the pulmonary capillary bed due to destruction of the lung parenchyma further increases pulmonary circulatory pressure. Studies have shown that small pulmonary artery and microvascular remodeling in patients with COPD-PH may occur prior to the clinical diagnosis of PH and is associated with dysregulation of the miR-126/ADAM9 axis.¹⁴ As COPD progresses, respiratory function gradually deteriorates. Lung hyperinflation reduces intrathoracic blood volume and cardiac filling pressure, and mechanically compresses the right

ventricle. Expiratory airflow limitation leads to expiratory phase blood flow restriction in the pulmonary circulation, thereby impairing stroke volume at rest and during exercise.¹⁵ These hemodynamic changes contribute to heart failure and promote the development of PH.

Clinically, COPD patients, particularly those with prolonged disease duration, should be highly vigilant for PH. Regular monitoring through diagnostic tests such as echocardiography and electrocardiography is essential for early detection of PH. LTDO2 serves as a protective factor against PH, a finding of significant clinical importance. For severe COPD patients with long-standing hypercapnia, home oxygen therapy has been actively recommended to alleviate symptoms and reduce mortality.^{16–18} Recent studies suggest that LTDO2 has minimal or no significant impact on 3-year mortality in COPD patients with mild to moderate hypoxemia.¹⁸ However, that study did not include patients with severe disease, nor did it evaluate disease progression in the included patients. Despite the above controversy, this study suggests that home oxygen therapy should be considered for COPD patients even with mild-to-moderate hypoxemia, from the perspective of reducing PH risk. LTDO2 may increase oxygen saturation, alleviate hypoxemia, reduce pulmonary vasoconstriction and remodeling, and lower pulmonary circulatory resistance, thereby decreasing the risk of PH.¹⁹ In clinical practice, LTDO2 should be actively recommended for COPD patients, particularly those at risk of PH. However, the current adoption rate of LTDO2 remains suboptimal in real-world settings. Some patients may fail to receive standardized LTDO2 due to financial constraints or insufficient awareness of oxygen therapy. Therefore, healthcare providers should enhance health education to raise patients' understanding of the importance of LTDO2, guide them in the proper use of oxygen equipment, and ensure the efficacy and safety of the treatment.

Perivascular inflammation is a significant pathological feature in most patients with PH. Increasing evidence suggests that perivascular inflammation plays a functional role in the initiation and/or progression of PH, as well as in pulmonary vascular remodeling.²⁰ The pathophysiological mechanisms of pulmonary vascular remodeling in PH involve not only pathological damage to endothelial cell function but also excessive infiltration of inflammatory cells into the perivascular space.²¹ RDW, as a quantification of red blood cell size heterogeneity, is a routinely reported laboratory parameter in complete blood count and is primarily used for the differential diagnosis of anemia. RDW serves as a biomarker of red blood cell dysfunction associated with high systemic inflammation²² and is also a mortality predictor for various diseases, such as acute pancreatitis and heart failure.²³ Recent studies have found that elevated RDW is correlated with the severity of acute exacerbation of chronic obstructive pulmonary disease²⁴ and mortality risk.²⁵ SIRI is a well-recognized biomarker of systemic inflammatory and stress responses. As a composite indicator, SIRI can assess systemic inflammation and disease prognosis more accurately than any single count of neutrophils, lymphocytes, or monocytes. In recent years, SIRI has been identified as a critical prognostic marker in cardiovascular diseases, tumors, and infectious diseases.^{26–28} Many immune - inflammatory biomarkers, such as neutrophil - to - lymphocyte ratio (NLR),²⁹ platelet - to - lymphocyte ratio (PLR),²⁹ RDW,³⁰ and Systemic Inflammatory Index (SII),^{29–31} have demonstrated predictive value for PH. SIRI can independently predict the severity and prognosis of PH, is associated with multiple PH severity indicators, and serves as a significant predictor of clinical deterioration, with higher predictive value than indicators such as SII, PLR, and NLR.³² For clinical application, RDW and SIRI, as routine testing indicators, offer the advantages of easy acquisition and low cost. Incorporating them into the predictive model for COPD with PH can provide clinicians with a more comprehensive and convenient assessment tool. Through simple complete blood count (CBC) tests, physicians can more sensitively detect the potential risk of PH in COPD patients, thereby enabling early intervention measures to improve patient outcomes.

Some studies on COPD-PH have excluded oxygen partial pressure due to concerns about the effect of inpatient oxygen therapy on blood gas results.²⁹ At our center, COPD patients admitted via the emergency department generally undergo routine blood tests and arterial blood gas analysis in the emergency setting, while those admitted through outpatient clinics undergo these tests immediately upon hospital admission. Therefore, the blood gas results are minimally affected by oxygen therapy. The core pathological feature of COPD is chronic hypoxia. Following alveolar hypoxia, the body initially constricts intrapulmonary arteries to redirect blood flow to better-oxygenated lung segments, thereby optimizing the ventilation/perfusion ratio and enhancing pulmonary oxygenation.³³ Hypoxia induces increased secretion of vascular endothelial growth factor (VEGF), leading to endothelial dysfunction, abnormal proliferation of

pulmonary artery smooth muscle cells, and vascular fibrosis.³⁴ Hypoxia also triggers various epigenetic modifications, including RNA methylation, DNA methylation, histone modification, and non-coding RNA molecular changes, which block programmed cell death pathways such as apoptosis, autophagy, and ferroptosis, thereby promoting the development of PH.³⁴

This study identified BNP as a risk factor for PH. In PH, ventricular cells are stimulated by transmural pressure, volume overload, hypoxia, or pro-inflammatory factors, inducing the transcription of the NPPB gene. The resulting preproBNP is subsequently cleaved in the sarcoplasmic reticulum into proBNP. Upon secretion into the bloodstream, proBNP is cleaved by serine proteases into the biologically active BNP and the non-biologically active NT-proBNP.³⁵ Multiple studies have confirmed the correlation between circulating BNP and NT-proBNP levels and pulmonary hemodynamics in PH patients: NT-proBNP levels are associated with mean pulmonary artery pressure (mPAP), pulmonary vascular resistance (PVR), right atrial pressure, and cardiac index.^{36–38}

The nomogram predictive model established in this study can effectively predict the risk of PH in COPD patients. However, since this is a single-center study, larger sample sizes and multicenter, prospective randomized controlled trials are still required to validate and optimize our findings. In summary, the nomogram can serve as a non-invasive clinical tool for predicting the risk of PH in COPD patients, demonstrating good clinical utility.

Conclusion

In conclusion, we constructed and internally validated a nomogram model incorporating six easily accessible clinical variables to predict PH risk in COPD patients. The model showed good performance and clinical utility, potentially aiding early identification and individualized management of high-risk patients. External validation in multicenter cohorts is needed.

Data Sharing Statement

This study supports the non-disclosure of data due to the inclusion of information that may compromise the privacy of study participants, but such data may be obtained from the corresponding author upon reasonable request.

Ethical Statement

This research protocol was reviewed and approved by Anhui NO.2 Provincial People's Hospital (Approval No: (R) 2025-152).

Acknowledgments

The authors would like to thank Associate Professor Changhong Yuan from the Department of Neurology, Anhui NO.2 Provincial People's Hospital, for his statistical assistance.

Disclosure

The authors declare no conflicts of interest in this work.

References

- Agusti A, Böhm M, Celli B, et al. GOLD COPD DOCUMENT 2023: a brief update for practicing cardiologists. *Clin Res Cardiol.* 2024;113(2):195–204. doi:10.1007/s00392-023-02217-0
- Soriano JB, Kendrick PJ, Paulson KR; GBD Chronic Respiratory Disease Collaborators. Prevalence and attributable health burden of chronic respiratory diseases, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Respir Med.* 2020;8(6):585–596. doi:10.1016/S2213-2600(20)30105-3
- Olsson KM, Corte TJ, Kamp JC, et al. Pulmonary hypertension associated with lung disease: new insights into pathomechanisms, diagnosis, and management. *Lancet Respir Med.* 2023;11(9):820–835. doi:10.1016/S2213-2600(23)00259-X
- Huang H, Yang L. Research progress of inflammatory factors in chronic obstructive pulmonary disease with pulmonary hypertension at high altitude. *Altern Ther Health Med.* 2023;29(8):689–693.
- Wang Z, Ma L, Liu M, Fan J, Hu S; Writing Committee of the Report on Cardiovascular Health and Diseases in China. Summary of the 2022 report on cardiovascular health and diseases in China. *Chin Med J.* 2023;136(24):2899–2908. doi:10.1097/CM9.0000000000000297
- Weitzenblum E, Chaouat A. Cor pulmonale. *Chron Respir Dis.* 2009;6(3):177–185. doi:10.1177/1479972309104664
- Prins KW, Rose L, Archer SL, et al. Disproportionate right ventricular dysfunction and poor survival in group 3 pulmonary hypertension. *Am J Respir Crit Care Med.* 2018;197(11):1496–1499. doi:10.1164/rccm.201712-2405LE

8. Minai OA, Chaouat A, Adnot S. Pulmonary hypertension in COPD: epidemiology, significance, and management: pulmonary vascular disease: the global perspective. *Chest*. 2010;137(6 Suppl):39S–51S. doi:10.1378/chest.10-0087
9. Agustí A, Celli BR, Criner GJ, et al. Global initiative for chronic obstructive lung disease 2023 report: GOLD executive summary. *Eur Respir J*. 2023;61(4):2300239. doi:10.1183/13993003.00239-2023
10. Humbert M, Kovacs G, Hoeper MM, et al. 2022 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension. *Eur Respir J*. 2023;61(1):3618–3731.
11. Zhou D, Liu C, Wang L, et al. Prediction of clinical risk assessment and survival in chronic obstructive pulmonary disease with pulmonary hypertension. *Clin Transl Med*. 2024;14(6):e1702. doi:10.1002/ctm2.1702
12. Chai T, Qiu C, Xian Z, Lu Y, Zeng Y, Li J. A narrative review of research advances in hypoxic pulmonary hypertension. *Ann Transl Med*. 2022;10(4):230. doi:10.21037/atm-22-259
13. Zangiabadi A, De Pasquale CG, Sajkov D. Pulmonary hypertension and right heart dysfunction in chronic lung disease. *Biomed Res Int*. 2014;2014:739674. doi:10.1155/2014/739674
14. Goel K, Egersdorf N, Gill A, et al. Characterization of pulmonary vascular remodeling and MicroRNA-126-targets in COPD-pulmonary hypertension. *Respir Res*. 2022;23(1):349. doi:10.1186/s12931-022-02267-4
15. Tseng S, Stanziola AA, Sultan S, Henry K, Saggat R, Rajan S. Pulmonary hypertension related to chronic obstructive pulmonary disease and diffuse parenchymal lung disease: a focus on right ventricular (Dys) Function. *Heart Fail Clin*. 2018;14(3):403–411. doi:10.1016/j.hfc.2018.03.006
16. Khor YH, Ekström M. The benefits and drawbacks of home oxygen therapy for COPD: what's next. *Expert Rev Respir Med*. 2024;18(7):469–483. doi:10.1080/17476348.2024.2379459
17. Freitas C, Serino M, Araújo D, et al. Long-term home non-invasive ventilation in chronic hypercapnic chronic obstructive pulmonary disease patients: real-world impact on lung function, acute exacerbations and survival. *Clin Respir J*. 2021;15(10):1130–1138. doi:10.1111/crj.13419
18. Lacasse Y, Casaburi R, Sliwinski P, et al. Home oxygen for moderate hypoxaemia in chronic obstructive pulmonary disease: a systematic review and meta-analysis. *Lancet Respir Med*. 2022;10(11):1029–1037. doi:10.1016/S2213-2600(22)00179-5
19. McCabe C. Therapeutic use of oxygen in pulmonary arterial hypertension: a forgotten friend. *Eur J Prev Cardiol*. 2021;28(1):124–126. doi:10.1093/eurjpc/zwaa020
20. Hu Y, Chi L, Kuebler WM, Goldenberg NM. Perivascular inflammation in pulmonary arterial hypertension. *Cells*. 2020;9(11):2338. doi:10.3390/cells9112338
21. Humbert M, Guignabert C, Bonnet S, et al. Pathology and pathobiology of pulmonary hypertension: state of the art and research perspectives. *Eur Respir J*. 2019;53(1):1801887. doi:10.1183/13993003.01887-2018
22. García-Escobar A, Lázaro-García R, Goicolea-Ruigómez J, et al. Red blood cell distribution width is a biomarker of red cell dysfunction associated with high systemic inflammation and a prognostic marker in heart failure and cardiovascular disease: a potential predictor of atrial fibrillation recurrence. *High Blood Press Cardiovasc Prev*. 2024;31(5):437–449. doi:10.1007/s40292-024-00662-0
23. Xanthopoulos A, Giamouzis G, Dimos A, et al. Red blood cell distribution width in heart failure: pathophysiology, prognostic role, controversies and dilemmas. *J Clin Med*. 2022;11(7):1951. doi:10.3390/jcm11071951
24. Huang Y, Wang J, Chen S, et al. Red cell distribution width related to the severity of the disease in patients with AECOPD. *Biomarker Med*. 2025;19(5):177–186. doi:10.1080/17520363.2025.2468152
25. Sheng P, Lei W, Cheng T, et al. The red cell distribution width-to-albumin ratio predicts 1-year mortality in patients with chronic obstructive pulmonary disease: evidence from MIMIC-IV and NHANES. *Int J Chron Obstruct Pulmon Dis*. 2025;20:3351–3360. doi:10.2147/COPD.S543479
26. Wu Y, Zhao J, Wang Z, et al. Association of systemic inflammatory markers and tertiary lymphoid structure with pathological complete response in gastric cancer patients receiving preoperative treatment: a retrospective cohort study. *Int J Surg*. 2023;109(12):4151–4161. doi:10.1097/JS9.0000000000000741
27. Han K, Shi D, Yang L, et al. Prognostic value of systemic inflammatory response index in patients with acute coronary syndrome undergoing percutaneous coronary intervention. *Ann Med*. 2022;54(1):1667–1677. doi:10.1080/07853890.2022.2083671
28. Xia Y, Xia C, Wu L, Li Z, Li H, Zhang J. Systemic Immune Inflammation Index (SII), System Inflammation Response Index (SIRI) and risk of all-cause mortality and cardiovascular mortality: a 20-year follow-up cohort study of 42,875 US adults. *J Clin Med*. 2023;12(3):1128. doi:10.3390/jcm12031128
29. Zuo H, Xie X, Peng J, Wang L, Zhu R. Predictive value of novel inflammation-based biomarkers for pulmonary hypertension in the acute exacerbation of chronic obstructive pulmonary disease. *Anal Cell Pathol*. 2019;2019:5189165. doi:10.1155/2019/5189165
30. Hu L, Liu Z, Yu J, Yang Z, Feng D. The correlation between NLR, RDW, and pulmonary hypertension in patients with bronchiectasis and chronic obstructive pulmonary disease overlap syndrome. *Clin Respir J*. 2025;19(10):e70128. doi:10.1111/crj.70128
31. Dziedzic EA, Gąsior JS, Tuzimek A, et al. Investigation of the associations of novel inflammatory biomarkers—Systemic Inflammation Index (SII) and Systemic Inflammation Response Index (SIRI)—with the severity of coronary artery disease and acute coronary syndrome occurrence. *Int J Mol Sci*. 2022;23(17):9553. doi:10.3390/ijms23179553
32. Gao L, Zhang S, Zhao Z, et al. Role of the systemic inflammatory response index in predicting disease severity and prognosis in idiopathic pulmonary arterial hypertension. *J Inflamm Res*. 2024;17:447–460. doi:10.2147/JIR.S434720
33. Grimmer B, Krauszman A, Hu X, et al. Pannexin 1: a novel regulator of acute hypoxic pulmonary vasoconstriction. *Cardiovasc Res*. 2022;118(11):2535–2547. doi:10.1093/cvr/cvab326
34. Jiang Y, Song S, Liu J, et al. Epigenetic regulation of programmed cell death in hypoxia-induced pulmonary arterial hypertension. *Front Immunol*. 2023;14:1206452. doi:10.3389/fimmu.2023.1206452
35. Lewis RA, Durrington C, Condliffe R, Kiely DG. BNP/NT-proBNP in pulmonary arterial hypertension: time for point-of-care testing. *Eur Respir Rev*. 2020;29(156):200009. doi:10.1183/16000617.0009-2020
36. Fu S, Ping P, Wang F, Luo L. Synthesis, secretion, function, metabolism and application of natriuretic peptides in heart failure. *J Biol Eng*. 2018;12:2. doi:10.1186/s13036-017-0093-0
37. Hendriks PM, van de Groep LD, Veen KM, et al. Prognostic value of brain natriuretic peptides in patients with pulmonary arterial hypertension: a systematic review and meta-analysis. *Am Heart J*. 2022;250:34–44. doi:10.1016/j.ahj.2022.05.006
38. Gan CT, McCann GP, Marcus JT, et al. NT-proBNP reflects right ventricular structure and function in pulmonary hypertension. *Eur Respir J*. 2006;28(6):1190–1194. doi:10.1183/09031936.00016006

International Journal of General Medicine

Publish your work in this journal

The International Journal of General Medicine is an international, peer-reviewed open-access journal that focuses on general and internal medicine, pathogenesis, epidemiology, diagnosis, monitoring and treatment protocols. The journal is characterized by the rapid reporting of reviews, original research and clinical studies across all disease areas. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/international-journal-of-general-medicine-journal>

Dovepress
Taylor & Francis Group