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.

Association of Eosinophil Count and Neutrophil Percentage-to-Albumin Ratio with 1-Year Readmission Risk in Patients with COPD and Comorbid Small Airway Dysfunction: A Retrospective Cohort Study with Machine Learning-Based Feature Selection

Ying Lu
Northern Jiangsu People's Hospital Affiliated to Yangzhou University
Rui Xu
Northern Jiangsu People's Hospital
Yijie Chen
Northern Jiangsu People's Hospital Affiliated to Yangzhou University
Yangyang Yang
Northern Jiangsu People's Hospital Affiliated to Yangzhou University
Yongqi Zhu
Northern Jiangsu People's Hospital
Lingfeng Min
minlingfeng@126.com

Northern Jiangsu People's Hospital Affiliated to Yangzhou University

Research Article

Keywords: chronic obstructive pulmonary disease, small airway dysfunction, eosinophil count, neutrophil percentage-to-albumin ratio, prediction model

Posted Date: October 22nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11147747/v1

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

Additional Declarations: No competing interests reported.

Abstract

Background

Patients with chronic obstructive pulmonary disease (COPD) and concomitant small airway dysfunction (SAD) exhibit more severe manifestations and increased exacerbation risk. However, robust biomarkers for predicting 1-year readmission in this population remain scarce. This study aimed to investigate associations of blood eosinophil count (Eos) and neutrophil percentage-to-albumin ratio (NPAR) with 1-year unplanned readmission for COPD exacerbations in patients with COPD and SAD, and to develop a machine learning-based model for individualized risk stratification.

Methods

This single-center retrospective cohort study consecutively enrolled patients diagnosed with both COPD and SAD between January 2016 and December 2023. Multivariable logistic regression and restricted cubic spline (RCS) analyses assessed the associations of Eos and NPAR with readmission risk and to evaluate potential nonlinear dose–response relationships. Prespecified subgroup analyses and sensitivity were conducted to evaluate the robustness and generalizability of the findings. Five machine learning algorithms—decision tree (DT), extreme gradient boosting (XGBoost), support vector machine (SVM), K-nearest neighbors (KNN), and random forest (RF)—were trained and validated to construct predictive models. Model discrimination was evaluated using the area under the receiver operating characteristic curve (AUC) and the area under the precision–recall curve (AUPRC). Model interpretability was enhanced through SHapley Additive exPlanations (SHAP) analysis.

Results

Among 768 patients, 342 (44.531%) had 1-year readmission. Eos (OR = 1.269, $P < 0.001$) and NPAR (OR = 6.283, $P < 0.001$) were independently associated with readmission risk. Eos showed a significant dose-response trend ($P < 0.001$; threshold: 0.25×10^9/L), while NPAR did not ($P > 0.05$). The RF model achieved the best performance (AUC = 0.801), and combining Eos with NPAR yielded the highest AUC of 0.802 (ΔAUC = 0.101, $P = 0.001$). SHAP identified Eos and NPAR as top predictors. Sensitivity and subgroup analyses supported the robustness of these findings.

Conclusion

In this single-center retrospective cohort, higher Eos and NPAR were associated with 1-year readmission in COPD patients with SAD. Their combined model showed internal discriminatory performance and requires external prospective validation before clinical implementation.

Background

Chronic obstructive pulmonary disease (COPD) remains a leading cause of morbidity and mortality worldwide, imposing substantial clinical and economic burdens on healthcare systems [1, 2]. Its pathophysiology is characterized by repeated injury and repair of the small airways and lung parenchyma, leading to early structural remodeling and alveolar destruction. These changes drive persistent airflow limitation, progressive decline in lung function, chronic respiratory symptoms, and recurrent acute exacerbations, which collectively impair patients' quality of life and accelerate disease progression [3, 4]. One-year hospital readmission is a critical indicator of disease burden and healthcare quality, and serves as an important benchmark for evaluating national health system performance. Although current therapeutic strategies have demonstrated efficacy in alleviating acute symptoms, post-discharge long-term management remains challenging. The burden following discharge for acute exacerbation of COPD (AECOPD) is substantial: approximately 15%–30% of patients are readmitted within 30 days due to recurrent exacerbations, with a pooled estimate of 18% [5–7]; this proportion rises to approximately 31% at 90 days [7, 8] and exceeds 50% at 1 year, with some studies reporting rates as high as 67% [9]. Frequent exacerbations not only accelerate lung

function decline but are also independently associated with increased all-cause mortality risk [5, 6, 10]. Thus, identifying reliable and readily accessible biomarkers to accurately predict readmission risk is essential for guiding clinical decision-making, optimizing post-discharge care strategies, and reducing healthcare expenditure.

In recent years, small airway dysfunction (SAD) has been increasingly recognized as an important clinical phenotype and core pathological feature of COPD, and has become a key determinant of disease severity and prognosis. Small airways are defined as bronchioles with an internal diameter of < 2 mm. Although they account for approximately 98% of total lung volume, they contribute only about 10% of total airway resistance under normal conditions, earning them the designation "silent zone" of the lung [11]. Multiple pathological mechanisms—including epithelial injury, airway wall fibrosis, smooth muscle dysfunction, mucus plugging, and external compression—contribute to the development of SAD [12, 13]. SAD is now acknowledged as an early and prevalent feature in patients with COPD [14], with its prevalence increasing alongside disease severity [15, 16], and is associated with airflow obstruction, air trapping, and enhanced inflammatory responses. Notably, COPD patients with coexisting SAD exhibit more severe clinical symptoms, greater lung function decline, higher exacerbation risk, and poorer treatment responses compared with those without SAD [17, 18], underscoring the urgent need for improved risk stratification and individualized management strategies in this high-risk subgroup.

Eosinophil count (Eos) has garnered extensive attention as a biomarker reflecting type 2 airway inflammation in COPD. Increasing evidence points to a type 2 inflammatory phenotype characterized by elevated eosinophils in approximately 20%–40% of COPD patients [19]. However, the prognostic role of Eos in COPD remains controversial. Substantial evidence suggests that elevated Eos is a risk factor for poor outcomes: meta-analyses and large cohort studies have confirmed that blood Eos ≥ 300 cells/μL predicts moderate-to-severe exacerbations and is associated with accelerated lung function decline, increased hospitalization, and higher mortality [20, 21]. Recent studies have further uncovered the potential value of Eos in assessing SAD. Abdo et al. found that airway eosinophilia correlates significantly with SAD in mild COPD and at-risk populations, implicating eosinophilic inflammation in the early pathogenesis of SAD [22]. Similarly, Wu et al., in a prospective Chinese cohort, reported that COPD patients with blood Eos ≥ 300 cells/μL exhibit more severe SAD, a higher burden of chronic respiratory symptoms, and a significantly elevated risk of exacerbations over 2-year follow-up [23]. Conversely, other studies have suggested a potential protective role for Eos. For instance, COPD patients with an eosinophilic phenotype receiving long-term non-invasive ventilation showed prolonged overall survival [24]. A Danish cohort study linked Eos < 150 cells/μL to higher mortality, and Eos < 50 cells/μL to longer hospital stays and lower 1-year survival [25]. In a Korean COPD cohort, persistently high Eos (≥ 300 cells/μL) was associated with better survival than persistently low Eos (< 150 cells/μL or < 2%), while low Eos correlated with more frequent hospitalizations and worse clinical outcomes [25, 26]. These discordant findings suggest that the prognostic significance of Eos in COPD may be modulated by disease stage, treatment context, and other factors, and its value specifically in COPD patients with SAD remains to be defined. Therefore, the present study was designed to evaluate the predictive utility of Eos for 1-year readmission in this specific population, aiming to inform precise risk stratification and individualized therapeutic decision-making.

The neutrophil percentage-to-albumin ratio (NPAR), as a composite biomarker integrating systemic inflammatory burden and nutritional status, has emerged as a promising prognostic indicator across various diseases, including cardiovascular disorders, sepsis, and malignancies, in recent years. Unlike single-parameter indices, NPAR simultaneously captures neutrophil-driven inflammation and hypoalbuminemia-related malnutrition—two interrelated processes frequently observed in hospitalized COPD patients, each independently linked to adverse outcomes. A study utilizing data from the National Health and Nutrition Examination Survey (NHANES) 2011–2018 demonstrated that elevated NPAR was significantly associated with all-cause mortality, cardiovascular mortality, and chronic lower respiratory disease mortality in community-dwelling COPD patients, with adjusted hazard ratios of 1.14, 1.15, and 1.16, respectively. Moreover, NPAR outperformed other blood-based inflammatory markers in predicting 5-year all-cause mortality [27]. Another study based on the MIMIC-IV database further identified NPAR as an independent risk factor for in-hospital mortality in critically ill COPD patients: each 1-unit increase in NPAR corresponded to a 5% increase in in-hospital death risk, and patients in the highest versus lowest quartile exhibited a significantly elevated risk [28]. Although the prognostic utility of NPAR in the general COPD population and in severe cases has gained preliminary recognition, its value in predicting 1-year readmission specifically among COPD patients with concomitant SAD remains unexplored. Given

that SAD represents an early core pathology of COPD characterized by inflammation predominantly affecting the small airways, the systemic inflammation and nutritional imbalance reflected by NPAR may plausibly contribute to its progression. Furthermore, the potential synergistic or complementary predictive value of combining NPAR with Eos—two readily available biomarkers—has yet to be systematically investigated.

In this context, the present study was designed to systematically evaluate the independent associations of Eos and NPAR with 1-year readmission risk in COPD patients with SAD, and to develop and validate machine learning-based prediction models incorporating these two biomarkers. These efforts are expected to enhance risk stratification, facilitate timely identification of high-risk individuals, guide targeted interventions, and ultimately slow disease progression and improve clinical outcomes.

Methods

Study Design and Participants

This single-center retrospective cohort study consecutively enrolled patients who were hospitalized in the Department of Respiratory Medicine at Northern Jiangsu People's Hospital between January 2016 and December 2023, and were diagnosed with chronic obstructive pulmonary disease (COPD) and comorbid small airway dysfunction (SAD) , with completed pulmonary function testing and bronchodilator reversibility testing. Inclusion criteria were: (1) age ≥ 30 years; (2) meeting established diagnostic criteria for both COPD and SAD; and (3) for patients with multiple hospitalizations, only the first hospitalization meeting the diagnostic criteria was included. Exclusion criteria included: (1) incomplete pulmonary function testing or bronchodilator reversibility testing during hospitalization (n = 41); (2) missing key clinical data (n = 48); (3) loss to follow-up or follow-up duration < 1 year after discharge (n = 34); (4) presence of other serious lung diseases (n = 115); and (5) severe cardiac, hepatic, or renal dysfunction (n = 78). All patients were followed up for one year after discharge, with the end of follow-up on December 31, 2024. A total of 1,084 patients were initially screened, and after applying the exclusion criteria, 768 patients were finally included in the analysis. Based on whether readmission occurred within 1 year after discharge, patients were divided into the readmission group (n = 342) and the non-readmission group (n = 426). The flowchart of patient screening, enrollment, and grouping is presented in Fig. 1.

Diagnostic criteria

SAD was defined by three indices: MMEF (maximal mid-expiratory flow), FEF50% (forced expiratory flow at 50% of forced vital capacity), and FEF75% (forced expiratory flow at 75% of forced vital capacity). According to a Chinese population study, SAD = at least two indices <65% predicted [29,30].

COPD was diagnosed per 2023 GOLD (Global Initiative for Chronic Obstructive Lung Disease) criteria: post-bronchodilator FEV_1 (forced expiratory volume in 1 second)/FVC (forced vital capacity) <0.70 [31].

Data Collection

All baseline data were extracted from the electronic medical record system and independently reviewed and cross-verified by two trained physicians to ensure data accuracy and consistency. Demographic and clinical variables included age, sex, body mass index (BMI), marital status (married vs. single/divorced/widowed), smoking history, alcohol consumption, length of hospital stay, and use of mechanical ventilation. Smoking history was recorded as pack-years and smoking status (never smoker, former smoker, or current smoker; former smokers were defined as those who had quit smoking for at least 1 year before enrollment). Alcohol consumption was categorized as either present or absent. Comorbidities recorded included asthma, hypertension, diabetes mellitus, and coronary artery disease. To minimize detection bias, all laboratory parameters were measured using peripheral venous blood samples collected within 24 hours after admission and before the administration of corticosteroids, antibiotics, nebulization therapy, fluid replacement, or albumin infusion, including complete blood count, C-reactive protein (CRP), serum albumin (ALB), D-dimer, and arterial blood gas parameters (PaO_2, $PaCO_2$, pH). Pulmonary

function parameters (MMEF, FEF_{50}%, FEF_{75}%) and bronchodilator reversibility testing indices (FEV_1, FVC, FEV_1/FVC ratio) were extracted from spirometry performed during hospitalization. The key outcome variable (readmission status) was ascertained through a combined search of inpatient and outpatient electronic medical records, supplemented by verification via the Yangzhou Regional Medical and Health Information Sharing Platform.

Assessment of study outcomes

The primary outcome was defined as the first unplanned hospital readmission due to acute exacerbation of COPD within 1 year after discharge. To comprehensively capture readmission events and minimize data loss resulting from patients seeking care at other institutions, readmission data were retrieved from the electronic medical record system of Northern Jiangsu People's Hospital and supplemented by the Yangzhou Regional Medical and Health Information Sharing Platform, which integrates inpatient and outpatient/emergency records from public medical institutions across the city. This dual-source strategy was intended to improve ascertainment of readmission events.

Definition and calculation of the NPAR

Neutrophil percentage-to-albumin ratio (NPAR) was calculated as: NPAR = (NE% × 100) / ALB, where NE% is the neutrophil percentage and ALB is serum albumin (g/L). Both parameters were measured from peripheral venous blood samples collected after an overnight fast within 24 hours of admission.

Statistical Analysis

All statistical analyses were performed using SPSS version 25.0 and R version 4.4.3, with a two-sided *P* value considered statistically significant.

Normality of continuous variables was assessed using the Kolmogorov–Smirnov test. Variables that followed a normal distribution were expressed as mean ± standard deviation (SD), and group comparisons were performed using independent samples t-tests; variables that did not follow a normal distribution were expressed as median and interquartile range (IQR), and compared using the Mann–Whitney U test. Categorical variables were presented as frequencies and percentages, and group comparisons were conducted using chi-square tests or Fisher's exact test, as appropriate.

To explore the associations of Eos and NPAR with 1-year readmission, three consecutive logistic regression models were constructed: Model 1 (unadjusted), Model 2 (adjusted for age, sex, and BMI), and Model 3 (further adjusted for smoking index, hypertension, diabetes, marital status, drinking status, coronary artery disease, and use of mechanical ventilation). Restricted cubic splines (RCS) were used to visualize dose-response relationships, and threshold analysis was performed to estimate an exploratory inflection point for eosinophil count (Eos). Subgroup analyses were conducted to evaluate heterogeneity across prespecified patient strata.

To control for confounding, 1:1 nearest-neighbor propensity score matching (PSM) with a caliper of 0.2 was performed based on all covariates (age, sex, BMI, smoking index, hypertension, diabetes, asthma, marital status, drinking status, coronary artery disease, and use of mechanical ventilation). A caliper of 0.2 on the propensity score scale was applied. Balance before and after matching was assessed using standardized mean differences (SMD), with |SMD| < 0.1 considered adequate; after matching, all covariates had |SMD| < 0.1. In the matched cohort, multivariable logistic regression was repeated to estimate the adjusted association between each exposure and 1-year readmission. Given that asthma may influence both eosinophil levels and readmission risk, we performed a sensitivity analysis excluding patients with asthma, followed by PSM in this subgroup using the same algorithm and covariates (except asthma). Kaplan-Meier (KM) curves and Cox proportional hazards regression were performed as sensitivity analyses to account for time-to-event data.

The discriminative ability of NPAR alone, Eos alone, and the combined model of NPAR and Eos was evaluated using receiver operating characteristic (ROC) curves and the area under the curve (AUC) on the testing set. The AUC values of the three models were compared using DeLong's testing, with a two-sided $P < 0.05$ considered statistically significant.

Of note, for Eos as a continuous variable, odds ratios are reported per 0.1×10^9/L increase (equivalent to 100 cells/μL) to improve clinical interpretability, whereas quartile-based and RCS analyses retain the original measurement unit. All statistical inferences, including *P* values and trend tests, are unaffected by this unit conversion.

Machine Learning Framework and Interpretability

To prevent data leakage, data were split into training (70%) and independent testing (30%) sets before feature selection. Training set was used for variable screening, model training, and internal validation; testing set was reserved solely for final performance evaluation, uninvolved in any screening or training.

Predictor selection was performed using least absolute shrinkage and selection operator (LASSO) regression combined with multicollinearity diagnostics. LASSO regression excluded WBC, NE, and ALB—the components used to calculate NPAR—and the optimal regularization parameter corresponding to the minimum mean squared error (MSE) was selected via 10-fold cross-validation. The lambda.min criterion was applied, and 10 variables were ultimately retained. All variables had variance inflation factor (VIF) values < 3 and were included in the final models. All predictors were standardized using Z scores (mean = 0, standard deviation = 1) with parameters estimated solely from the training set, which were then applied to the testing set to avoid data leakage.

To evaluate the incremental predictive value of Eos and NPAR, four progressive RF-based models were constructed: a basic model (including smoking index, RDW, HGB, PaO_2, diabetes, marital status, mechanical ventilation, and coronary artery disease), an +Eos model (basic model plus Eos), an +NPAR model (basic model plus NPAR), and the full model (basic model plus Eos and NPAR). All models were trained on the same training set and evaluated on the prespecified independent testing set. The DeLong test was used to compare AUC differences between models, with 95% confidence intervals calculated for each AUC increment.

Model interpretability was enhanced using SHapley Additive exPlanations (SHAP). Feature importance was quantified by the mean absolute SHAP values. Waterfall plots were employed to visualize the prediction decomposition for representative high-risk and low-risk samples, and dependence plots were generated to illustrate the impact of key variables on predicted probabilities. Additionally, decision curve analysis (DCA), along with the net reclassification improvement (NRI) and integrated discrimination improvement (IDI), was performed to evaluate the clinical net benefit of the models across different risk thresholds, thereby assessing their practical clinical utility.

Results

Basic Characteristics of the Study Population

Among the 768 enrolled patients with COPD and comorbid SAD, 342 (44.531%) were readmitted within 1 year, while 426 (55.569%) were not. Baseline characteristics between the two groups are compared in Table 1. Compared with the non-readmission group, the readmission group had significantly higher smoking index, a higher proportion of former smokers, more frequent mechanical ventilation use, and a higher prevalence of diabetes (all $P < 0.05$). Laboratory findings revealed that readmitted patients had significantly elevated levels of WBC, NE, Eos, RDW, CRP, D-dimer, and NPAR, along with significantly lower levels of L and ALB (all $P < 0.05$). No significant differences were observed between the two groups in age, sex, marital status, BMI, length of hospital stay, drinking status, GOLD grade, asthma, hypertension, coronary artery disease, PLT, HGB, PaO_2, $PaCO_2$, or pH (all $P > 0.05$).

Table 1 Characteristics of patients with COPD and comorbid SAD according to 1-year readmission

Clinical Characteristic	Non-readmission within 1 year (N = 426)	Readmission within 1 year (N = 342)	Statistic	*P*
Age (years)	71.000(65.000-76.000)	71.500(67.000-76.000)	1.433	0.152
Sex, n (%)			0.463	0.496
Male	325(76.235%)	268(78.426%)		
Female	101(23.765%)	74(21.574%)		
Marital status, n (%)			0.381	0.537
Married	395(92.706%)	313(91.545%)		
Single/divorced/widowed	31(7.294%)	29(8.454%)		
BMI (kg/m2)	23.519(21.091-26.107)	23.321(20.697-25.970)	1.333	0.183
Length of hospital stay (days)	8.000(6.000-9.000)	8.000 (6.000-10.000)	1.790	0.074
Smoking index (pack-years)	20.000 (0.000-40.000)	20.000 (0.000-45.000)	3.096	0.002
Smoking status, n (%)			6.344	0.042
Never smoker	142(33.412%)	94(27.405%)		
Former smoker	113(26.353%)	118(34.684%)		
Current smoker	171(40.235%)	130(37.901%)		
Drinking status, n (%)			0.186	0.666
Ever drinker	247(58.118%)	193(56.268%)		
Never drinker	179(41.882%)	149(43.732%)		
GOLD grade, n (%)			4.613	0.202
Grade 1 (Mild)	26(6.118%)	16(4.665%)		
Grade 2 (Moderate)	178(41.882%)	128(37.318%)		
Grade 3 (Severe)	173(40.706%)	143(41.691%)		
Grade 4 (Very severe)	49(11.294%)	55(16.326%)		
Mechanical ventilation, n (%)			6.133	0.013
No	306(71.765%)	217(63.557%)		
Yes	120(28.235%)	125(36.443%)		
Asthma, n (%)			2.735	0.098
No	328(76.941%)	280(81.924%)		
Yes	98(23.059%)	62(18.076%)		
Hypertension, n (%)			1.001	0.317
No	221(51.765%)	165(48.397%)		
Yes	205(48.235%)	177(51.603%)		
Diabetes, n (%)			4.498	0.034
No	387(90.824%)	294(86.006%)		

Yes	39(9.176%)	48(13.994%)		
Coronary artery disease, n (%)			1.693	0.193
No	385(90.353%)	299(87.464%)		
Yes	41(9.647%)	43(12.536%)		
WBC ($\times10^9$/L))	6.880 (5.358-8.683)	7.465 (5.688-9.700)	3.081	0.002
NE ($\times10^9$/L)	4.545(3.410-6.363)	5.145(3.858-7.420)	4.248	<0.001
L ($\times10^9$/L)	1.385 (0.960-1.850)	1.225 (0.910-1.673)	2.424	0.015
Eos ($\times10^9$/L)	0.130 (0.050-0.290)	0.320 (0.148-0.530)	9.190	<0.001
PLT ($\times10^9$/L)	182.000(144.000-240.000)	200.000(149.750-250.000)	1.543	0.123
RDW (%)	13.200(12.600-13.800)	14.000(13.200-15.425)	9.764	<0.001
ALB (g/L)	39.850(37.200-43.100)	36.200(32.100-40.630)	9.192	<0.001
HGB (g/L)	131(121.000-143.000)	132.000(121.000-144.250)	0.727	0.467
CRP (mg/L)	5.450(1.428-20.188)	9.030(2.210-36.373)	3.518	<0.001
DD (μg/L FEU)	0.370(0.248-0.650)	0.440(0.270-0.920)	2.900	0.004
PaO_2(mmHg)	85.700 (75.000-102.250)	87.000(72.875-110.900)	0.732	0.464
$PaCO_2$(mmHg)	43.000 (39.000-48.000)	43.000(37.000-48.000)	0.967	0.333
pH	7.420(7.390-7.450)	7.430(7.400-7.450)	0.947	0.344
NPAR	1.694(1.469-1.987)	2.006(1.697-2.365)	9.603	<0.001

Notes: Data are presented as median (interquartile range) for continuous variables and as n (%) for categorical variables.

Abbreviations: COPD, chronic obstructive pulmonary disease; SAD, small airway dysfunction; BMI, body mass index; WBC, white blood cell count; NE, neutrophil count; L, lymphocyte count; Eos, eosinophil count; PLT, Platelets; RDW, red cell distribution width; ALB, albumin; HGB, hemoglobin; CRP, C-reactive protein; DD, D-dimer; NPAR, neutrophil percentage-to-albumin ratio.

Multivariable logistic regression analysis of Eos and NPAR associated with 1-year readmission

To evaluate the associations of Eos and NPAR with 1-year readmission among COPD patients with SAD, each indicator was entered separately into three logistic regression models as both a continuous variable and a quartile-based categorical variable, as shown in the Table 2.

In the unadjusted analysis (Model 1), both Eos and NPAR were positively associated with 1-year readmission. These associations remained statistically significant after adjustment for multiple confounding factors. In the fully adjusted model (Model 3), each 0.1×10⁹/L increase in Eos was associated with a 26.9% increase in the odds of 1-year readmission (OR = 1.269, 95% CI: 1.194–1.353, $P < 0.001$). For each one-unit increase in NPAR, the odds of readmission increased by 528.3% (OR = 6.283, 95% CI: 4.238-9.491, $P < 0.001$). Furthermore, the quartile-based categorical analysis yielded results consistent with the continuous variable analysis. Compared with the lowest quartile (Q1), the highest quartile (Q4) of Eos was significantly associated with increased odds of readmission (OR = 6.223, 95% CI: 3.953–9.949); similarly, the Q4 group of NPAR also showed

a positive association with readmission (OR = 7.281, 95% CI: 4.573-11.790) (all *P* < 0.05). E-value analysis further indicated that an unmeasured confounder would need to be associated with both the exposure and 1-year readmission by a risk ratio of at least 4.43 (Eos) and 4.84 (NPAR) to explain away the observed associations (Supplementary Table S1).

Table 2 Multivariable logistic regression analysis of Eos and NPAR associated with 1-year readmission

Variable	Model 1		Model 2		Model 3	
	OR (95% CI)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*
Eos	1.273(1.200-1.356)	<0.001	1.275(1.201-1.358)	<0.001	1.269(1.194-1.353)	<0.001
Q1	Reference		Reference		Reference	
Q2	1.610(1.035-2.520)	0.036	1.635(1.046-2.570)	0.032	1.729(1.100-2.738)	0.019
Q3	3.966(2.569-6.196)	<0.001	4.037(2.607-6.330)	<0.001	4.034(2.590-6.362)	<0.001
Q4	6.244(4.025-9.832)	<0.001	6.319(4.058-9.990)	<0.001	6.223(3.953-9.949)	<0.001
P	<0.001		<0.001		<0.001	
NPAR	6.307(4.319-9.378)	<0.001	6.182(4.215-9.231)	<0.001	6.283(4.238-9.491)	<0.001
Q1	Reference		Reference		Reference	
Q2	1.750(1.120-2.753)	0.015	1.720(1.098-2.710)	0.019	1.771(1.120-2.820)	0.048
Q3	3.487(2.265-5.433)	<0.001	3.397(2.194-5.320)	<0.001	3.480(2.219-5.530)	<0.001
Q4	7.340(4.707-11.631)	<0.001	7.137(4.552-11.370)	<0.001	7.281(4.573-11.790)	<0.001
P	<0.001		<0.001		<0.001	

Notes: Model 1: Unadjusted for any covariates. Model 2: Adjusted for age, sex, and BMI as confounding variables. Model 3: Further adjusted for smoking index, hypertension, diabetes, marital status, drinking status, coronary artery disease, and use of mechanical ventilation on the basis of Model 2. For continuous Eos, OR is presented per 0.1×10⁹/L increase. For continuous NPAR, OR is presented per 1-unit increase. Quartiles are referenced to Q1.

Abbreviations: Eos, eosinophil count; NPAR, neutrophil percentage-to-albumin ratio; OR, odds ratio; CI, confidence interval; P, P-value.

Dose-response relationship and threshold effect analysis of Eos and NPAR on 1-year readmission

Restricted cubic spline (RCS) regression was employed to evaluate the dose–response relationship between each indicator and the risk of 1-year readmission (Fig. 2). In the unadjusted model (Figure 2A, 2B), Eos exhibited a significant non-linear dose–response association with readmission risk (overall association test: *P* < 0.001; non-linearity test: *P* < 0.05), whereas NPAR did not demonstrate a significant non-linear relationship (overall association test: *P* < 0.001; non-linearity test: *P* > 0.05), suggesting a predominantly linear association with the outcome. After full adjustment for potential confounding factors using Model 3 (Figure 2C, 2D), the non-linear association for Eos remained statistically significant (overall association test: *P* < 0.001; non-linearity test: *P* < 0.05), while NPAR continued to show a linear dose–response pattern without evidence of non-linearity (overall association test: *P* < 0.001; non-linearity test: *P* > 0.05). These findings collectively indicate that the association between Eos and readmission risk follows a non-linear pattern, whereas the effect of NPAR is consistent with a linear trend, irrespective of covariate adjustment.

Given the observed non-linear association, threshold effect analysis was performed based on Model 3 to estimate an exploratory inflection point for Eos. As shown in Table 3, the estimated inflection point was 0.25 × 10⁹/L. Below this value, each

0.1×10⁹/L increase in Eos was associated with an OR of 1.888 (95% CI: 1.549–2.311, P < 0.001); above this value, the OR was 1.097 (95% CI: 1.009–1.199, P = 0.034). The log-likelihood ratio test indicated better fit for the segmented than the linear model (P < 0.001).

Table 3 Threshold effect analysis of Eos on 1-year readmission risk.

Variable	Inflection point	Effect segment	Adjusted OR (95% CI)	*P* value
Eos (per 0.1×10⁹/L)	0.25			
		< 0.25	1.888(1.549-2.311)	<0.001
		≥ 0.25	1.097(1.009-1.199)	0.034
P for non-linearity				<0.001

Notes: Adjusted for variables in Model 3, including age, sex, BMI, smoking index, hypertension, diabetes, marital status, drinking status, coronary artery disease and use of mechanical ventilation. Eos was analyzed per 0.1 × 10⁹/L. *P* for non-linearity was derived from log-likelihood ratio test comparing the one-piece and two-piece models.

Abbreviations: Eos, eosinophil count; OR, odds ratio; CI, confidence interval.

Subgroup and Interaction Analyses

Subgroup analyses were performed to examine potential differences in the association between exposures and outcomes across various patient strata, and multiplicative interaction terms were used to assess effect modification. The results are presented as forest plots (Fig. 3).

Overall, the associations of Eos and NPAR with 1-year readmission risk were generally consistent across clinical subgroups (including age, sex, BMI, hypertension, coronary artery disease, smoking, and drinking status), with all interaction *P* values exceeding 0.05, indicating no significant effect modification. Notably, the interaction between Eos and asthma status was borderline (*P* for interaction = 0.050, Figure 3A), and that between NPAR and asthma status was also close to borderline (*P* for interaction = 0.054, Figure 3B). Given the limited sample size of the asthma subgroup and the exploratory nature of subgroup analyses, these findings warrant cautious interpretation and require validation in independent asthma cohorts.

Sensitivity Analysis

The propensity-score-matched sensitivity analysis is presented in Table 4. In the matched cohort, the associations of Eos and NPAR with 1-year readmission remained directionally consistent with the primary analysis. Given that asthma may influence both eosinophil levels and readmission risk, potentially introducing residual confounding, we further performed a sensitivity analysis excluding patients with asthma, as shown in Supplementary Table S4; the results were consistent with the primary analysis, suggesting that these findings were not driven by comorbid asthma. The balance of baseline characteristics after matching is reported in Supplementary Table S2, and the balance of baseline characteristics after matching in the asthma-excluded subgroup is reported in Supplementary Table S3; in both tables, all covariates had absolute standardized mean differences (SMD) less than 0.1, indicating good comparability between the two groups.

Furthermore, Cox proportional hazards regression (Supplementary Table S5) and KM curves (Supplementary Figure S1) were performed as sensitivity analyses to account for time-to-event data. In the fully adjusted model, Eos (per 0.1×10⁹/L increase: HR = 1.088, 95% CI: 1.046–1.130, *P* < 0.001) and NPAR (per unit increase: HR = 1.390, 95% CI: 1.107–1.750, *P* = 0.005) remained significantly associated with readmission risk, with consistent dose-response trends across quartiles (*P* for trend < 0.05). The KM curves further demonstrated a positive dose-response trend (Log-rank *P* < 0.05). Collectively, these sensitivity analyses support the robustness of our primary findings.

Table 4 Sensitivity Analysis: OR (95% CI) in the Propensity Score-Matched Cohort

Variable		OR (95% CI)	*P*
Eos			
	Continuous	1.271(1.191-1.362)	<0.001
	Q1	reference	
	Q2	1.901(1.205-3.015)	0.006
	Q3	4.349(2.758-6.943)	<0.001
	Q4	5.806(3.658-9.354)	<0.001
	P for trend		<0.001
NPAR			
	Continuous	5.761(3.839-8.826)	<0.001
	Q1		
	Q2	1.587(1.004-2.520)	0.049
	Q3	3.442(2.182-5.489)	<0.001
	Q4	6.391(3.984-10.425)	<0.001
	P for trend		<0.001

Note: Propensity scores were calculated adjusting for all covariates (age, sex, BMI, smoking index, hypertension, diabetes, marital status, drinking status, coronary artery disease, and use of mechanical ventilation). Doubly robust models combine propensity score weighting with covariate adjustment. For continuous Eos, OR is presented per 0.1×10^9/L increase. For continuous NPAR, OR is presented per 1-unit increase. Quartiles are referenced to Q1.

Abbreviations: Eos, eosinophil count; NPAR, neutrophil percentage to albumin ratio; OR, odds ratio; CI, confidence interval; P, P-value.

Predictive Value of Eos, NPAR, and the Combined Model for 1-year readmission

ROC curve analysis (Fig. 4) revealed that NPAR and Eos exhibited comparable discriminative ability for predicting 1-year readmission risk, with AUCs of 0.707 (95% CI: 0.670–0.744) and 0.693 (95% CI: 0.655–0.730), respectively. Notably, combining these two parameters significantly improved the AUC to 0.776 (95% CI: 0.743–0.808), and the DeLong test confirmed that the combined model was significantly superior to either single indicator alone (both $P < 0.001$). These findings suggest that the combined assessment of NPAR and Eos may facilitate more precise risk stratification for 1-year readmission in this population.

Establishment and validation of the machine learning prediction model

Patients were randomly assigned to the training group (n = 537) and the testing group (n = 231) at a 7:3 ratio. As shown in Supplementary Table S6, baseline characteristics were generally comparable between the two groups, with a statistically significant difference observed only for use of mechanical ventilation (P = 0.038).

Predictor selection was performed using Least Absolute Shrinkage and Selection Operator (LASSO) regression combined with collinearity diagnostics on the training set, with the results shown in Fig. 5. In the LASSO procedure, 10-fold cross-validation was used to determine the optimal penalty parameter, and the λ value that minimized the cross-validated error (λmin =

0.01657194) was selected to balance predictive performance and clinical interpretability. This approach identified 10 robust predictors with nonzero coefficients from the training cohort (Figure 5A, 5B). Subsequently, variance inflation factor (VIF) analysis was performed to assess multicollinearity among these predictors. The VIF values ranged from 1.014 to 2.869, all well below the conventional threshold of 5, indicating no significant multicollinearity. Consequently, all 10 variables were incorporated into the subsequent prediction model development (Supplementary Table S7).

Based on these 10 predictors, five candidate machine learning models—Decision Tree (DT), Random Forest (RF), eXtreme Gradient Boosting (XGBoost), Support Vector Machine (SVM), and K-Nearest Neighbors (KNN)—were constructed. The performance metrics of these models in the training set are summarized in Table 5, and the corresponding ROC curves for the testing set are displayed in Fig. 6. In the training set, the RF model achieved the highest discriminative ability (AUC = 0.801, 95% CI: 0.749–0.852; AUPRC = 0.773, 95% CI: 0.696–0.840), with a sensitivity of 0.653 (95% CI: 0.567–0.742) and a specificity of 0.792 (95% CI: 0.720–0.859), indicating the best balance between sensitivity and specificity. In the testing set, the RF model maintained superior discriminative performance (AUC = 0.801, 95% CI: 0.745–0.856), closely consistent with the training-set results, suggesting no significant overfitting and satisfactory generalizability. Taking into account model fit, generalizability, and clinical utility, the RF model was ultimately selected as the core model for subsequent incremental value assessment and SHAP-based interpretability analysis.

Table 5 Performance comparison of five machine learning models in the training set

Model	AUC (95% CI)	AUPRC (95% CI)	F1-score (95% CI)	sensitivity (95% CI)	specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
DT	0.748 (0.685–0.818)	0.732 (0.633–0.804)	0.649 (0.566–0.724)	0.665 (0.562–0.759)	0.694 (0.622–0.786)	0.636 (0.543–0.727)	0.721 (0.645–0.800)
RF	0.801 (0.749–0.852)	0.773 (0.696–0.840)	0.682 (0.610–0.751)	0.653 (0.567–0.742)	0.792 (0.720–0.859)	0.716 (0.631–0.804)	0.739 (0.662–0.814)
XGBoost	0.784 (0.735–0.840)	0.758 (0.672–0.826)	0.660 (0.590–0.724)	0.668 (0.585–0.763)	0.716 (0.636–0.788)	0.654 (0.564–0.735)	0.728 (0.661–0.808)
SVM	0.795 (0.744–0.847)	0.777 (0.705–0.837)	0.641 (0.555–0.713)	0.607 (0.509–0.707)	0.773 (0.692–0.833)	0.682 (0.582–0.770)	0.710 (0.638–0.777)
KNN	0.737 (0.676–0.796)	0.717 (0.639–0.787)	0.617 (0.541–0.697)	0.624 (0.532–0.726)	0.684 (0.613–0.754)	0.613 (0.518–0.696)	0.693 (0.615–0.784)

Notes: RF model demonstrated the best overall performance among the five models.

Abbreviations: AUC, area under the receiver operating characteristic curve; AUPRC, area under the precision-recall curve; CI, confidence interval; F1-score, harmonic mean of precision and recall; PPV, positive predictive value; NPV, negative predictive value; DT, Decision Tree; RF, random forest; XGBoost, extreme gradient boosting; SVM, support vector machine; KNN, k-nearest neighbors.

To further quantify the independent and combined incremental value of Eos and NPAR, we constructed four stepwise models based on the RF algorithm and evaluated their predictive performance on the testing set (Table 6). The baseline model, which included smoking index, RDW, HGB, PaO_2, diabetes, marital status, mechanical ventilation, and coronary artery disease, achieved an AUC of 0.701 (95% CI: 0.629–0.766), with a sensitivity of 0.553 (95% CI: 0.460–0.652) and a specificity of 0.742 (95% CI: 0.667–0.816). When NPAR was added alone, the AUC increased to 0.742 (95% CI: 0.676–0.806), with a ΔAUC of 0.041 (95% CI: −0.011 to 0.093, P = 0.134), which did not reach statistical significance; sensitivity improved to 0.583 (95% CI: 0.490–0.677), while specificity slightly decreased to 0.734 (95% CI: 0.658–0.813). Adding Eos alone significantly raised the AUC to 0.772 (95% CI: 0.711–0.829; ΔAUC = 0.066, 95% CI: 0.012–0.123, P = 0.016), accompanied by an increase in sensitivity to 0.631 (95% CI: 0.545–0.725) and a decrease in specificity to 0.711 (95% CI: 0.633–0.791). The full model combining Eos and NPAR

achieved the highest AUC of 0.802 (95% CI: 0.744–0.853; ΔAUC = 0.101, 95% CI: 0.042–0.161, *P* = 0.001), with an AUPRC of 0.774 (95% CI: 0.696–0.841), a sensitivity of 0.621 (95% CI: 0.528–0.716), and a specificity of 0.766 (95% CI: 0.687–0.843). Notably, the full model not only achieved the highest AUC but also restored specificity to a level surpassing that of the baseline model (0.766 vs. 0.742), while maintaining a relatively high sensitivity. These findings indicate that the combination of Eos and NPAR provides superior overall predictive performance compared with either indicator alone.

Table 6 Predictive performance of progressive machine learning models in the testing set

Model	Sensitivity (95% CI)	Specificity (95% CI)	AUPRC (95% CI)	AUC (95% CI)	AUC difference vs Base (95% CI)	*P* value
Basic	0.553 (0.460–0.652)	0.742 (0.667–0.816)	0.705 (0.616–0.780)	0.701 (0.629–0.766)	Ref	Ref
+NPAR	0.583 (0.490–0.677)	0.734 (0.658–0.813)	0.736 (0.653–0.810)	0.742 (0.676–0.806)	0.041 (-0.011–0.093)	0.134
+Eos	0.631 (0.545–0.725)	0.711 (0.633–0.791)	0.757 (0.675–0.830)	0.772 (0.711–0.829)	0.066 (0.012–0.123)	0.016
Full	0.621 (0.528–0.716)	0.766 (0.687–0.843)	0.774 (0.696–0.841)	0.802 (0.744–0.853)	0.101 (0.042–0.161)	0.001

Notes: All models were based on the RF algorithm and evaluated on the testing set. The Basic model included smoking index, RDW, HGB, PaO_2, diabetes, marital status, mechanical ventilation, and coronary artery disease. The +NPAR model added NPAR; the +Eos model added Eos; the Full model included both. AUC differences (ΔAUC) and their 95% confidence intervals were estimated using 500 bootstrap resamples. *P* values were derived from DeLong's test comparing the AUC of each model to that of the Basic model.

Abbreviations: AUC, area under the receiver operating characteristic curve; AUPRC, area under the precision-recall curve; CI, confidence interval; Eos, eosinophil count; NPAR, neutrophil percentage-to-albumin ratio; Ref, reference category (Basic model); ΔAUC, difference in AUC.

To further evaluate the incremental clinical value of Eos and NPAR, we assessed improvements in individualized risk stratification using the categorical net reclassification improvement (NRI) and integrated discrimination improvement (IDI) (Table 7). Adding Eos alone was associated with significant improvements in reclassification and discrimination, with a categorical NRI of 0.204 (95% CI, 0.058–0.351; P = 0.006) and an IDI of 0.062 (95% CI, 0.019–0.105; P = 0.006). Adding NPAR alone was associated with a significant improvement in discrimination (IDI, 0.057; 95% CI, 0.016–0.099; P = 0.007), but not reclassification (NRI, 0.119; 95% CI, −0.025 to 0.254; P = 0.092). The full model combining Eos and NPAR showed the greatest improvements in both reclassification and discrimination, with an NRI of 0.255 (95% CI, 0.102–0.404; P = 0.001) and an IDI of 0.104 (95% CI, 0.052–0.155; P < 0.001). These findings, supported by complementary metrics, suggest that the combined use of Eos and NPAR may improve individualized clinical risk stratification. However, because these analyses were performed in an internal testing set, they should be considered exploratory until externally validated.

Table 7 NRI and IDI for incremental risk stratification by adding Eos and NPAR

Model comparison	NRI (95% CI)	*P* value	IDI (95% CI)	*P* value
Basic	Ref		Ref	
+Eos	0.204 (0.058–0.351)	0.006	0.062 (0.019–0.105)	0.006
+NPAR	0.119 (-0.025–0.254)	0.094	0.057 (0.016–0.099)	0.007
Full	0.255 (0.102–0.404)	0.001	0.104 (0.052–0.155)	< 0.001

Notes: All models were based on the RF algorithm and evaluated in the testing set. The Basic model included smoking index, RDW, HGB, PaO_2, diabetes, marital status, mechanical ventilation, and coronary artery disease. The +Eos model added Eos; the +NPAR model added NPAR; the Full model included both. 95% CIs and *P* values were estimated via 1000 bootstrap resamples; *P* values correspond to two-sided tests for NRI/IDI > 0 versus the Basic model.

Abbreviations: NRI, net reclassification improvement; IDI, integrated discrimination improvement; CI, confidence interval; Ref, reference; Eos, eosinophil count; NPAR, neutrophil percentage-to-albumin ratio.

The calibration and clinical utility of the Full model were further evaluated using calibration curves and decision curve analysis (DCA) in the testing set (Fig. 7). As shown in Figure 7A, the Full model achieved the lowest Brier score (0.1788, 95% CI: 0.1553–0.2039) among all models, with its calibration curve closest to the ideal diagonal line and demonstrating no obvious calibration bias across the entire risk spectrum. In contrast, the Basic model and the single-indicator models (Eos and NPAR) showed varying degrees of deviation from the ideal line at both low- and high-risk regions. DCA (Figure 7B) further revealed that the Full model provided the highest net benefit across the clinically reasonable threshold range compared with the Basic model and the single-indicator models. These findings suggest that the combination of Eos and NPAR offers improved calibration and superior clinical decision utility for predicting 1-year readmission risk.

SHAP analysis was performed to interpret the predictions of the optimized random forest (RF) full model (Fig. 8). Feature importance ranking (Figure 8A) identified Eos and NPAR as the most influential predictors, and the summary plot (Figure 8B) further showed that elevated levels of both were associated with an increased 1-year readmission risk. Waterfall plots (Figure 8C, 8D) decomposed the predictions for representative high- and low-risk samples. In the high-risk sample, RDW, NPAR, and Eos positively contributed to the prediction, driving the predicted probability from a baseline of 0.379 to 0.979; in the low-risk sample, these variables exerted negative contributions, lowering the predicted probability to 0.019. The dependence plots (Supplementary Figure S2) further revealed positive associations for both Eos and NPAR. Collectively, these findings support the core predictive role of Eos and NPAR in the model.

Discussion

This retrospective cohort study evaluated the associations of eosinophil count (Eos) and neutrophil percentage-to-albumin ratio (NPAR) with 1-year readmission in patients with COPD and comorbid SAD and developed an internally validated machine-learning prediction model. Elevated Eos and NPAR were associated with a higher risk of readmission. Eos showed evidence of a non-linear association, with an estimated inflection point at 0.25×10^9/L, whereas NPAR showed a significant overall association without evidence of non-linearity. The random forest (RF) model achieved the highest AUC in the internal testing set, and the full model including Eos and NPAR showed higher discrimination than the basic model. Overall, these findings suggest that these routine inflammatory-nutritional biomarkers may help improve post-discharge risk stratification in this high-risk phenotype; however, further external validation is required before clinical implementation.

Our study found that elevated Eos was independently associated with an increased 1-year readmission risk in patients with COPD and comorbid SAD, and that it showed a statistically significant dose-response relationship. After adjustment for multiple confounders, this association remained robust and was consistent across multivariable logistic regression and Cox proportional hazards models; however, it still requires external and prospective validation. Although some clinical contexts have reported potential protective effects of Eos [24-26], our findings are consistent with most literature supporting Eos as a predictor of poor prognosis [22,23] and support its potential value in assessing type 2 inflammation and risk stratification in

COPD. This study extends the prognostic significance of Eos from its previous focus mainly on predicting acute exacerbations in the general COPD population to a clinically distinct subgroup with unique pathophysiological characteristics—COPD with comorbid SAD—and preliminarily quantifies its dose-response relationship and threshold effect in predicting readmission risk. Notably, the Eos cutoff value preliminarily identified in this study for predicting readmission risk was 0.25×10⁹/L, which is lower than the threshold recommended by the Global Initiative for Chronic Obstructive Lung Disease (GOLD) guidelines for initiating inhaled corticosteroid (ICS) therapy (300 cells/μL) [31]. This difference may reflect greater biological sensitivity of patients with SAD to eosinophil-mediated airway inflammation. Given that SAD is an early core pathological change in COPD [14], its characteristic small-airway eosinophilic infiltration may precede large-airway involvement; therefore, a lower Eos threshold may be needed to predict readmission risk. The SHAP dependence plot further suggests a non-linear contribution pattern of Eos to model predictions—at low Eos levels, SHAP values remained relatively stable, indicating a limited impact on readmission risk within this range; however, once Eos exceeded a certain level, SHAP values began to rise steadily, and the rate of increase accelerated as Eos levels rose, suggesting that Eos may be positively associated with readmission risk, with this association becoming more pronounced at higher levels. This trend was generally consistent with the results of our multivariable logistic regression and restricted cubic spline (RCS) analyses. These findings have potential clinical implications: in patients with COPD and comorbid SAD, readmission risk may already be elevated even when peripheral blood Eos levels have not reached the conventional ICS treatment threshold (300 cells/μL). In clinical practice, routine monitoring of Eos might be considered in such patients, and further exploration could focus on refining risk stratification and individualized intervention decisions based on this marker; however, its clinical utility still needs to be confirmed by prospective and external validation studies.

The predictive value of eosinophil count for readmission risk can be attributed to its role as a biomarker of type 2 inflammation, which drives airway remodeling and mucus hypersecretion. Specifically, previous studies have confirmed that elevated peripheral blood eosinophil counts can serve as an easily accessible surrogate marker of small airway eosinophilic inflammation [32]. Activated eosinophils release cytotoxic granules that directly damage small airway epithelium, secrete cysteinyl leukotrienes that induce smooth muscle contraction, and promote structural injury via pro-fibrotic factors [32,33]. In the context of small airway disease (SAD), characterized by luminal narrowing, thin walls, and lack of cartilaginous support, the small airways represent the earliest and most vulnerable site in COPD [12]. Persistent eosinophilic inflammation leads to irreversible wall fibrosis, smooth muscle hypertrophy, and luminal narrowing [32,34], forming the structural basis for repeated hospitalizations and increasing susceptibility to exacerbations—the most direct cause of readmission. By integrating both "structural damage" and "acute event risk" pathways, Eos exerts its prognostic value specifically within the SAD phenotype. Furthermore, this study employed continuous Eos measurements combined with restricted cubic spline (RCS) analysis, enabling depiction of the full dose-response relationship rather than relying on dichotomous methods, thereby potentially allowing more accurate identification of risk thresholds and providing a preliminary reference for clinical decision-making.

This study found that NPAR was associated with 1-year readmission risk in patients with COPD and comorbid SAD, and, to our knowledge, this is among the earlier studies to explore its prognostic value in this specific population. Although NPAR showed a significant association with readmission risk, no clear dose-response trend was observed, suggesting that risk increased approximately linearly with rising NPAR levels. SHAP dependence plots further showed that the contribution of NPAR to model predictions increased monotonically—SHAP values continuously rose with higher NPAR, without evident plateaus or inflection points. This pattern was consistent with the RCS analysis results, suggesting that in patients with COPD and comorbid SAD, elevated NPAR may be consistently and positively associated with increased readmission risk, rather than showing a sudden risk surge only after crossing a certain threshold. The difference in risk contribution patterns between NPAR and eosinophils (Eos) may stem from their distinct biological characteristics: Eos, as a single-cell count, may have a physiological upper limit, whereas NPAR, as a ratio-based metric, can increase with worsening inflammation and malnutrition, possibly exhibiting a more continuous, incremental risk effect. This suggests that NPAR might be evaluated as a continuous variable in clinical practice, although this still requires validation. Previous studies have also supported this finding [27,28]. In this study, NPAR showed an approximately linear positive association with 1-year readmission risk; its relatively large effect size (OR = 6.28) may be partly because NPAR simultaneously reflects systemic inflammation and nutritional deterioration, but the specific mechanisms still require further validation. In univariate analysis, each 1 g/L decrease in albumin was associated with a 16% increase in risk (OR = 0.862), suggesting that hypoalbuminemia itself may be an important risk factor. NPAR integrates two risk signals—elevated

inflammation and nutritional depletion—and its effect size may reflect the combined contribution of these two pathways rather than simple amplification of either component. This study may extend the prognostic utility of NPAR from mortality to readmission outcomes and provide preliminary evidence for its approximately linear relationship in patients with COPD and comorbid SAD; however, these findings require external and prospective validation, and residual confounding cannot be excluded.

The predictive value of NPAR for readmission risk can be explained by its composite nature, which integrates two distinct physiological axes: neutrophil-driven systemic inflammation and albumin-mediated nutritional reserve. Small airways are among the earliest sites affected in COPD and may respond more sensitively to systemic inflammation and nutritional imbalance. In the context of SAD, elevated NPAR may increase readmission risk through several mechanisms. First, neutrophils are key effector cells of the innate immune system, whose numbers and percentages rise significantly during COPD exacerbations; neutrophil-mediated inflammation can worsen damage and remodeling of small airway walls, impair airway clearance, and increase susceptibility to infection [35]. Second, serum albumin is a critical marker of nutritional status and systemic inflammation; hypoalbuminemia not only reflects malnutrition but also correlates with persistent systemic inflammation—wherein inflammation suppresses albumin synthesis, while low albumin levels further amplify inflammatory responses, creating a complex bidirectional interaction. Additionally, hypoalbuminemia reduces plasma oncotic pressure and drug-binding capacity, affecting effective circulating blood volume and pharmacokinetics, leading to suboptimal therapeutic response [36]. Third, these factors may synergistically enhance oxidative stress and promote protease-antiprotease imbalance, accelerating destruction of lung tissue architecture [37]. Together, these mechanisms support the theoretical rationale for NPAR as a multidimensional prognostic biomarker for readmission risk in COPD patients with SAD.

This study found that combining Eos with NPAR significantly improved predictive performance, indicating complementary value in forecasting 1-year readmission risk in COPD patients with SAD. This complementarity reflects their biological distinctiveness: Eos primarily identifies patients with an "eosinophilic phenotype" driven by type 2 airway inflammation, associated with enhanced corticosteroid sensitivity, mucus hypersecretion, and bronchial hyperresponsiveness [20,32,33]; in contrast, NPAR broadly captures the "catabolic phenotype" resulting from systemic inflammation and nutritional imbalance, processes linked to overall frailty, impaired immune recovery, and increased susceptibility to infection [27,35]. A single biomarker cannot fully cover all risk subgroups, whereas their combination characterizes the pathophysiology of COPD with SAD across two dimensions—"local airway inflammation" and "systemic host status"—each capturing different aspects of disease progression. The additive predictive value of the two markers suggests that they identify distinct yet partially overlapping susceptible subgroups: some patients are readmitted due to eosinophil-mediated airway instability, others due to inflammation-driven nutritional depletion, and the highest-risk patients exhibit both conditions. This complementarity supports a dual-biomarker testing strategy over reliance on a single laboratory measure, enabling more comprehensive risk assessment.

Subgroup analyses revealed consistent associations between Eos and NPAR with 1-year readmission risk across clinical subgroups, without significant effect modification. However, interactions between both markers and asthma status were at borderline levels, suggesting potentially stronger predictive trends in COPD-SAD patients with comorbid asthma. From a pathophysiological perspective, asthma and COPD share partial overlap in eosinophil-driven type 2 inflammatory pathways, although asthmatic patients typically exhibit more pronounced eosinophilic infiltration [38,39]. In patients with COPD and comorbid asthma, airway eosinophilic inflammation is often higher than in those with COPD alone; thus, Eos as a type 2 inflammation biomarker may offer greater dynamic range and stronger risk discrimination in this subgroup. Moreover, COPD patients with comorbid asthma inherently face higher risks of exacerbations and poorer clinical outcomes, and elevated Eos levels may further compound these risks [40,41]. Regarding NPAR, although asthma is traditionally considered a type 2 inflammation-driven disease, it frequently involves systemic inflammation and metabolic disturbances; NPAR integrates systemic inflammation and nutritional status, potentially capturing richer risk information in patients with asthma combined with systemic inflammation or malnutrition. Additionally, COPD-asthma overlap patients tend to have more complex disease courses, higher hospitalization rates, and heavier systemic inflammatory burdens, suggesting that NPAR may possess enhanced predictive potential in this subgroup [41-43]. Nevertheless, given the exploratory nature of subgroup analyses and the limited sample size of the asthma subgroup, these interactions have not yet reached confirmatory evidence standards, limiting

the reliability of conclusions. These findings provide important hypotheses for future research—validation in independent, adequately powered cohorts of asthma-COPD overlap populations could enable more refined risk stratification and support precision-based decision-making for this high-risk group.

This study employed five machine learning algorithms to build prediction models, with random forest (RF) demonstrating the best performance. As an ensemble learning method, RF constructs multiple decision trees and aggregates their predictions, effectively handling nonlinear interactions among variables and offering significant advantages in high-dimensional medical data modeling. Through a stepwise model-building approach, we systematically evaluated the incremental predictive value of Eos and NPAR—when combined, these markers achieved optimal model performance beyond conventional clinical variables. SHAP analysis identified Eos and NPAR as the most important predictors, consistent with multivariable logistic regression results, thereby strengthening the credibility of our findings. SHAP values further decompose individual patient predictions into specific feature contributions, clearly revealing the drivers behind high-risk predictions at the individual level, providing clinicians with actionable insights beyond population-level averages and bridging the gap between "population statistics" and "individual clinical decision-making."

Compared to previous studies, this research both confirms existing evidence and provides new insights across multiple levels. Regarding Eos, although some studies have questioned its predictive value in COPD [20,21,24-26], our findings are consistent with a positive association between Eos and readmission risk, and, in an early exploratory analysis, we identified a lower threshold (0.25×10^9/L) within the SAD subgroup, suggesting that the airways in this population may be more sensitive to eosinophilic inflammation. Regarding NPAR, this study may extend its predictive utility from mortality, which has been emphasized in previous literature, to readmission outcomes, providing preliminary evidence for its independent role in patients with SAD and revealing an approximately linear trend distinct from that observed in critically ill patients, which may reflect differences in patient populations and outcome measures. The main contribution of this study lies in the combined assessment of Eos, representing type 2 airway inflammation, and NPAR, representing systemic inflammation-nutritional status, suggesting that they may have complementary and incremental predictive value. By integrating an interpretable machine learning model, this exploratory analysis may provide clinicians with a more precise and understandable decision-support tool, helping to fill a gap in the literature regarding joint biomarker prediction strategies for the specific COPD-SAD subgroup.

These exploratory findings may have potential clinical implications. First, because Eos and NPAR are routine laboratory tests that are low-cost and widely accessible, if their utility is confirmed, they may be suitable for implementation across different healthcare settings. Clinicians might use their levels to rapidly stratify readmission risk before discharge: for patients with Eos > 0.25×10^9/L, closer monitoring may be needed for risks driven primarily by airway eosinophilic inflammation, and optimization of anti-inflammatory treatments such as ICS could be considered; for those with significantly elevated NPAR, benefits may be gained from controlling systemic inflammation and improving nutritional status, which may help facilitate timely adjustments to post-discharge management plans. Second, Eos and NPAR reflect two distinct dimensions—airway eosinophilic inflammation and systemic inflammation-nutritional imbalance—and their combination may provide complementary information, helping to assess risk more comprehensively. Categorizing patients based on combinations of Eos and NPAR levels (e.g., high Eos/high NPAR vs. only one elevated vs. neither elevated) may help define different risk strata corresponding to different follow-up frequencies and intervention intensities, but this strategy still requires prospective evaluation. Finally, machine learning models could potentially be integrated into electronic health record systems to enable automated risk scoring, support clinical decision-making, and improve efficiency.

This study has several notable strengths. First, in variable selection, we comprehensively considered statistical significance, predictive contribution, and feature independence, effectively reducing overfitting and mitigating the impact of multicollinearity on model stability. Second, we integrated traditional regression models with machine learning approaches, supplemented by RCS analysis and nested model comparisons, establishing a systematic and rigorous analytical framework spanning from association exploration to prediction optimization and clinical decision support. Third, we identified a risk threshold for Eos at 0.25×10^9/L, providing clinicians with a clear and actionable cutoff for risk stratification, aiding in early identification of high-risk individuals and guiding personalized intervention strategies.

This study also has several limitations. First, the single-center retrospective design may introduce selection bias and information bias, and the generalizability of findings requires validation through multi-center prospective studies. Second, Eos and NPAR were measured only once using blood samples after admission, without dynamic monitoring of changes over time; longitudinal trends in inflammatory markers might offer additional prognostic information. Third, the sample size in some subgroups (e.g., asthma group) was limited, leading to imprecise effect estimates and interaction *P* values at borderline levels, which require validation in larger cohorts. Fourth, due to data unavailability, certain potential confounders (such as medication adherence, socioeconomic status, and family support) were not considered. Finally, the performance of machine learning models needs further validation in independent external datasets.

Future research could advance in several directions. Multicenter prospective cohort studies are warranted to validate the predictive value and thresholds of Eos and NPAR. Dynamic monitoring of changes in Eos and NPAR over time could explore their trajectories and associations with readmission risk. Incorporating Eos and NPAR into more comprehensive multimodal prediction models that integrate clinical, laboratory, imaging, and genomic data may further improve predictive performance. Additionally, randomized controlled trials are needed to evaluate whether risk stratification strategies based on Eos and NPAR can effectively reduce readmission rates and improve long-term outcomes in this high-risk population.

Conclusion

In summary, elevated Eos and NPAR were independently associated with an increased risk of 1-year readmission in patients with COPD and comorbid SAD, and their combination was associated with improved predictive performance in the internal testing set. The machine learning model incorporating Eos and NPAR showed promising discriminative ability, suggesting that these readily available biomarkers may aid individualized risk stratification. However, prospective studies are warranted to further validate their clinical utility and to explore whether biomarker-guided intervention strategies can improve patient outcomes.

Abbreviations

ALB, Albumin; AUC, Area under the curve; BMI, Body mass index; CI, Confidence interval; COPD, Chronic obstructive pulmonary disease; CRP, C-reactive protein; DCA, Decision curve analysis; DD, D-dimer; Eos, Eosinophil count; FEF50%, Forced expiratory flow at 50% of forced vital capacity; FEF75%, Forced expiratory flow at 75% of forced vital capacity; FEV_1, Forced expiratory volume in 1 second; FVC, Forced vital capacity; HGB, Hemoglobin; IQR, Interquartile range; L, Lymphocyte count; LASSO, Least Absolute Shrinkage and Selection Operator; MMEF, Maximal mid-expiratory flow; NE, Neutrophil count; NPAR, Neutrophil Percentage-to-Albumin Ratio; OR, Odds ratio; P, P-value; PLT, Platelets; PR AUC, Precision-recall area under the curve; PSM, Propensity score matching; RCS, Restricted cubic spline; RDW, Red cell distribution width; ROC, Receiver operating characteristic; SAD, Small airway dysfunction; SD, Standard deviation; SHAP, SHapley Additive exPlanations; SII, Systemic Immune-Inflammation Index; SMD, Standard mean difference; VIF, Variance inflation factor; WBC, White blood cell count.

Declarations

Ethics approval and consent to participate: The study protocol was reviewed and approved by the Ethics Committee of Northern Jiangsu People's Hospital (Approval Number:2026KY232). Considering that this study is retrospective in design, informed consent was waived.

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: This work was supported by the National Natural Science Foundation of China [Grant numbers 81870033]; the Six Talent Peaks Project of Jiangsu Province [Grant number WSN-106]; the Medical Scientific Research Foundation of Jiangsu Province of China [Grant numbers QNRC2016340].

Authors' contributions: YL contributed to data collection and statistical analysis, and was a major contributor in writing the manuscript. RX, YJC, YYY, YQZ contributed to data collection and manuscript review. MLF contributed to study conception, study framework design, funding acquisition, manuscript drafting, and manuscript revision. All authors read and approved the final manuscript.

Acknowledgements: We would like to thank all the participating authors.

References

1. de Oca MM, Perez-Padilla R, Celli B, Aaron SD, Wehrmeister FC, Amaral AFS, et al. The global burden of COPD: epidemiology and effect of prevention strategies. Lancet Respir Med. 2025;13(8):709–24.
2. GBD 2021 COPD Collaborators. Burden of chronic obstructive pulmonary disease and its attributable risk factors in 204 countries and territories, 1990–2021: results from the Global Burden of Disease Study 2021. BMJ Public Health. 2026;4:e002489.
3. Hogg JC, Timens W. The pathology of chronic obstructive pulmonary disease. Annu Rev Pathol. 2009;4:435–59.
4. Zhou R, Wang H, Zhang Y, Mai J, Yang L. Small airway disease as a key factor in COPD: new perspectives and insights. Front Med (Lausanne). 2025;12:1648612.
5. Fakhraei R, Matelski J, Gershon A, Kendzerska T, Lapointe-Shaw L, Kaneswaran L, et al. Development of multivariable prediction models for the identification of patients admitted to hospital with an exacerbation of COPD and the prediction of risk of readmission: a retrospective cohort study using electronic medical record data. COPD. 2023;20(1):274–83.
6. Li M, Cheng K, Ku K, Li J, Hu H, Ung COL. Modelling 30-day hospital readmission after discharge for COPD patients based on electronic health records. NPJ Prim Care Respir Med. 2023;33(1):16.
7. Ruan H, Zhao H, Wang J, Zhang H, Li J. All-cause readmission rate and risk factors of 30- and 90-day after discharge in patients with chronic obstructive pulmonary disease: a systematic review and meta-analysis. Ther Adv Respir Dis. 2023;17:17534666231202742.
8. Kong CW, Wilkinson TMA. Predicting and preventing hospital readmission for exacerbations of COPD. ERJ Open Res. 2020;6(2):00325–2019.
9. Cao Z, Ong KC, Eng P, Tan WC, Ng TP. Frequent hospital readmissions for acute exacerbation of COPD and their associated factors. Respirology. 2006;11(2):188–95.
10. Yılmaz C, Özkan S, Erer OF. Risk assessment and rate of readmission within 30 days of discharge after hospitalization for acute exacerbation of chronic obstructive pulmonary disease. Tuberk Toraks. 2021;69(3):328–37.
11. Neder JA, Berton DC, O'Donnell DE. Small airway disease: when the silent zone speaks up. J Bras Pneumol. 2023;48(6):e20220414.
12. Hogg JC, Chu F, Utokaparch S, Woods R, Elliott WM, Buzatu L, et al. The nature of small-airway obstruction in chronic obstructive pulmonary disease. N Engl J Med. 2004;350(26):2645–53.
13. Lazarinis N, Fouka E, Linden A, Bossios A. Small airways disease in chronic obstructive pulmonary disease. Expert Rev Respir Med. 2024;18(7):539–52.
14. Tanabe N, Verleden SE. From Web to Collapse: Small Airway Disease Trajectory in Chronic Obstructive Pulmonary Disease. Am J Respir Crit Care Med. 2025;211(12):2231–3.
15. Wen G, Meng J, Xu Y, Wang R, Wang H, Peng P, et al. Prevalence and risk factors of spirometry-defined small airway dysfunction in the high-risk population for COPD in Yunnan Province, China: a population-based cross-sectional study. Int J Chron Obstruct Pulmon Dis. 2025;20:3005–17.

16. Rath AK, Sahu D, De S. Oscillometry-defined small airway dysfunction in patients with chronic obstructive pulmonary disease. Tuberc Respir Dis (Seoul). 2024;87(2):165–75.
17. Abdo M, Watz H, Trinkmann F, Bohnet S, Guess MAM, Roeben J, et al. Oscillometry-defined small airway dysfunction in tobacco-exposed adults with impaired or preserved airflow. Am J Respir Crit Care Med. 2025;211(9):1652–61.
18. Toumpanakis D, Kim Y, Usmani OS. Small airways disease in patients with COPD: a question-and-answer approach for everyday clinical practice. Chest. 2026;169(3):641–51.
19. Suter P, Greig R, Chan R, Lipworth BJ. Revisiting type 2 inflammation and airway hyper-responsiveness in COPD. J Allergy Clin Immunol Pract. 2026;14(1):67–75.
20. Pirera E, Di Raimondo D, Tuttolomondo A. Triple therapy de-escalation and withdrawal of inhaled corticosteroids to dual bronchodilator therapy in patients with chronic obstructive pulmonary disease (COPD): a systematic review and meta-analysis. J Clin Med. 2024;13(20):6199.
21. Chen F, Yang M, Wang H, Liu L, Shen Y, Chen L. High blood eosinophils predict the risk of COPD exacerbation: a systematic review and meta-analysis. PLoS ONE. 2024;19(10):e0302318.
22. Abdo M, Pedersen F, Trinkmann F, Herth FJF, Rabe KF, Kirsten AM, et al. Association of airway eosinophilia with small airway dysfunction in patients with mild and at risk for COPD. Int J Chron Obstruct Pulmon Dis. 2022;17:1403–8.
23. Wu X, Deng Z, Wu F, Zheng Y, Huang P, Yang H, et al. Clinical characteristics and 2-year outcomes of chronic obstructive pulmonary disease patients with high blood eosinophil counts: a population-based prospective cohort study in China. Arch Bronconeumol. 2024;60(7):402–9.
24. Bikov A, Csoma B, Chai A, Croft E, Lazar Z, Bentley A. The association between the eosinophilic COPD phenotype with overall survival and exacerbations in patients on long-term non-invasive ventilation. Biomolecules. 2025;15(12):1728.
25. Håkansson KEJ, Ulrik CS, Godtfredsen NS, Kallemose T, Andersen O, Eugen-Olsen J, et al. High suPAR and low blood eosinophil count are risk factors for hospital readmission and mortality in patients with COPD. Int J Chron Obstruct Pulmon Dis. 2020;15:733–43.
26. Shin SH, Park HY, Kang D, Cho J, Kwon SO, Park JH, et al. Serial blood eosinophils and clinical outcome in patients with chronic obstructive pulmonary disease. Respir Res. 2018;19(1):134.
27. Lan CC, Su WL, Yang MC, Chen SY, Wu YK. Predictive role of neutrophil-percentage-to-albumin, neutrophil-to-lymphocyte and eosinophil-to-lymphocyte ratios for mortality in patients with COPD: evidence from NHANES 2011–2018. Respirology. 2023;28(12):1136–46.
28. Shi Y, Shi Y, Liu Y, Wang C, Qi M, Li C. Association between neutrophil percentage to serum albumin ratio and in-hospital mortality of patients with chronic obstructive pulmonary disease in intensive care unit: a retrospective cohort study. Int J Chron Obstruct Pulmon Dis. 2025;20:1227–37.
29. Xiao D, Chen Z, Wu S, Huang K, Xu J, Yang L, et al. Prevalence and risk factors of small airway dysfunction, and association with smoking, in China: findings from a national cross-sectional study. Lancet Respir Med. 2020;8(11):1081–93.
30. Wang Z, Gao Y, Jian W, Li Y, Wang Z, Lin J, et al. Establishment and application of reference equations for FEF(50) and FEF(75) in the Chinese population. J Thorac Dis. 2024;16(1):379–90.
31. Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: 2023 report. Fontana, WI: Global Initiative for Chronic Obstructive Lung Disease (GOLD); 2023.
32. Maetani T, Tanabe N, Sato A, Shiraishi Y, Sakamoto R, Ogawa E, et al. Association between blood eosinophil count and small airway eosinophils in smokers with and without COPD. ERJ Open Res. 2023;9(5):00235–2023.
33. Balzano G, Stefanelli F, Iorio C, De Felice A, Melillo EM, Martucci M, et al. Eosinophilic inflammation in stable chronic obstructive pulmonary disease. Relationship with neutrophils and airway function. Am J Respir Crit Care Med. 1999;160(5 Pt 1):1486–92.
34. Li M, Zheng A, Liu Y. Molecular mechanisms linking eosinophils to lung function impairment in respiratory diseases. J Asthma Allergy. 2026;19:625304.

35. Lapperre TS, Willems LN, Timens W, Rabe KF, Hiemstra PS, Postma DS, et al. Small airways dysfunction and neutrophilic inflammation in bronchial biopsies and BAL in COPD. Chest. 2007;131(1):53–9.
36. Yamaya M, Usami O, Nakayama S, Tode N, Yamada A, Ito S, et al. Malnutrition, airflow limitation and severe emphysema are risks for exacerbation of chronic obstructive pulmonary disease in Japanese subjects: a retrospective single-center study. Int J Chron Obstruct Pulmon Dis. 2020;15:857–68.
37. Fischer BM, Pavlisko E, Voynow JA. Pathogenic triad in COPD: oxidative stress, protease-antiprotease imbalance, and inflammation. Int J Chron Obstruct Pulmon Dis. 2011;6:413–21.
38. Fricker M, McDonald VM, Winter NA, Baines KJ, Wark PAB, Simpson JL, et al. Molecular markers of type 2 airway inflammation are similar between eosinophilic severe asthma and eosinophilic chronic obstructive pulmonary disease. Allergy. 2021;76(7):2079–89.
39. Gorska K, Krenke R, Korczynski P, Kosciuch J, Domagala-Kulawik J, Chazan R. Eosinophilic airway inflammation in chronic obstructive pulmonary disease and asthma. J Physiol Pharmacol. 2008;59(Suppl 6):261–70.
40. Ko FWS, Chan KKP, Ng TCC, Lo RLP, Ngai JCL, To KW, et al. A 3-year prospective study to assess clinical characteristics and risk factors for exacerbations in patients with asthma-COPD overlap based on the GINA guideline compared with patients with asthma and COPD. Respir Res. 2026;27(1):202.
41. Menezes AMB, Montes de Oca M, Pérez-Padilla R, Nadeau G, Wehrmeister FC, Lopez-Varela MV, et al. Increased risk of exacerbation and hospitalization in subjects with an overlap phenotype: COPD-asthma. Chest. 2014;145(2):297–304.
42. Kim MA, Noh CS, Chang YJ, Hong YK, Lee JS, Lee SW, et al. Asthma and COPD overlap syndrome is associated with increased risk of hospitalisation. Int J Tuberc Lung Dis. 2015;19(7):864–9.
43. Zhou XL, Zhao LY. Comparison of clinical features and outcomes for asthma-COPD overlap syndrome vs. COPD patients: a systematic review and meta-analysis. Eur Rev Med Pharmacol Sci. 2021;25(3):1495–510.

Figures

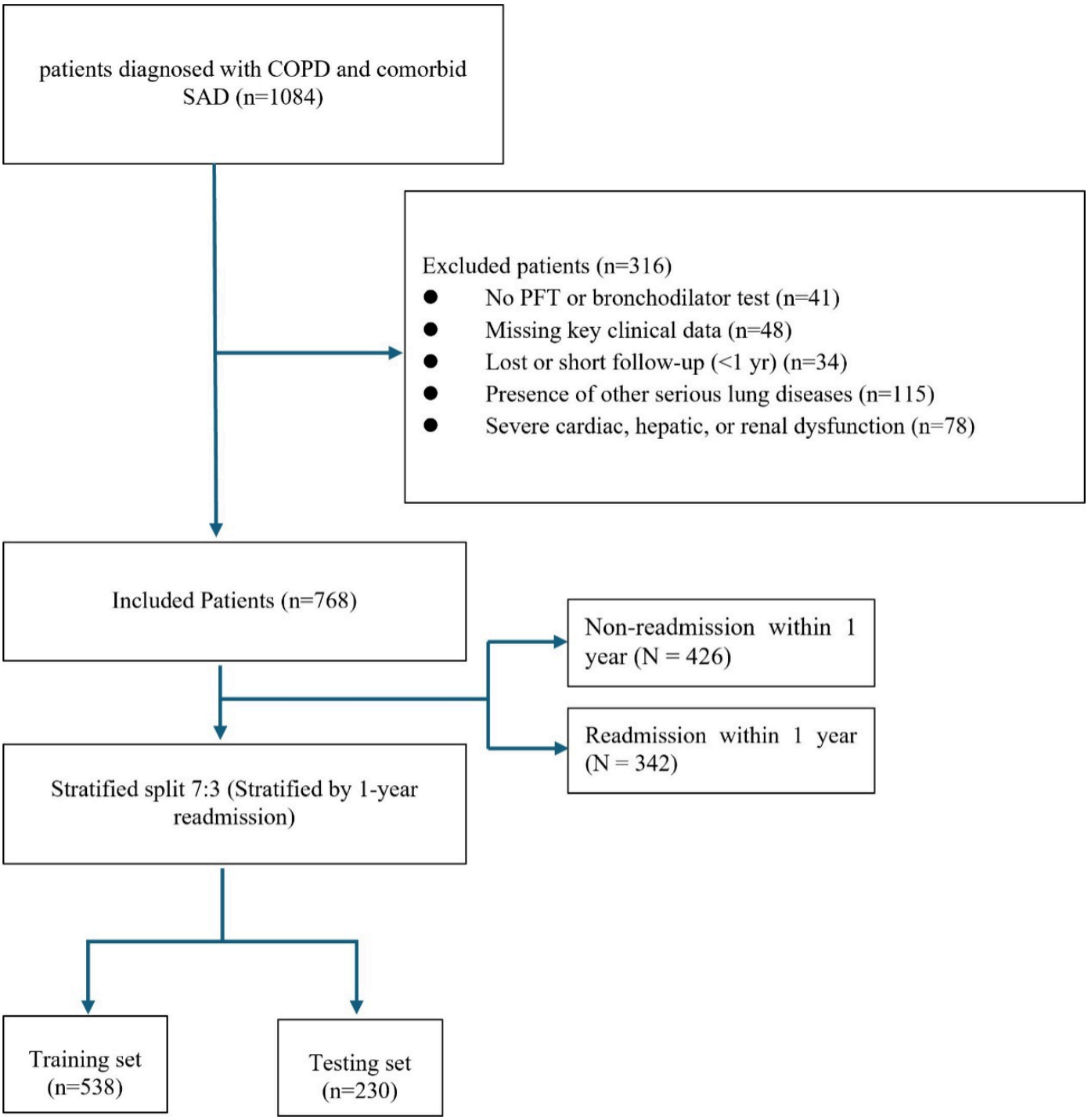


Figure 1

Flowchart of patient selection

Abbreviations: COPD, chronic obstructive pulmonary disease; SAD, small airway dysfunction; PFT, pulmonary function test.

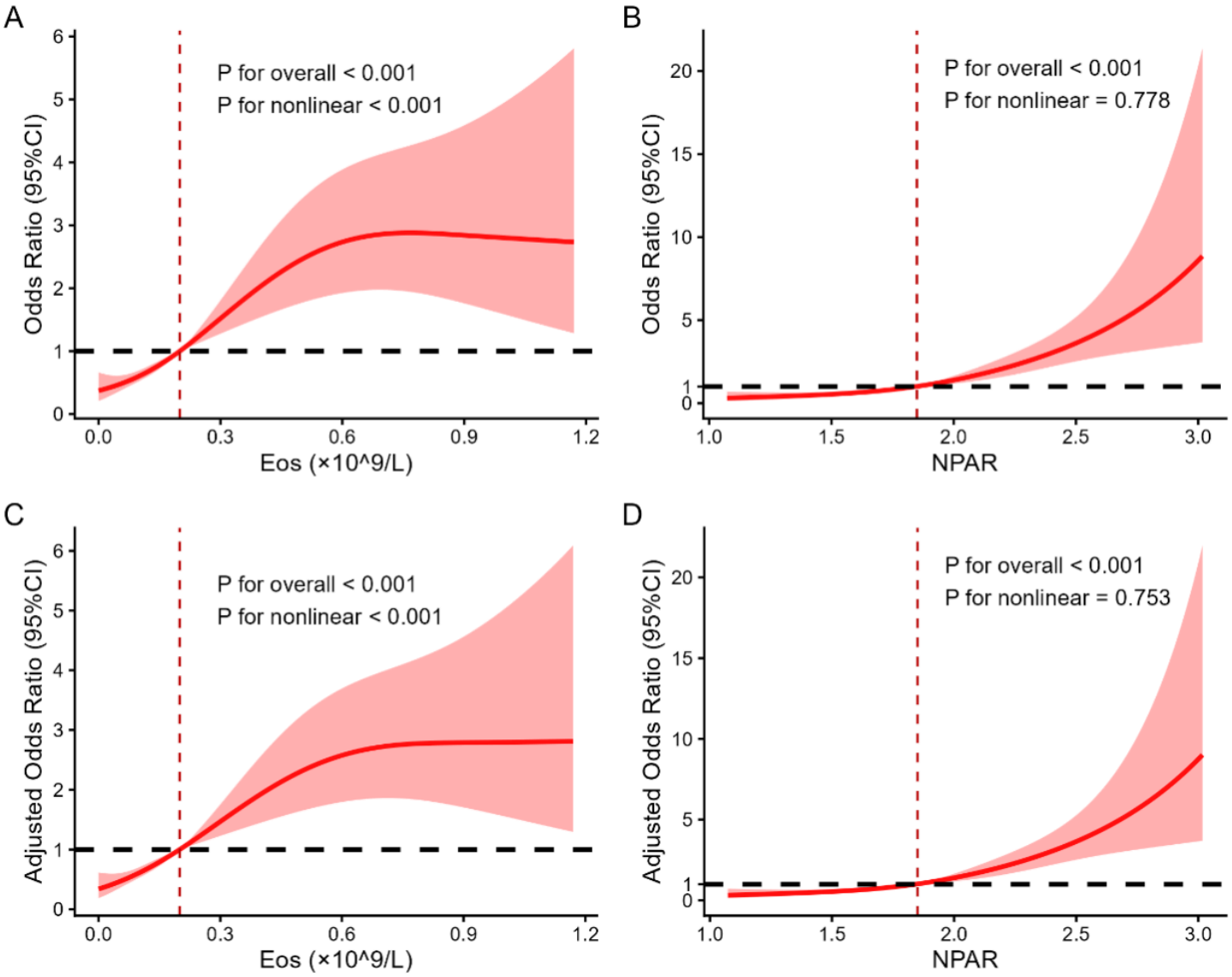


Figure 2

RCS analysis of Eos and NPAR with risk of 1-year readmission. (A) Eos without adjustment; (B) NPAR without adjustment; (C) Eos after adjustment for Model 3; (D) NPAR after adjustment for Model 3

Abbreviations: RCS, restricted cubic splines; Eos, eosinophil count; NPAR, neutrophil percentage-to-albumin ratio.

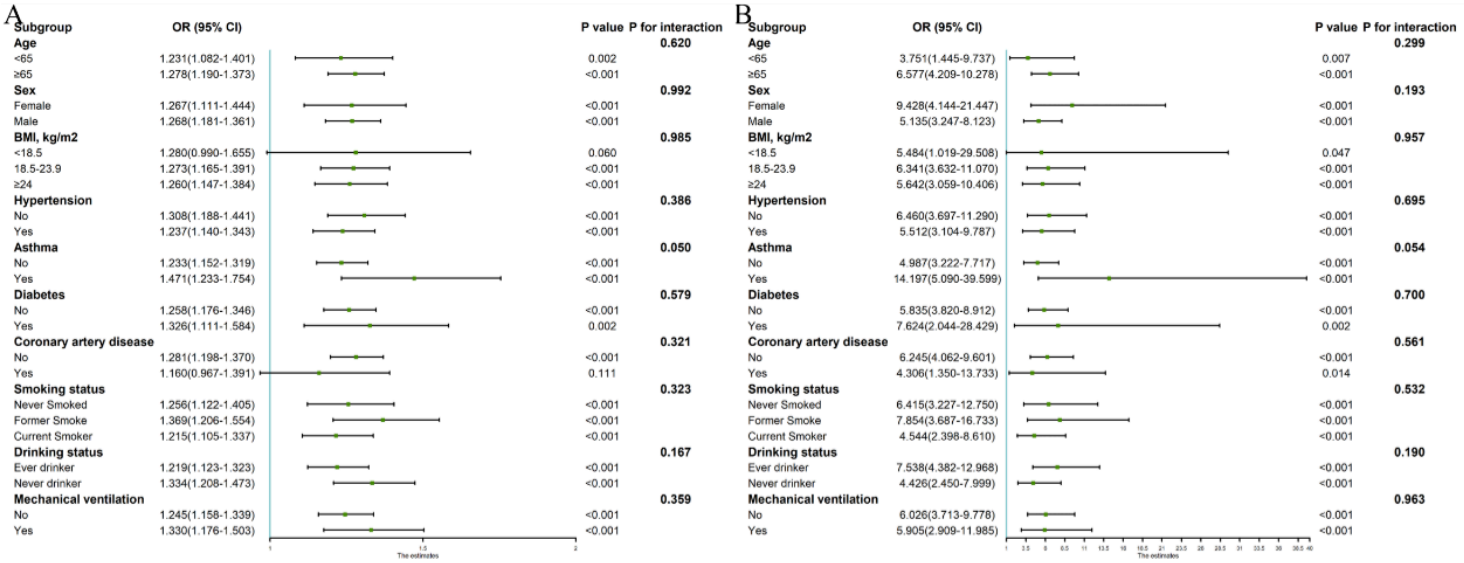


Figure 3

Forest plots for subgroup analyses of Eos and NPAR. (A) Subgroup analysis of Eos; (B) Subgroup analysis of NPAR

Notes: ORs are presented for Eos per 0.1×10⁹/L (100 cells/μL) and for NPAR per 1-unit increase. All models were adjusted for age, sex, BMI, smoking index, hypertension, diabetes, marital status, drinking status, coronary artery disease and use of mechanical ventilation. (same covariates as Model 3 of the primary analysis). *P* values for interaction were derived from multiplicative interaction terms between each exposure and subgroup variable.

Abbreviations: Eos, eosinophil count; NPAR, neutrophil percentage-to-albumin ratio; OR, odds ratio; CI, confidence interval.

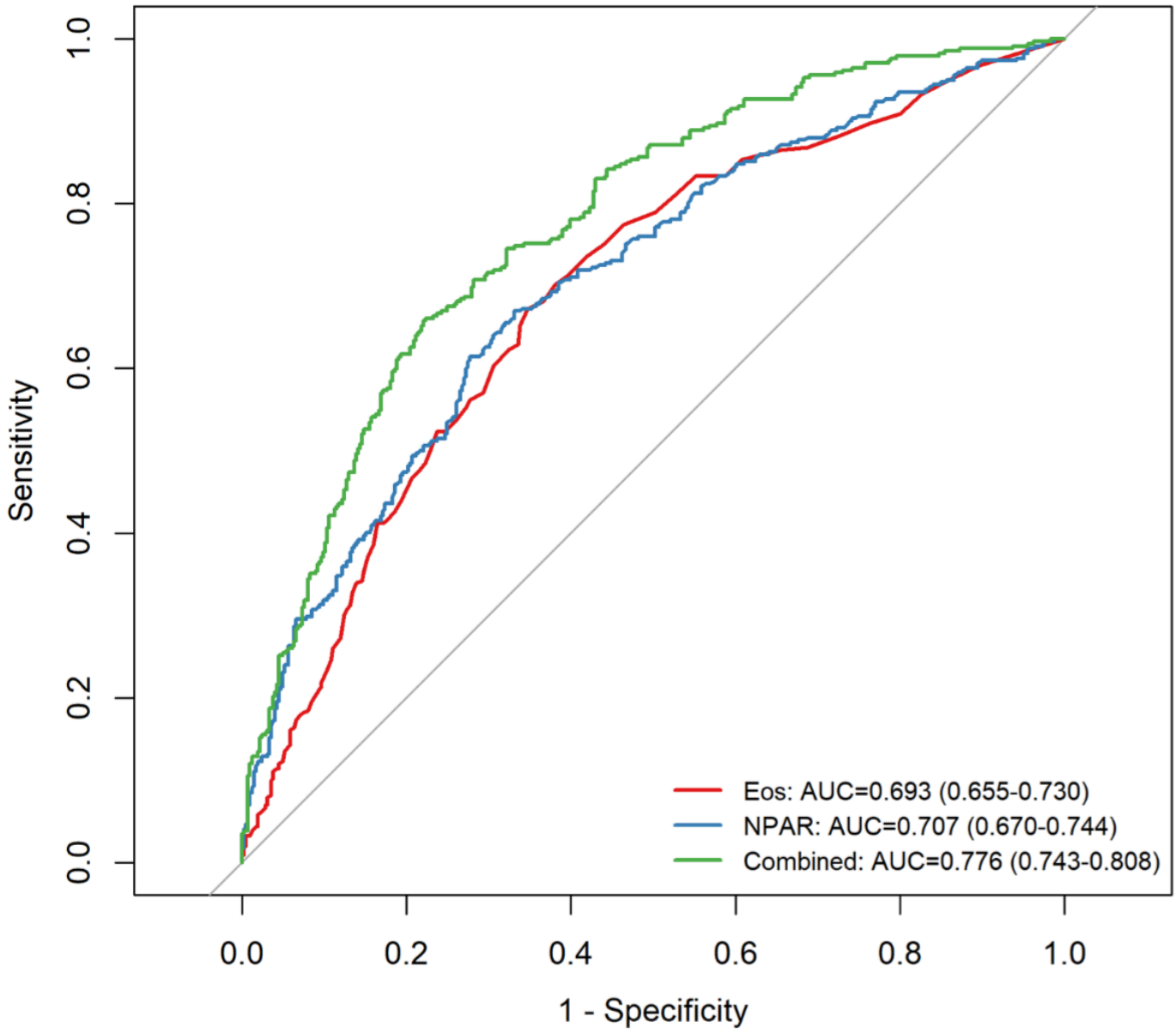


Figure 4

ROC curves of Eos, NPAR, and the combined model for predicting 1-year readmission

Notes: The combined model includes both Eos and NPAR as predictors. AUC values with 95% confidence intervals are shown in the legend. The DeLong test was used to compare AUC differences between models.

Abbreviations: ROC, receiver operating characteristic; AUC, area under the curve; Eos, eosinophil count; NPAR, neutrophil percentage-to-albumin ratio.

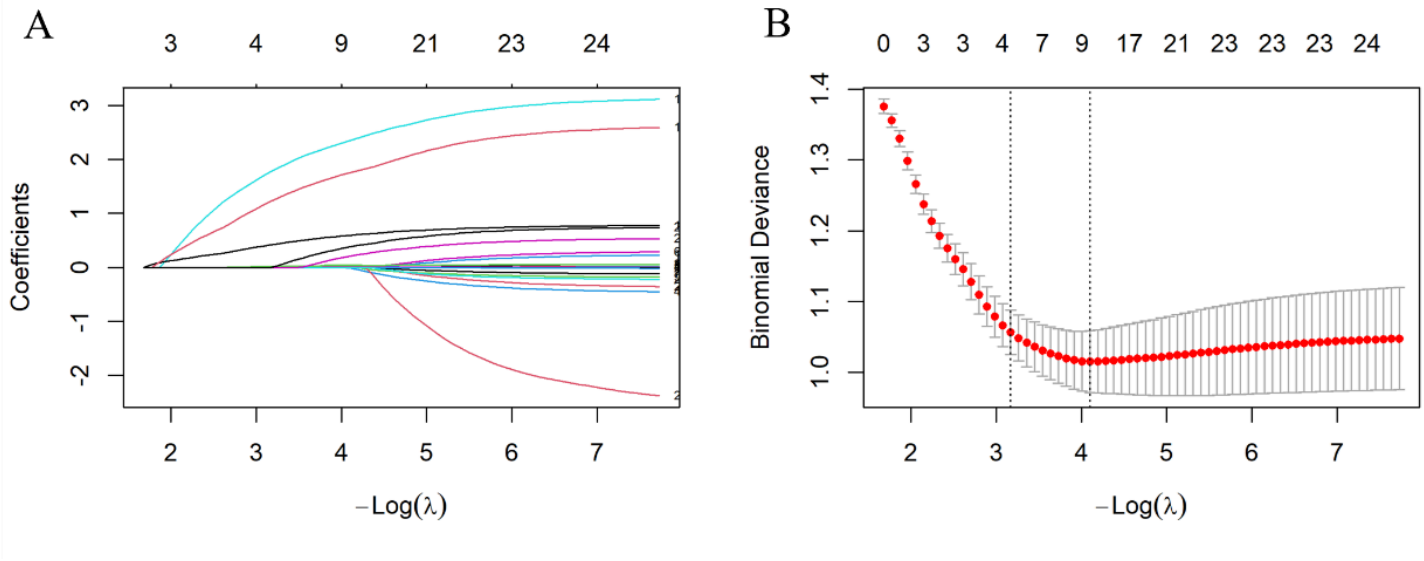


Figure 5

Selection of final predictors. (A) LASSO coefficient paths for the 25 covariates. (B) Cross-validation plot for penalty parameter selection.

Notes: Vertical dashed lines in (B) represent λmin = 0.01657194 (right) and $\lambda_1$se = 0.04201591 (left). λmin was selected as the optimal penalty parameter.

Abbreviations: LASSO, least absolute shrinkage and selection operator.

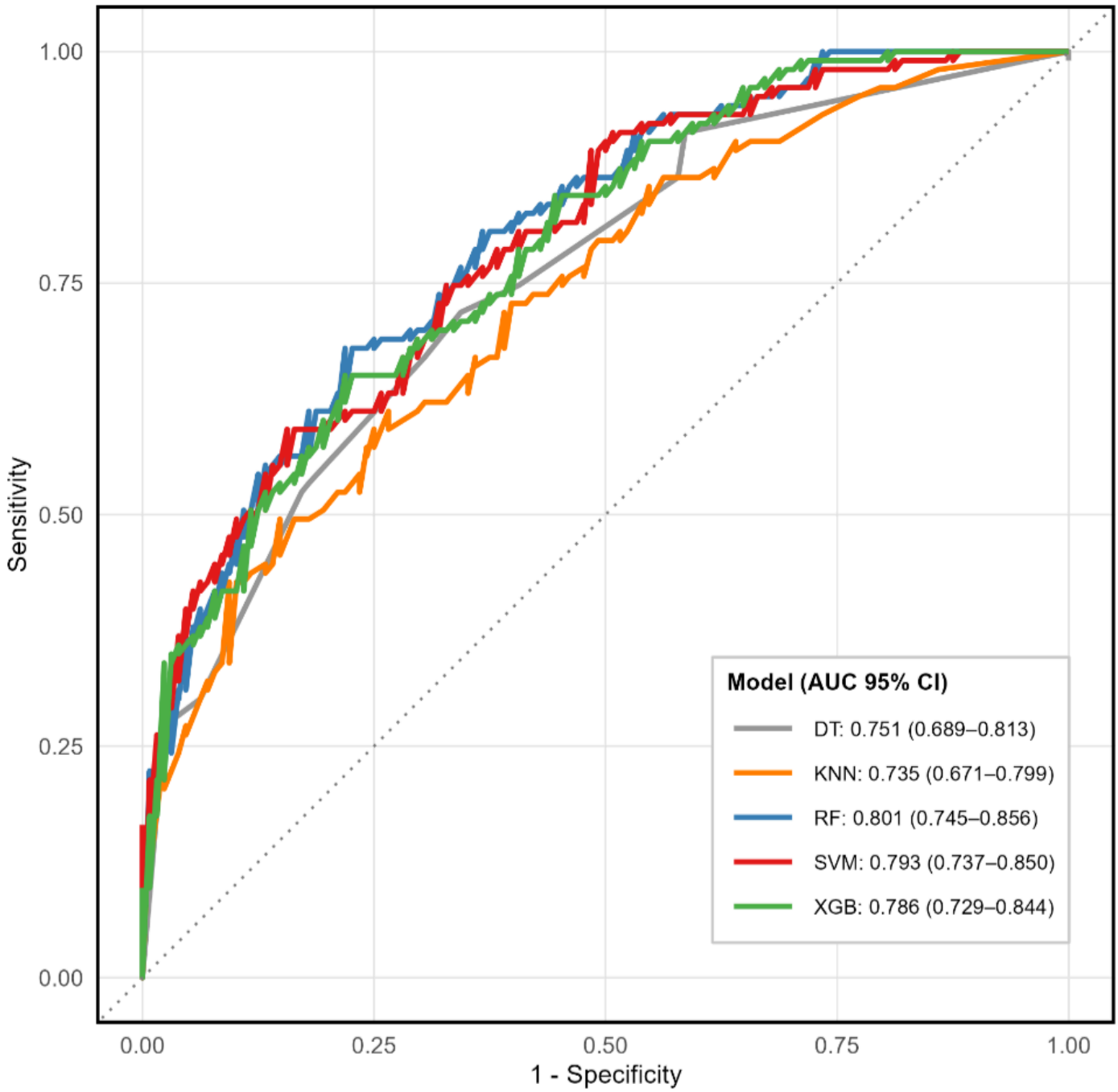


Figure 6

ROC curves of five machine learning models in the testing set

Notes: AUC values are displayed in the legend. The diagonal dashed line indicates the reference level (AUC = 0.5).

Abbreviations: ROC, receiver operating characteristic; AUC, area under the curve; DT, Decision Tree; RF, random forest; XGBoost, extreme gradient boosting; SVM, support vector machine; KNN, k-nearest neighbors.

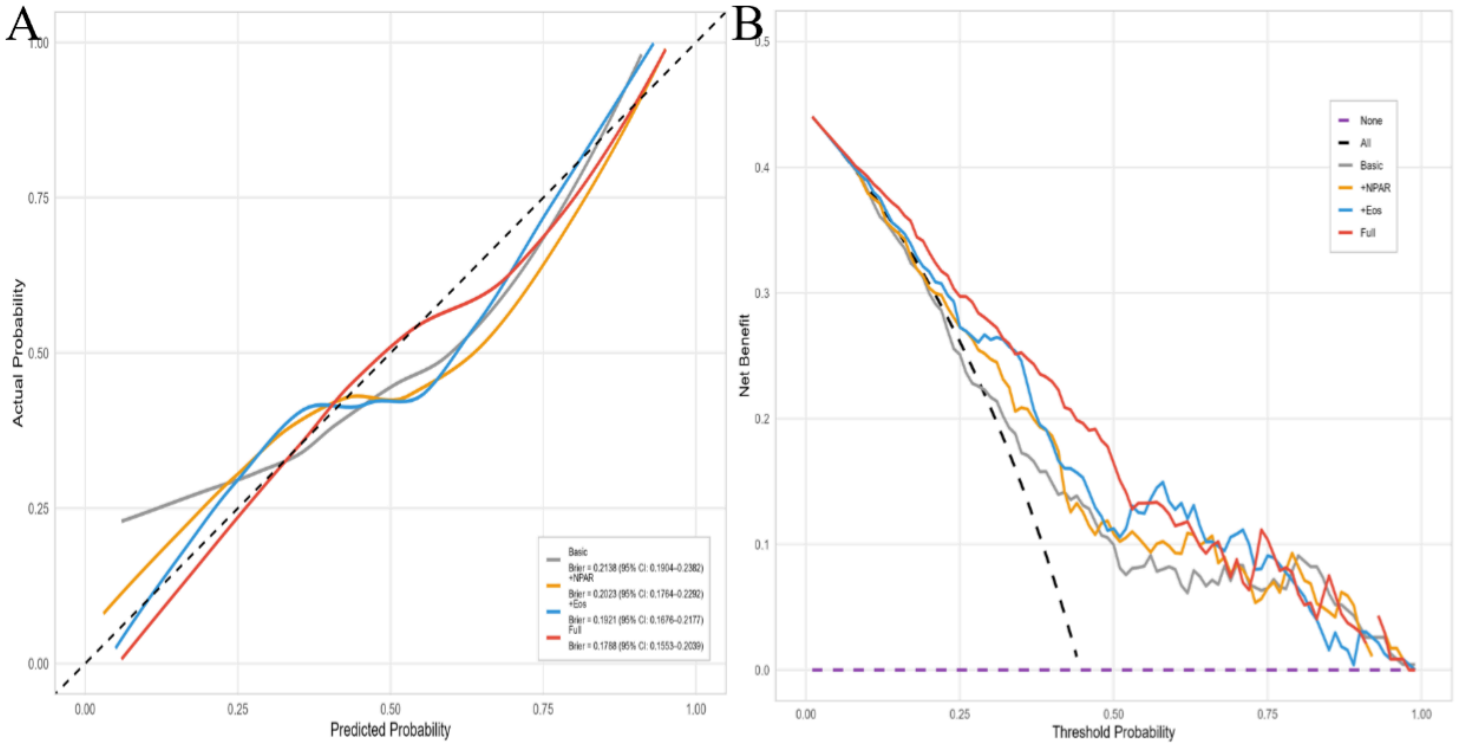


Figure 7

Calibration curves and decision curve analysis of the prediction models. (A) Calibration curves of four models. (B) Decision curve analysis (DCA) of the four models

Notes: The diagonal dashed line represents the ideal calibration in (A). The Full model showed the closest alignment with the ideal line and achieved the lowest Brier score. The y-axis represents the net benefit, and the x-axis represents the threshold probability in (B). The Full model demonstrated the highest net benefit across the clinically relevant threshold range. The gray lines represent the assumptions that all patients are readmitted ("All") or none are readmitted ("None").

Abbreviations: DCA, decision curve analysis; Brier score, mean squared difference between predicted probabilities and observed outcomes.

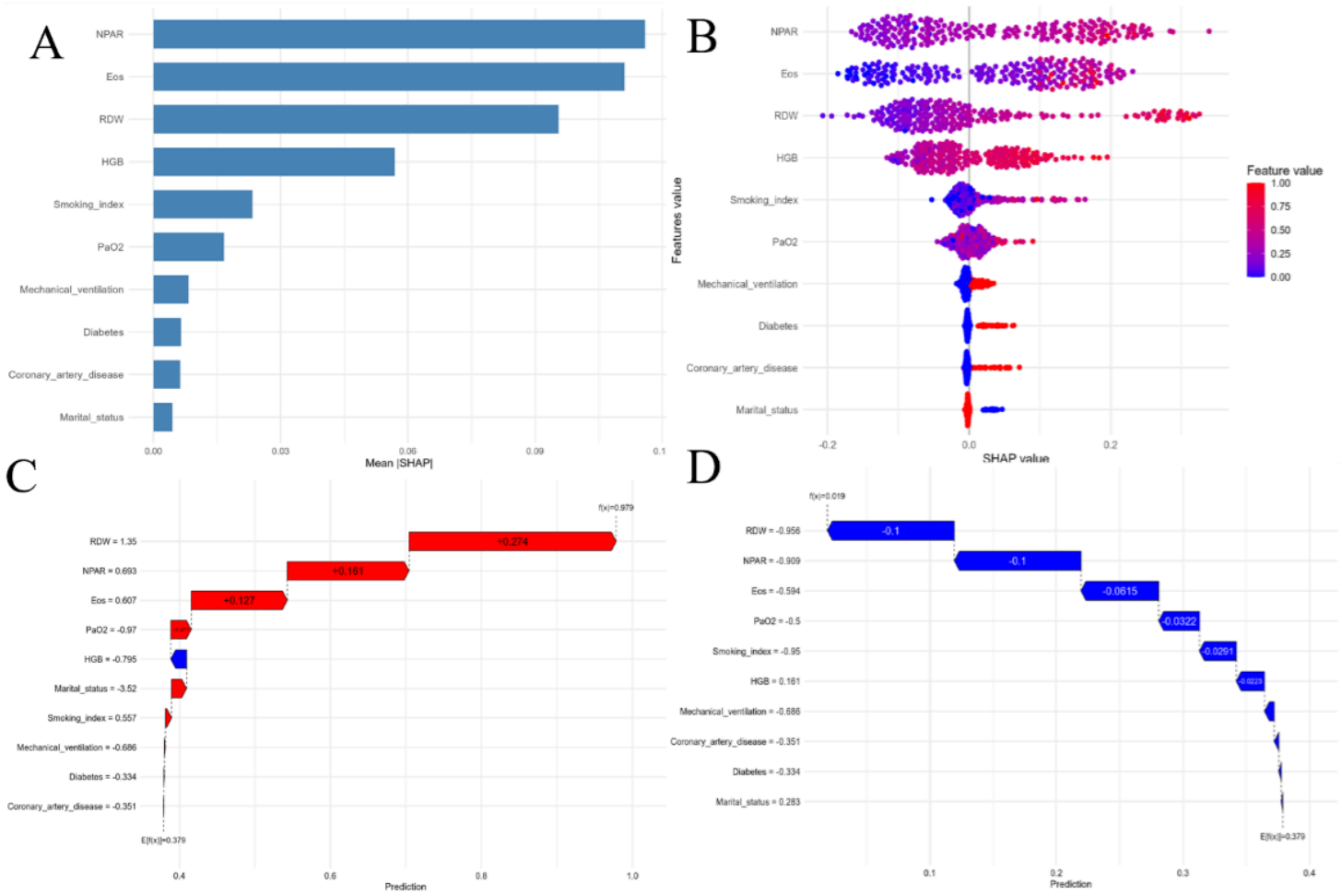


Figure 8

SHAP analysis of the random forest full model. (A) Feature importance ranking based on mean absolute SHAP values. (B) SHAP summary plot showing the distribution and direction of feature contributions. (C) SHAP waterfall plot for the high-risk sample. (D) SHAP waterfall plot for the low-risk sample

Notes: In (B), the color gradient represents SHAP values ranging from low (blue) to high (red). In (C) and (D), red bars indicate variables that push the predicted probability toward high risk, while blue bars indicate variables that push it toward low risk.

Abbreviations: SHAP, Shapley additive explanations; NPAR, neutrophil percentage-to-albumin ratio; Eos, eosinophil count; RDW, red cell distribution width; HGB, hemoglobin.

Supplementary Files

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

- Additionalfile1.docx