Journal of Inflammation Research

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Predicting Superaverage Length of Stay in COPD Patients with Hypercapnic Respiratory Failure Using Machine Learning

Bingqing Zuo[1,*], Lin Jin[2,*], Zhixiao Sun[1], Hang Hu[1], Yuan Yin[3], Shuanying Yang[4], Zhongxiang Liu[1,4]

[1]Department of Pulmonary and Critical Care Medicine, The Yancheng Clinical College of Xuzhou Medical University, The First People's Hospital of Yancheng, Yancheng, Jiangsu, 224006, People's Republic of China; [2]Third Department of Respiratory and Critical Care Medicine, First Affiliated Hospital of Kunming Medical University, Kuming, Yunnan, 650000, People's Republic of China; [3]Department of Respiratory and Critical Care Medicine, The First Affiliated Hospital of Nanjing Medical University, Nanjing, Jiangsu, 210029, People's Republic of China; [4]Department of Respiratory and Critical Care Medicine, The Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shanxi, 710004, People's Republic of China

*These authors contributed equally to this work

Correspondence: Shuanying Yang; Zhongxiang Liu, Email yangshuanying112@163.com; liuzhongxiang711@163.com

Objective: The purpose of this study was to develop and validate machine learning models that can predict superaverage length of stay in hypercapnic-type respiratory failure and to compare the performance of each model. Furthermore, screen and select the optimal individualized risk assessment model. This model is capable of predicting in advance whether an inpatient's length of stay will exceed the average duration, thereby enhancing its clinical application and utility.

Methods: The study included 568 COPD patients with hypercapnic respiratory failure, 426 inpatients from the Department of Respiratory and Critical Care Medicine of Yancheng First People's Hospital in the modeling group and 142 inpatients from the Department of Respiratory and Critical Care Medicine of Jiangsu Provincial People's Hospital in the external validation group. Ten machine learning algorithms were used to develop and validate a model for predicting superaverage length of stay, and the best model was evaluated and selected.

Results: We screened 83 candidate variables using the Boruta algorithm and identified 9 potentially important variables, including: cerebrovascular disease, white blood cell count, hematocrit, D-dimer, activated partial thromboplastin time, fibrin degradation products, partial pressure of carbon dioxide, reduced hemoglobin, and oxyhemoglobin. Cerebrovascular disease, hematocrit, activated partial thromboplastin time, partial pressure of carbon dioxide, reduced hemoglobin and oxyhemoglobin were independent risk factors for superaverage length of stay in COPD patients with hypercapnic respiratory failure. The Catboost model is the optimal model on both the modeling dataset and the external validation set. The interactive web calculator was developed using the Shiny framework, leveraging a predictive model based on Catboost.

Conclusion: The Catboost model has the most advantages and can be used for clinical evaluation and patient monitoring.

Keywords: chronic obstructive pulmonary disease, COPD, hypercapnic respiratory failure, HRF, superaverage length of stay, machine learning, Catboost model

Introduction

Hypercapnic respiratory failure (HRF), usually defined as arterial partial pressure of carbon dioxide (PaCO2) ≥45 mmHg and often accompanied by a decrease in arterial partial pressure of oxygen (PaO2) can occur in a variety of etiologies, primarily in chronic respiratory diseases such as exacerbations of chronic obstructive pulmonary disease (COPD), cystic fibrosis, thoracic deformities, and other conditions such as neuromuscular disease.[1,2] The end-stage of COPD often leads to HRF, which is associated with debilitating symptoms and a low survival rate.[3]

However, the majority of patients with HRF require hospitalization, and many require ventilatory support in a dedicated intensive care unit, incurring considerable healthcare costs.[4,5] In the United States, acute respiratory failure (ARF) treatment costs more than $50 billion annually due to high mortality rates and long hospital stays, especially for older patients.[6–8]

Received: 8 December 2024
Accepted: 19 April 2025
Published: 8 May 2025

© 2025 Zuo et al. This work is published and licensed by Dove Medical Press Limited. The full terms of this license are available at https://www.dovepress.com/terms.php and incorporate the Creative Commons Attribution – Non Commercial (unported, v4.0) License (http://creativecommons.org/licenses/by-nc/4.0/). By accessing the work you hereby accept the Terms. Non-commercial uses of the work are permitted without any further permission from Dove Medical Press Limited, provided the work is properly attributed. For permission for commercial use of this work, please see paragraphs 4.2 and 5 of our Terms (https://www.dovepress.com/terms.php).

Especially as the global healthcare industry gradually and steadily shifts from fee-for-service to value-based care agreements, length of stay (LOS) is a useful indicator of resource utilization and cost-effectiveness, and has been likened to an indicator for reducing Medicare expenditures.[9,10] Therefore, the prolongation of hospital stay has an important impact on the resource consumption and revenue management of medical institutions, and should not be discarded.[11]

However, in order to better clarify the reasons for prolonged hospital stay in order to improve long-term care strategies for these patients, at the same time, increased knowledge of prediction and prediction of prolonged length of stay may contribute to earlier and better patient treatment, optimal disease planning, and shorter length of stay, which will ultimately help in cost assessment for hospitals.[12] There are no studies investigating risk factors for superaverage length of stay in COPD patients with HRF, so it is important to look for a cause in patients with HRF to reduce the length of hospital stay and mitigate high medical costs.[13]

Machine learning (ML) algorithms can screen for risk factors that affect the length of hospital stay in patients with HRF by analyzing large amounts of data from electronic health records. These data may include patient comorbidities, demographics, and routine examination findings. These cutting-edge analytical methods can process high-dimensional data and analyze complex relationships. They are more flexible than traditional modeling techniques.[14,15]

This study aims to be used to develop and validate an ML model that can predict the superaverage length of stay in COPD patients with HRF.

Materials and Methods

Data Source

Modeling dataset: We selected COPD patients combined with HRF who were hospitalized in the Department of Respiratory and Critical Care Medicine of Yancheng First People's Hospital from October 2020 to September 2021, and the external validation dataset: COPD patients combined with HRF who were hospitalized in the Department of Respiratory and Critical Care Medicine of Jiangsu Provincial People's Hospital from October 2021 to December 2021. The diagnosis of COPD was determined according to the Global Initiative for Chronic Obstructive Pulmonary Disease (GOLD) strategy.[16] Diagnostic criteria for respiratory failure were as follows: arterial partial pressure of oxygen (PaO2) < 8.0 kPa (60 mmHg) and arterial partial pressure of carbon dioxide (PaCO2)>6.0 kPa (45 mmHg) as measured by blood gas analysis.[17]

Patients who had incomplete clinical data, were under the age of 60, experienced death during hospitalization or abandoned treatment, had hearing or speech impairment that hindered their ability to answer questions adequately, required tracheotomy or intubation, suffered from organ failure, or were unable to provide informed consent were excluded. Exclusions also applied to patients with trauma, malignancy (including hematological malignancy), and pregnancy.

Study Variables

The following clinical data were collected within 24 hours of admission: demographic information, clinical characteristics, comorbidities, various rating scales at the time of admission, laboratory test results (blood sample tests), etc. The length of stay in COPD patients combined with HRF was defined as the duration between admission and discharge.

The flowchart illustrating the experimental procedure of the multicenter post-study was depicted in Figure 1. Firstly, 426 COPD patients combined with HRF in Yancheng First People's Hospital were used as a modeling dataset, and 142 COPD patients combined with HRF provided by Jiangsu Provincial People's Hospital were used for external validation. The current project followed the principles of the Declaration of Helsinki. This work was approved by the First People's Hospital of Yancheng City (No. 2020-K062) and the Ethics Committee of Jiangsu Provincial People's Hospital (No. 2021-SR-346). In addition, participants at both hospitals provided informed written consent to support the clinical study.

Model Building and Validation Methods

The study data included both modeled datasets and external validation datasets. Based on the modeling dataset, the Boruta method was used to screen the variables, and then the Random Forest (RF), Categorical Boosting (CatBoost), Light Gradient Boosting Machine (LightGBM) and Extreme Gradient were used respectively Boosting (XGBoost), Gradient Boosting Machine (GBM), Neural Network (NNET), Support Vector Machine (SVM), K-Nearest Neighbor

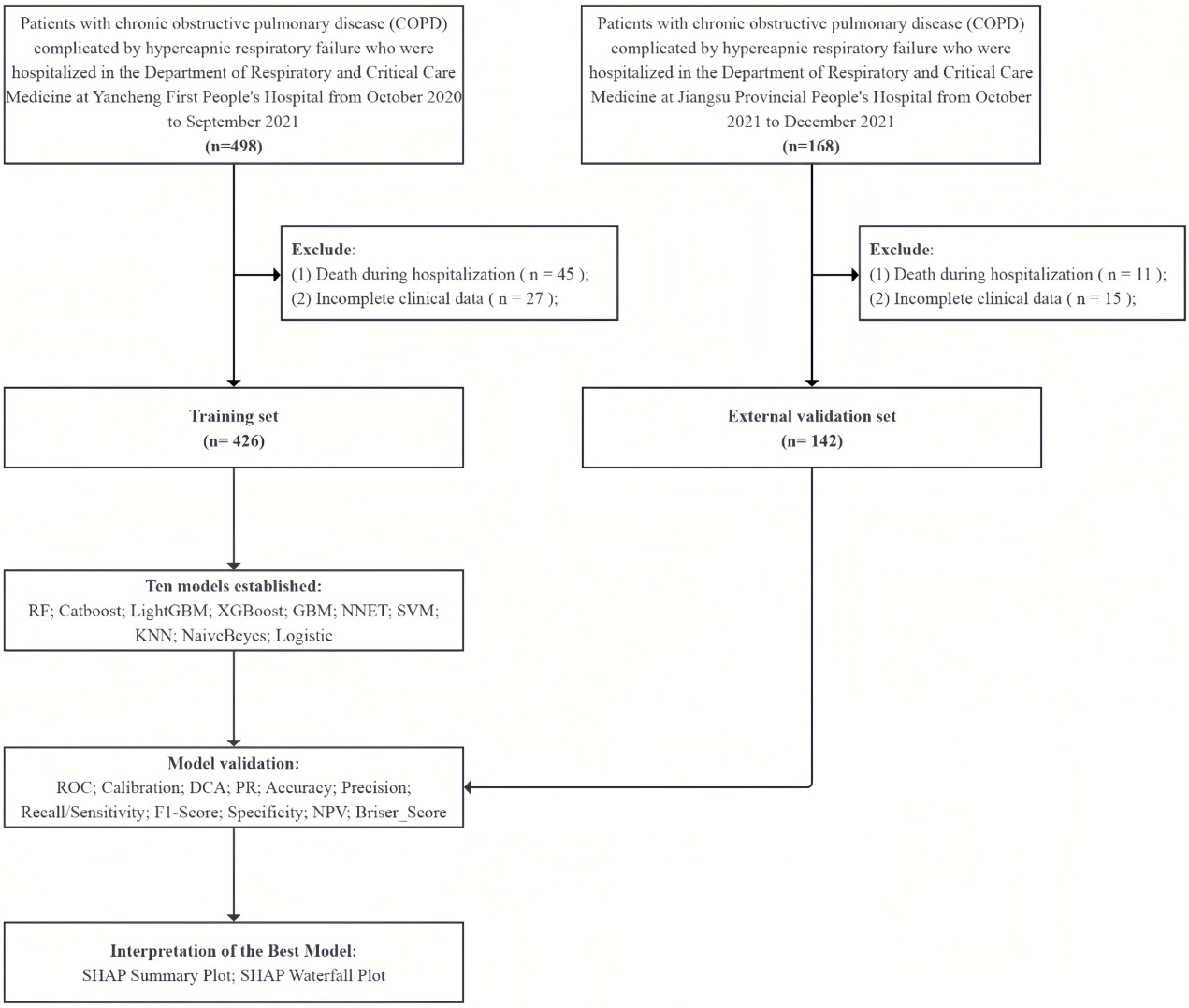

Figure 1 The flowchart of this study.

Algorithm (KNN), Naive Bayes Bayes and Logistic Regression, respectively, constructed models for predicting super-average length of stay. Each model was trained to determine the optimal hyperparameters through grid search and 10-fold cross-validation.

After the model was established, the performance of 10 models was evaluated on the modeling dataset and the external validation dataset. The performance of these models was evaluated by several metrics, including accuracy, precision (positive predictive value), recall, F1-score, sensitivity, specificity, negative predictive value, and Brier score. Subsequently, the area under the receiver operating characteristic curve (AUC) was used as the main criterion to measure the performance of the models, and the discrimination ability of the 10 models was evaluated. In addition, calibration curves were used to verify the consistency of the prediction effect of each model with the actual situation. At the same time, the decision curve analysis (DCA) was used to calculate the clinical net benefit of each model under different clinical decision thresholds. Finally, to evaluate model performance more comprehensively, a precision-recall (PR) curve was used, a tool that was particularly useful for class-imbalance datasets, to measure the performance of each model in terms of positive class predictions.

SHAP plots were used to explain machine learning models, where the length of each variable on the horizontal axis represents its contribution to the outcome, and the color of the dots indicated the size or category of the variable's

https://doi.org/10.2147/JIR.S511092

features. Through SHAP analysis, the contribution value of the features in each sample to the prediction, ie, the Shapley value, can be obtained. Based on the Shapley value, the SHAP summary graph and the single-sample SHAP waterfall chart were constructed. The SHAP summary chart ranked the risk factors in importance, while the single-sample SHAP waterfall chart explains the predictions of a single sample.

Statistical Methods

All variables were included in the comparison of the modeled dataset and the external validation dataset. In the comparison of the differences between the two groups, the chi-square test was used for categorical variables, and the results were displayed as frequency and percentage. Normally distributed continuous variables were expressed using the *t*-test, and the results were expressed as mean ± standard deviations; Continuous variables that do not conform to the normal distribution were performed using the Mann–Whitney *U*-test, and the results were expressed in median and interquartile ranges. All statistical analysis and machine learning algorithms were performed in R language (version 4.1.3), and a two-sided p-value of less than 0.05 was considered statistically significant.

Results

Clinical Features of COPD Patients Combined with Hypercapnic Respiratory Failure

A total of 568 COPD patients combined with HRF were included in the study, with 426 in the modelling dataset and 142 in the external validation set. The study defined a hospital stay exceeding 10 days as superaverage length of stay, based on findings indicating that the average length of stay for hypercapnia respiratory failure was approximately 10 days.[18,19] The findings of our study revealed that, on the whole, 60.56% of patients had a hospitalization duration of less than 10 days, while 39.44% required a hospital stay exceeding 10 days. There was no significant difference in the number of days of hospital stay between the two groups (p = 0.766). The median age of the patients was 74 years, and 66.55% were male, and there were no significant differences between the two groups in terms of gender, age, BMI, fall score, pressure ulcer score, self-care ability score, pain score, and VTE score (P > 0.05). There were significant differences between the two groups in terms of smoking history, cardiovascular disease, MMRC score, pneumonia, and white blood cell count (p < 0.05). In general, the distribution of most variables between the two groups was similar and comparable, as detailed in Table 1.

Table 1 Clinical Characteristics of COPD Patients Combined with Hypercapnic Respiratory Failure

Variables	Total (n = 568)	Training Set (n = 426)	External Validation Set (n = 142)	p
Length of stay				0.766
≤10	344 (60.56)	256 (60.09)	88 (61.97)	
>10	224 (39.44)	170 (39.91)	54 (38.03)	
Sex				0.837
0	190 (33.45)	144 (33.8)	46 (32.39)	
1	378 (66.55)	282 (66.2)	96 (67.61)	
Age	74 [69, 80]	75 [69, 80]	73.5 [69, 79.75]	0.618
BMI	21.26 [18.37, 24.61]	21.02 [18.29, 24.49]	21.84 [18.94, 25.03]	0.311
Smoking history				< 0.001
No	224 (39.44)	138 (32.39)	86 (60.56)	
Yes	344 (60.56)	288 (67.61)	56 (39.44)	
Treatment				0.265
0	313 (55.11)	235 (55.16)	78 (54.93)	
1	232 (40.85)	177 (41.55)	55 (38.73)	
2	23 (4.05)	14 (3.29)	9 (6.34)	

(*Continued*)

Table 1 (Continued).

Variables	Total (n = 568)	Training Set (n = 426)	External Validation Set (n = 142)	p
Fall score	4 [3, 5]	4 [3, 5]	4 [2.25, 5]	0.873
Pressure ulcer score	18 [15, 20]	18 [15, 20]	18 [15.25, 20]	0.984
Self-care score	58 [38, 78]	58 [38.75, 80]	60 [38, 71]	0.568
Pain score	0 [0, 0]	0 [0, 0]	0 [0, 0]	0.504
MMRC score	4 [3, 4]	4 [3, 4]	3 [3, 4]	0.010
VTE score	3 [1, 5]	3 [1, 5]	3 [2, 5]	0.321
Hypertension				0.233
No	346 (60.92)	266 (62.44)	80 (56.34)	
Yes	222 (39.08)	160 (37.56)	62 (43.66)	
Diabetes				0.239
No	487 (85.74)	370 (86.85)	117 (82.39)	
Yes	81 (14.26)	56 (13.15)	25 (17.61)	
Cerebrovascular disease				0.938
No	507 (89.26)	381 (89.44)	126 (88.73)	
Yes	61 (10.74)	45 (10.56)	16 (11.27)	
Cardiovascular disease				0.009
No	439 (77.29)	341 (80.05)	98 (69.01)	
Yes	129 (22.71)	85 (19.95)	44 (30.99)	
Asthma				0.168
No	562 (98.94)	423 (99.3)	139 (97.89)	
Yes	6 (1.06)	3 (0.7)	3 (2.11)	
Interstitial lung disease				0.375
No	561 (98.77)	422 (99.06)	139 (97.89)	
Yes	7 (1.23)	4 (0.94)	3 (2.11)	
Bronchiectasis				0.672
No	517 (91.02)	386 (90.61)	131 (92.25)	
Yes	51 (8.98)	40 (9.39)	11 (7.75)	
Pneumonia				< 0.001
No	479 (84.33)	373 (87.56)	106 (74.65)	
Yes	89 (15.67)	53 (12.44)	36 (25.35)	
RDW-CV	13.5 [12.9, 14.6]	13.5 [12.9, 14.6]	13.5 [12.9, 14.4]	0.584
RDW-SD	45.9 [42.8, 49.8]	45.9 [42.9, 49.77]	45.85 [42.5, 49.72]	0.920
White blood cell count	7.8 [6.02, 10.36]	7.69 [5.72, 10.16]	8.43 [6.7, 11.39]	0.026
Large platelet ratio	35 [28.67, 43.2]	35.2 [28.8, 43.1]	34.6 [28.63, 43.63]	0.917
Percentage of monocytes	7.3 [4.88, 9.6]	7.3 [4.82, 9.47]	7.35 [5.03, 9.95]	0.881
Monocyte absolute value	0.54 [0.39, 0.75]	0.53 [0.37, 0.74]	0.58 [0.42, 0.8]	0.076
Erythrocyte volume	41.2 [36.8, 45.42]	41.15 [36.8, 45.4]	41.25 [36.8, 45.75]	0.742
Red blood cell count	4.46 [3.99, 4.94]	4.46 [4, 4.94]	4.43 [3.95, 4.94]	0.845
Percentage of lymphocytes	11.3 [5.88, 18.05]	11.6 [6.3, 17.85]	10.6 [4.9, 18.58]	0.310
Lymphocyte absolute value	0.86 [0.5, 1.26]	0.86 [0.52, 1.22]	0.86 [0.46, 1.42]	0.865
Mean erythrocyte volume	92.45 [88.9, 95.93]	92.25 [88.82, 96.18]	93.35 [89.3, 95.75]	0.298
Mean erythrocyte hemoglobin volume	30.15 [28.8, 31.3]	30.2 [28.8, 31.6]	29.9 [28.8, 30.9]	0.057
Mean erythrocyte hemoglobin concentration	325 [315, 333]	326 [317, 335]	322 [314, 329]	< 0.001
Mean platelet volume	11.3 [10.5, 12.3]	11.3 [10.5, 12.3]	11.25 [10.5, 12.38]	0.975
Hemoglobin	134.06 ± 21.23	134.55 ± 21.46	132.58 ± 20.52	0.327
Platelet distribution width	13.9 [11.9, 16.2]	13.9 [11.9, 16.2]	13.85 [12.03, 16.53]	0.917
Platelet count	170 [127, 213.25]	163.5 [125, 211.75]	179.5 [141.5, 217]	0.030
Platelet accumulation	0.19 [0.15, 0.23]	0.19 [0.15, 0.23]	0.2 [0.17, 0.24]	0.006

(*Continued*)

https://doi.org/10.2147/JIR.S511092

Table 1 (Continued).

Variables	Total (n = 568)	Training Set (n = 426)	External Validation Set (n = 142)	p
Percentage of neutrophils	79.4 [70.7, 87.3]	79.35 [70.7, 87.25]	79.5 [70.4, 88.65]	0.615
Neutrophil absolute value	6.15 [4.39, 8.7]	6.12 [4.22, 8.59]	6.43 [5, 9.04]	0.065
D-dimer	0.62 [0.35, 1.22]	0.6 [0.35, 1.21]	0.72 [0.42, 1.3]	0.119
APTT	27.7 [25.7, 30.22]	27.8 [25.63, 30.5]	27.6 [25.8, 29.08]	0.160
Antithrombin III	76.72 ± 14.26	75.68 ± 14.38	79.82 ± 13.47	0.002
TT	15.8 [14.97, 17.1]	15.7 [14.93, 16.9]	16.4 [15, 17.48]	0.019
PT	11.8 [11.1, 12.6]	11.7 [11, 12.6]	12 [11.2, 12.67]	0.138
FDP	2.9 [1.9, 4.73]	2.9 [1.9, 4.8]	2.6 [1.8, 4.45]	0.318
Fibrinogen	3.53 [2.7, 4.67]	3.51 [2.65, 4.52]	3.89 [2.77, 5.07]	0.046
γ- glutamyltransferase	23 [16.92, 37]	23 [16.27, 37]	24 [17, 38.85]	0.714
Albumin	36.33 ± 4.6	36.35 ± 4.56	36.28 ± 4.71	0.872
Alanine aminotransferase	25 [17, 36]	24 [15.85, 36]	28 [21, 38]	< 0.001
Cholinesterase	5000 [4386.5, 6170]	5000 [4365.5, 6119.6]	5317.1 [4488.25, 6343.48]	0.051
Calcium	2.18 ± 0.12	2.19 ± 0.12	2.18 ± 0.13	0.792
Triglyceride	0.96 [0.74, 1.27]	0.92 [0.72, 1.25]	1.02 [0.82, 1.38]	0.024
Creatinine	64.3 [51.95, 79.6]	62.95 [51.15, 77.68]	66.7 [54.35, 83.3]	0.041
CK	39 [30, 66.1]	37.5 [30, 62]	47.5 [34, 79.75]	< 0.001
CK-MB	11.45 [9.93, 16]	11 [9, 16]	13 [10, 17]	0.021
Alkaline phosphatase	77 [65, 93.93]	76.9 [66, 93.25]	78.25 [64.17, 94.3]	0.973
Phosphorus	1.2 [1.05, 1.38]	1.2 [1.06, 1.38]	1.19 [1.03, 1.39]	0.685
Urea	6.98 [5.25, 8.93]	7.06 [5.27, 8.89]	6.72 [5.11, 9.3]	0.840
Uric acid	309.6 [231.7, 396.8]	296.95 [226.55, 390.18]	326.85 [244.45, 418.02]	0.150
Globulin	28.4 [25.7, 31.5]	28.25 [25.6, 31.37]	29.25 [26.45, 31.78]	0.062
LDH	320.5 [208.75, 473.25]	356.5 [217.98, 490.5]	239.5 [199.5, 355.25]	< 0.001
Bicarbonate	35.85 [32.08, 39.2]	36.25 [32.7, 39.6]	33.9 [29.47, 37.98]	< 0.001
Aspartate aminotransferase	25 [20, 33.4]	25 [20, 33.77]	26 [20, 33.3]	0.452
Total cholesterol	4.11 [3.45, 4.9]	4.11 [3.4, 4.93]	4.09 [3.59, 4.85]	0.568
Total bilirubin	11.8 [8.42, 16.7]	12.13 [8.6, 16.9]	11.1 [8.2, 16.19]	0.298
Total protein	65.11 ± 6.89	64.96 ± 6.84	65.54 ± 7.05	0.396
Myoglobin	37.45 [27.3, 59.85]	38.7 [27.4, 60.15]	35.8 [27.1, 59.38]	0.612
N-telencephalic natriuretic peptide	433 [130, 2152.5]	481 [130, 2300]	319.5 [132.25, 1855]	0.275
PH value	7.37 [7.32, 7.41]	7.37 [7.32, 7.41]	7.37 [7.33, 7.4]	0.960
PaO2	48 [40, 54]	50 [41, 56]	44 [37, 50]	< 0.001
PaCO2	65 [56, 78]	65 [56, 78]	65 [56, 76.75]	0.378
SB	32.3 [29.4, 35.4]	32.4 [29.52, 35.7]	31.8 [28.9, 34.9]	0.355
Calcium _ vitality	1.16 [1.12, 1.19]	1.16 [1.13, 1.19]	1.15 [1.1, 1.18]	0.014
Methemoglobin	1.2 [1, 1.4]	1.2 [1, 1.4]	1.2 [1, 1.4]	0.843
Reduced hemoglobin	4 [2, 9.1]	3.7 [1.92, 7.88]	6.15 [2.55, 11.47]	< 0.001
Hematocrit	44 [38, 49]	44 [38, 48]	43 [39, 49]	0.955
Base surplus	9.9 [6, 13.72]	10.1 [6.2, 14.1]	9.35 [5.73, 13.3]	0.426
Lactic acid	1.5 [1.2, 1.9]	1.5 [1.2, 1.9]	1.6 [1.2, 2]	0.195
AB	37.9 [33.5, 42.82]	38.1 [33.6, 43.13]	37.65 [32.7, 42.48]	0.340
Carboxyhemoglobin	2.4 [1.9, 2.9]	2.4 [2, 2.9]	2.3 [1.8, 2.8]	0.205
Oxygenated hemoglobin	92.2 [87.35, 94.32]	92.6 [87.93, 94.4]	90.25 [84.52, 93.6]	0.001
Anion gap	4 [2, 7]	4 [2, 7]	4 [1, 7]	0.627
Total carbon dioxide	40.05 [35.2, 45.23]	40.15 [35.3, 45.4]	39.5 [34.4, 44.75]	0.315

Abbreviations: BMI, Body Mass Index; MMRC, modified British medical research council; VTE, Venous Thrombus Embolism; RDW-CV, Red Cell Distribution Width-variation coefficient; RDW-SD, Red Cell Distribution Width-standard deviation; APTT, Activated Partial Thromboplastin Time; TT, Thrombin Time; PT, Prothrombin Time; FDP, Fibrinogen Degradation Products; CK, Creatine Kinase; LDH, Lactate Dehydrogenase; PaO2, Arterial partial pressure of oxygen; PaCO2, Arterial partial pressure of carbon dioxide; SB, Standard bicarbonate.

Variable Screening and Identification of Independent Influencing Factors

In this study, we used the Boruta algorithm to screen 83 candidate variables with the aim of identifying key indicators that are closely related to the dependent variable. The results showed that nine potentially important variables were screened from the modeled dataset, including: cerebrovascular disease, white blood cell count, hematocrit, D-dimer, activated partial thromboplastin time, fibrin degradation products, carbon dioxide partial pressure, reduced hemoglobin and oxyhemoglobin, as shown in Figure 2.

In order to conduct an initial exploration of the factors influencing prolonged hospitalization in COPD patients with HRF, we employed multivariate logistic regression analysis based on the 9 indicators selected by the Boruta algorithm mentioned above. The findings revealed that cerebrovascular diseases, erythrocyte specific volume, activated partial thromboplastin time, partial pressure of carbon dioxide, reduced hemoglobin and oxygenated hemoglobin were identified as independent factors significantly impacting the excessively long duration of hospital stay among COPD patients combined with HRF, as presented in Table 2.

Establishment and Evaluation of 10 Machine Learning Models

In this study, we used 10 algorithms, including RF, Catboost, LightGBM, XGBoost, GBM, NNET, SVM, KNN, NaiveBeyes, and Logistic regression to train the model, and used 8 evaluation indicators to compare the evaluation effect of each model on the modeling dataset and the external validation set (Table 3). The results indicated that the Catboost model demonstrated superior performance in both the modeling dataset and the external validation set.

Then, in this study, we compared the performance of 10 different prediction models using the area under the receiver operating characteristic curve (AUC) as an evaluation metric. The Catboost model exhibited the highest AUC values on both the modeled data set and the external validation set, as illustrated in Figure 3. The analysis of calibration curves further confirmed the superiority of the Catboost model, particularly in the external validation set. The calibration curves of the Catboost model demonstrated a significantly higher level of prediction accuracy compared to other models, as clearly

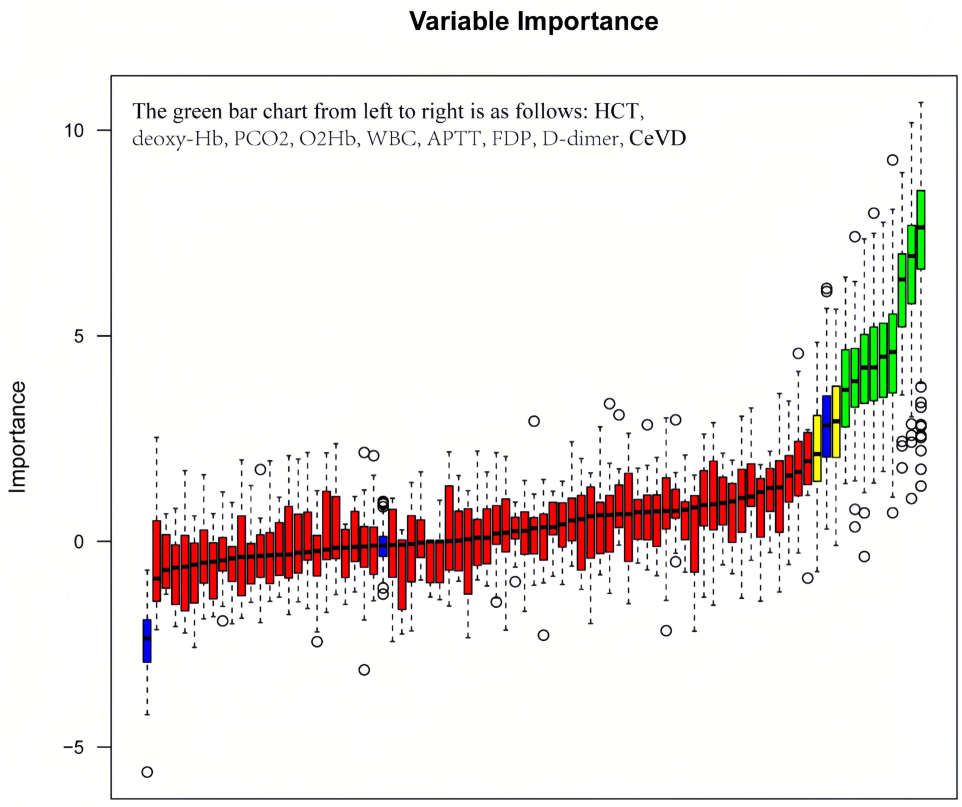

Figure 2 The process of filtering variables using the Boruta method.

https://doi.org/10.2147/JIR.S511092

Table 2 Multivariate Logistic Regression Results for Predicting the Risk of Prolonged Hospital Stay in COPD Patients with Hypercapnia Respiratory Failure

Variables	OR	95% CI	P	SE	Wald	B
CEVD						
No	Reference					
Yes	2.994	1.521–5.892	0.002	0.346	10.070	1.096
WBC	1.027	0.980–1.076	0.272	0.024	1.208	0.026
HCT	0.954	0.924–0.986	0.005	0.016	8.060	−0.047
D-dimer	0.946	0.710–1.260	0.704	0.146	0.144	−0.056
APTT	1.078	1.022–1.137	0.006	0.027	7.566	0.075
FDP	1.087	0.964–1.227	0.172	0.061	1.861	0.084
PCO2	1.018	1.006–1.031	0.004	0.006	8.237	0.018
Deoxy-Hb	1.381	1.042–1.831	0.025	0.144	5.038	0.323
O2Hb	1.381	1.042–1.832	0.025	0.144	5.035	0.323

Abbreviations: CEVD, Cerebrovascular disease; WBC, White Blood Cells; HCT, Hematocrit; APTT, Activated Partial Thromboplastin Time; FDP, Fibrinogen Degradation Products; PCO2, Partial Pressure of Carbon Dioxide; Hb, Hemoglobin.

depicted in Figure 4. Decision curve analysis (DCA) demonstrated that each model exhibited excellent clinical applicability in both the modeling data set and external validation set, with the specific DCA results presented in Figure 5.

The Catboost model also showed strong performance in terms of positive predictions in the evaluation of precision rate-recall curve (PR curve), as depicted in Figure 6. Taken together, the Catboost model was considered to be the best model due to its performance on several key performance indicators. To further demonstrate this, we also provided the Catboost model's confusion matrix on the modeling dataset and the external validation set, as shown in Figure 7. This matrix provided a comprehensive breakdown of the model's accurate predictions and misclassifications across different categories, thereby offering an insightful assessment of its predictive capabilities.

SHAP Interpretation of the Catboost Model

In order to better understand the influence of the Catboost model on the variable in the analysis of the factor of superaverage length of stay, we used SHAP (SHapley Additive exPlanations) values to explain the prediction results of the model.

First, we generated a SHAP summary plot that summarizes the impact of all variables on the model's predictions. As can be observed from Figure 8, the effects of variables on superaverage length of stay were oxyhemoglobin (O2Hb), activated partial thromboplastin time (APTT), D-dimer, reduced hemoglobin (deoxy-Hb), fibrin degradation products (FDP), white blood cell count (WBC), hematocrit (HCT), cerebrovascular disease (CeVD), and partial pressure of carbon dioxide (PCO2). These results pointed to the level of oxyhemoglobin as the most important factor in predicting superaverage length of stay, followed by coagulation system-related indicators such as APTT and D-dimer.

Further, we used SHAP waterfall plots to explain in detail the predictions of the model at the individual patient level. The waterfall plot details the risk probability of 0.259 for the first patient (Figure 9), which meaned that the patient has a lower risk of superaverage length of stay. The figure showed that a lower APTT is a factor that increases the probability of superaverage length of stay in this patient. On the contrary, D-dimer, O2Hb, and deoxy-Hb were the protective factors that reduced the probability of superaverage length of stay in this patient, indicating that the higher values of these indicators may help mitigate the risk of prolonged hospitalization.

Webpage Calculator for Catboost Model

In order to provide a practical tool to predict the risk of a patient's superaverage length of stay, we builted an interactive web-based calculator using the Shiny framework based on Catboost's prediction model. In the Shiny app, users can enter the patient's medical parameters such as oxyhemoglobin levels, D-Dimer values, etc., which are key variables in the

Table 3 Evaluation Results of 10 Models on Modeling Datasets and External Validation Sets

Dataset	Model	Accuracy	Precision	Recall	F1-Score	Sensitivity	Specificity	NPV	Briser_Score
Training	Catboost	0.993(0.985–1.001)	0.983(0.964–1.002)	1.000(1.000–1.000)	0.991(0.982–1.001)	1.000(1.000–1.000)	0.988(0.976–1.001)	1.000(1.000–1.000)	0.049(0.044–0.055)
Training	GBM	0.737(0.694–0.780)	0.656(0.590–0.724)	0.718(0.646–0.788)	0.685(0.629–0.743)	0.718(0.646–0.788)	0.750(0.699–0.803)	0.800(0.747–0.851)	0.182(0.167–0.196)
Training	KNN	0.692(0.648–0.735)	0.634(0.554–0.714)	0.541(0.467–0.614)	0.584(0.520–0.649)	0.541(0.467–0.614)	0.793(0.742–0.842)	0.722(0.671–0.771)	0.212(0.196–0.229)
Training	LightGBM	0.735(0.694–0.776)	0.667(0.601–0.736)	0.671(0.600–0.741)	0.669(0.613–0.727)	0.671(0.600–0.741)	0.777(0.729–0.828)	0.780(0.730–0.830)	0.192(0.179–0.205)
Training	Logistic	0.695(0.652–0.739)	0.630(0.554–0.709)	0.571(0.499–0.643)	0.599(0.538–0.664)	0.571(0.499–0.643)	0.777(0.726–0.830)	0.732(0.680–0.782)	0.211(0.196–0.227)
Training	NNET	0.695(0.652–0.737)	0.635(0.560–0.711)	0.553(0.477–0.627)	0.591(0.529–0.655)	0.553(0.477–0.627)	0.789(0.740–0.839)	0.727(0.674–0.776)	0.193(0.176–0.211)
Training	NaiveBayes	0.671(0.627–0.716)	0.606(0.525–0.689)	0.506(0.432–0.584)	0.551(0.486–0.621)	0.506(0.432–0.584)	0.781(0.729–0.832)	0.704(0.652–0.756)	0.258(0.226–0.291)
Training	RF	0.866(0.835–0.897)	0.842(0.789–0.896)	0.818(0.760–0.873)	0.830(0.788–0.872)	0.818(0.760–0.873)	0.898(0.862–0.934)	0.881(0.842–0.918)	0.140(0.129–0.152)
Training	SVM	0.859(0.826–0.893)	0.904(0.857–0.953)	0.724(0.655–0.793)	0.804(0.755–0.855)	0.724(0.655–0.793)	0.949(0.923–0.975)	0.838(0.795–0.880)	0.222(0.212–0.232)
Training	XGBoost	0.756(0.716–0.796)	0.699(0.629–0.769)	0.682(0.617–0.752)	0.690(0.637–0.748)	0.682(0.617–0.752)	0.805(0.754–0.853)	0.792(0.745–0.840)	0.213(0.207–0.219)
External	Catboost	0.810(0.745–0.877)	0.745(0.630–0.867)	0.759(0.639–0.883)	0.752(0.662–0.852)	0.759(0.639–0.883)	0.841(0.762–0.921)	0.851(0.774–0.926)	0.155(0.120–0.187)
External	GBM	0.676(0.598–0.755)	0.569(0.443–0.701)	0.611(0.480–0.744)	0.589(0.484–0.703)	0.611(0.480–0.744)	0.716(0.621–0.812)	0.750(0.655–0.842)	0.205(0.176–0.235)
External	KNN	0.627(0.548–0.703)	0.519(0.334–0.713)	0.259(0.137–0.383)	0.346(0.212–0.489)	0.259(0.137–0.383)	0.852(0.782–0.923)	0.652(0.564–0.737)	0.228(0.195–0.260)
External	LightGBM	0.634(0.558–0.711)	0.520(0.387–0.663)	0.481(0.349–0.615)	0.500(0.388–0.622)	0.481(0.349–0.615)	0.727(0.638–0.821)	0.696(0.602–0.787)	0.215(0.189–0.242)
External	Logistic	0.683(0.603–0.761)	0.600(0.453–0.751)	0.500(0.363–0.633)	0.545(0.427–0.669)	0.500(0.363–0.633)	0.795(0.709–0.884)	0.722(0.629–0.809)	0.224(0.193–0.256)
External	NNET	0.634(0.556–0.713)	0.524(0.380–0.678)	0.407(0.278–0.537)	0.458(0.341–0.585)	0.407(0.278–0.537)	0.773(0.686–0.864)	0.680(0.588–0.769)	0.212(0.179–0.245)
External	NaiveBayes	0.606(0.528–0.686)	0.479(0.336–0.620)	0.426(0.296–0.554)	0.451(0.333–0.572)	0.426(0.296–0.554)	0.716(0.623–0.812)	0.670(0.579–0.766)	0.309(0.247–0.368)
External	RF	0.739(0.667–0.815)	0.667(0.535–0.807)	0.630(0.501–0.766)	0.648(0.543–0.765)	0.630(0.501–0.766)	0.807(0.721–0.895)	0.780(0.696–0.864)	0.187(0.160–0.214)
External	SVM	0.655(0.574–0.738)	0.551(0.409–0.687)	0.500(0.366–0.637)	0.524(0.410–0.644)	0.500(0.366–0.637)	0.750(0.653–0.845)	0.710(0.615–0.810)	0.233(0.212–0.253)
External	XGBoost	0.648(0.570–0.727)	0.542(0.402–0.690)	0.481(0.348–0.619)	0.510(0.394–0.637)	0.481(0.348–0.619)	0.750(0.661–0.840)	0.702(0.609–0.793)	0.230(0.219–0.241)

Note: Sort by the first letter of the model name. The 95% CI for this measure is in parentheses.

https://doi.org/10.2147/JIR.S511092

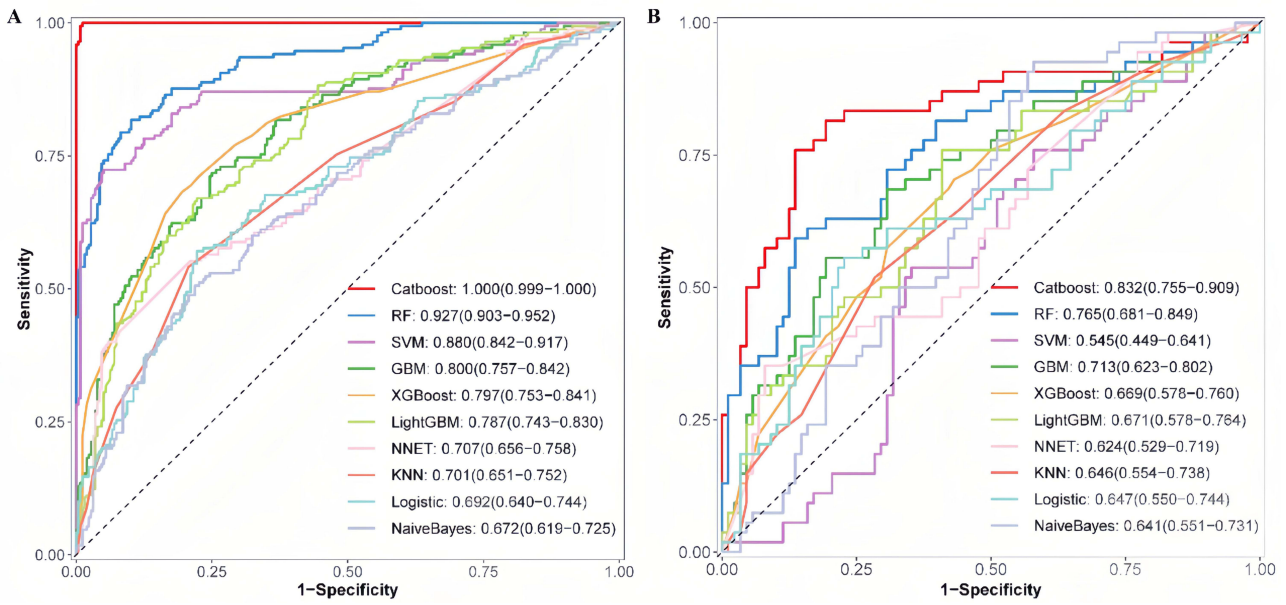

Figure 3 The ROC curve was used to evaluate the 10 models in modeling data set (**A**) and external validation set (**B**).

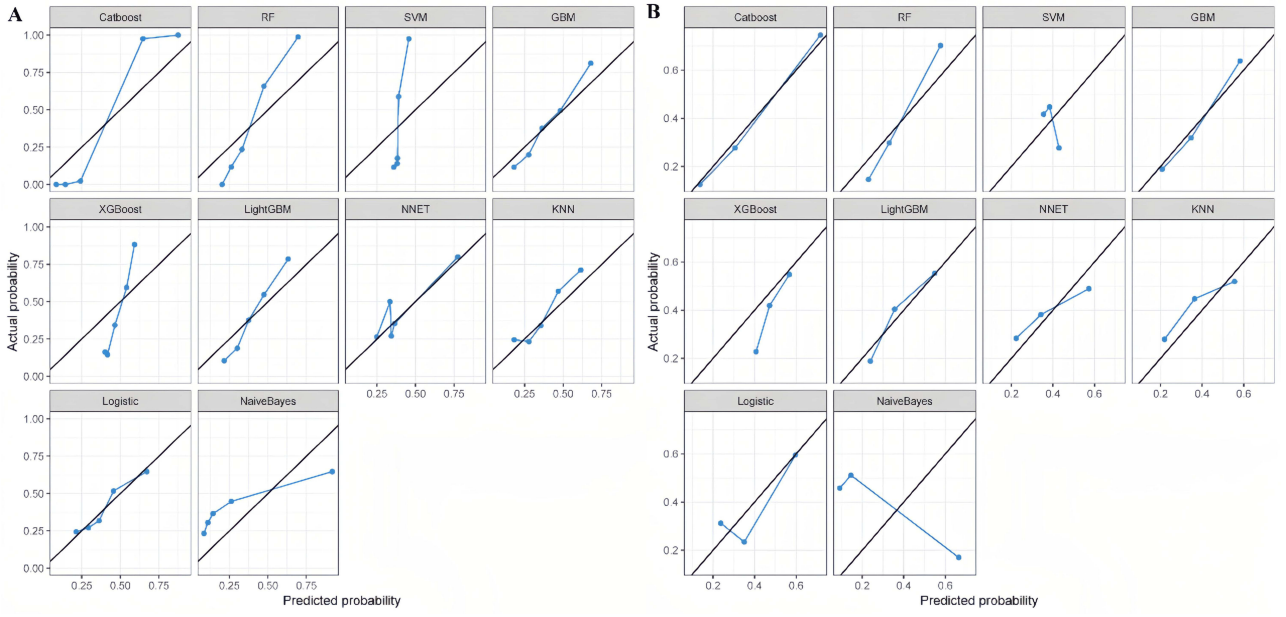

Figure 4 The Calibration curve was used to evaluate the 10 models in modeling data set (**A**) and external validation set (**B**).

model's decision-making process through a simple and intuitive interface. The application reads these inputs in real-time and uses a pre-trained Catboost model to calculate the patient's superaverage length of stay and displays the risk probability on a web page at: https://prolonged.shinyapps.io/Catboost-model/.

Discussion

As far as we know, this is the first study based on ML techniques to generate multiple models, evaluate performance, and select the highest-performing model to predict the superaverage length of stay in COPD patients combined with HRF. This study showed that the Catboost model exhibited superior performance and clinical utility compared to other

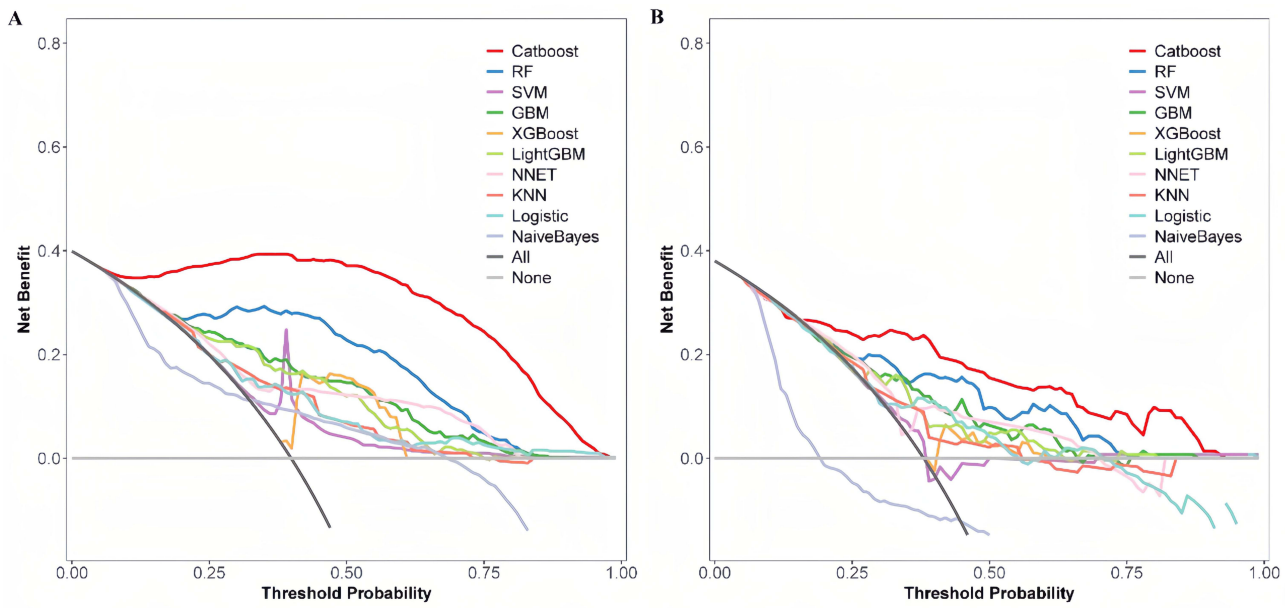

Figure 5 The DCA curve was used to evaluate the 10 models in modeling data set (**A**) and external validation set (**B**).

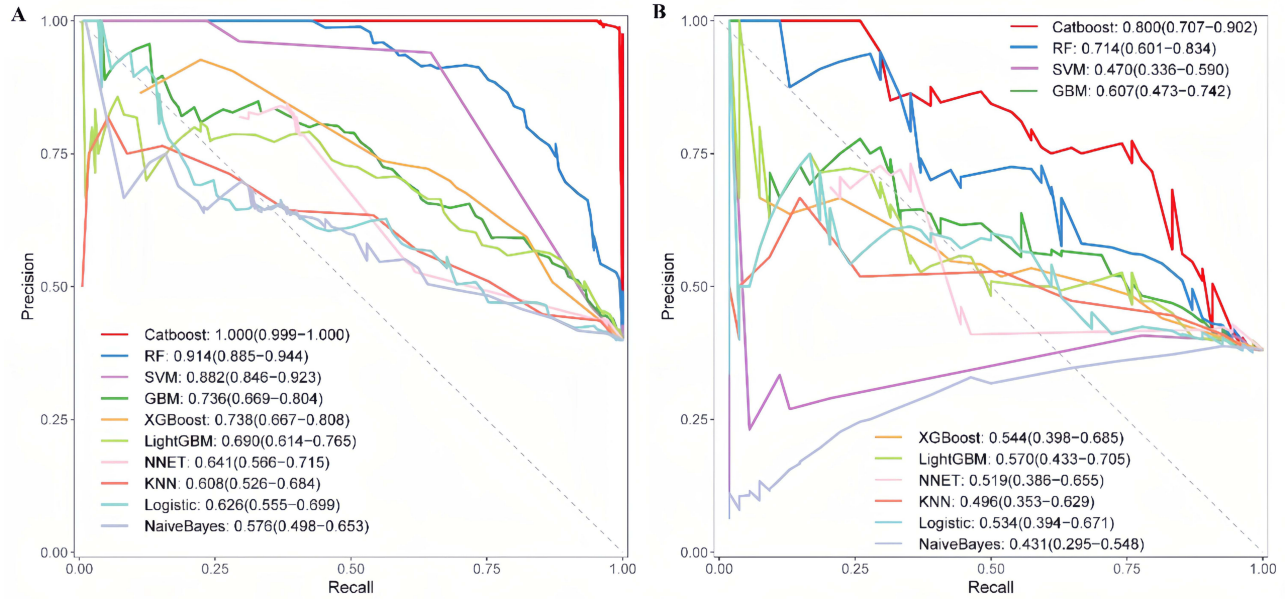

Figure 6 The PR curve was used to evaluate the 10 models in modeling data set (**A**) and external validation set (**B**).

independent machine learning models, including RF, LightGBM, XGBoost, GBM, NNET, SVM, KNN, NaiveBeyes, and Logistic regression models.

CatBoost is a novel machine learning algorithm that has the advantage of automatically processing categorical features and missing values during training, not during preprocessing.[20,21] As a result, in some previous studies, CatBoost has significantly outperformed other machine learning models in various data analysis.[22]

In this study, the CatBoost model was a machine learning technique that could prevent overfitting by using unbiased gradient estimation.[23] The CatBoost algorithm was chosen because our dataset contains many categorical variables (eg, gender, smoking status, comorbidities, etc.) and ensured generalization of the model by minimizing overfitting.[23]

https://doi.org/10.2147/JIR.S511092

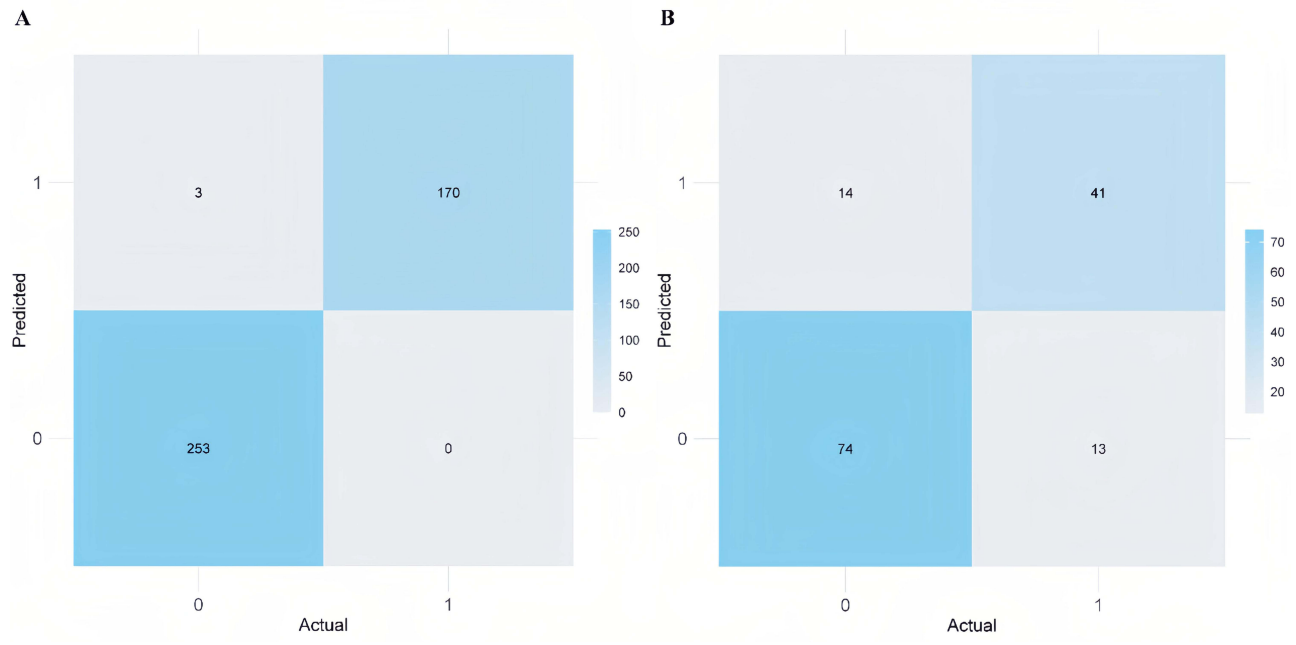

Figure 7 The Confusion matrix was used to evaluate the Catboost model in modeling data set (**A**) and external validation set (**B**).

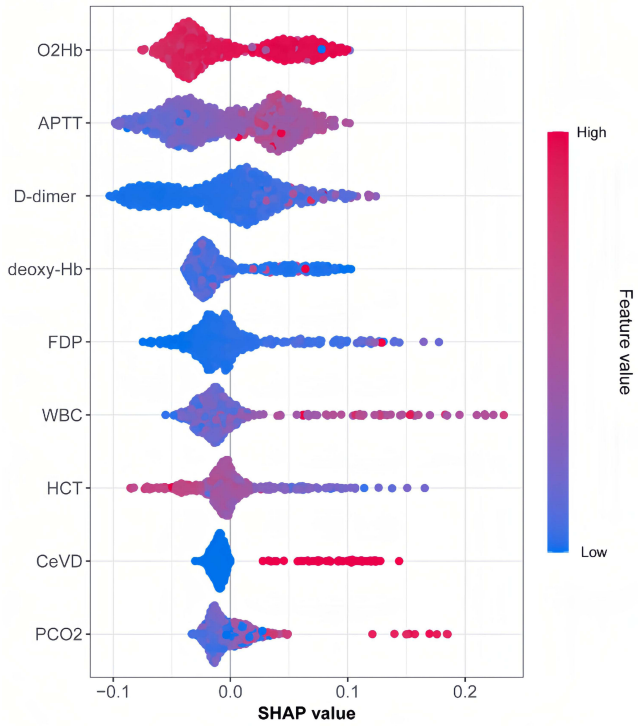

Figure 8 The SHAP summary diagram of Catboost model.

In addition, machine learning-based AI methods, such as CatBoost, tend to favor black-box models, which offer significant advantages over traditional models in obtaining accurate predictions, but cannot be reasonably explained.[24] Therefore, the SHAP method could be used to calculate the SHAP value of each feature in the prediction model to provide consistent and accurate attribution for each feature, and the interpretable analysis of the machine learning model was carried out.[25,26] The SHAP analysis

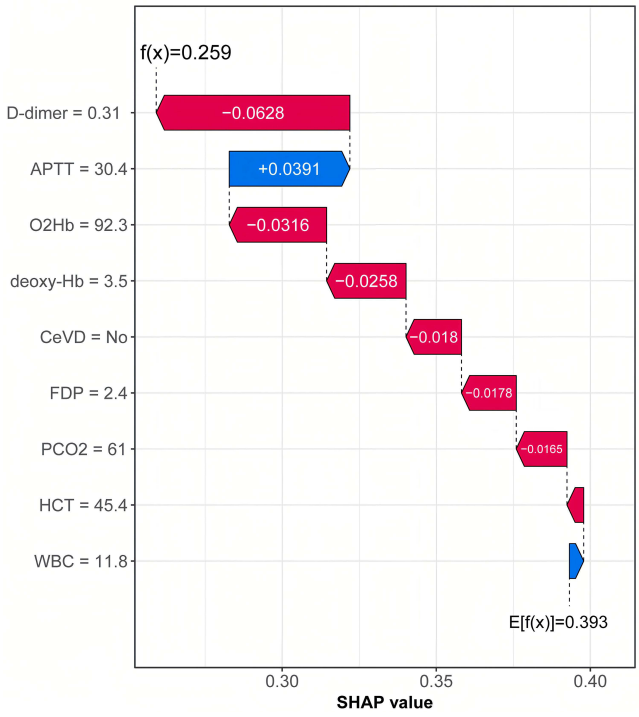

Figure 9 The SHAP waterfall diagram of the first patient in the Catboost model.

of a single sample enables the identification of high-risk samples and patients, thereby enhancing physicians' comprehension of the decision-making process employed by the CatBoost model and facilitating the utilization of predictive outcomes.

According to the results of the study, we found that cerebrovascular disease,White blood cell count, hematocrit, activated partial thromboplastin time, partial pressure of carbon dioxide, reduced hemoglobin and oxyhemoglobin were risk factors for the superaverage length of stay in COPD patients with HRF. Some results have been shown to affect the length of stay in other types of respiratory failure[27–30] or in patients with COPD,[31–33] but reduced hemoglobin and oxyhemoglobin have not been studied.

Based on the CatBoost algorithm, we found that oxyhemoglobin, APTT, and D-dimer were the top 3 strongest predictors of superaverage length of stay in HRF patients. There are no studies on the effect of these measures on the superaverage length of stay in COPD patients with HRF, but in one study it was found that patients with COPD often resulted in prolonged hospital stay and poor prognosis. The main reason is that high concentrations of blood carbon dioxide and acidosis can lead to malfunction of coagulation factors[34] and damage to blood endothelial cells.[35] Specifically, APTT may be significantly prolonged[36,37] and D-dimer levels may increase.[38] Prolongation of APTT may be caused by the depletion of coagulation factors,[39,40] while D-dimer is an indicator of thrombosis and is suggested as a prognostic biomarker of mortality in AECOPD.[41] These results suggest a correlation between coagulopathy and HRF, and that this relationship may lead to a longer hospital stay.[42,43] This is consistent with our findings.

Severe impairment of respiratory function is associated with oxyhemoglobin affinity (P50).[44] In particular, when hypoxia or hypercapnia is present, P50 has some effect.[45] Hypercapnia has been shown to inhibit arterial oxygen-carrying capacity.[46] Moreover, some studies have substantiated the correlation between oxyhemoglobin concentration and the long-term mortality risk associated with COPD.[47,48] Perhaps this is the reason for the extraordinarily long hospital stay in HRF patients. These findings can identify the risk of long hospital stay in HRF patients at an early stage, and enable targeted clinical care through timely intervention.

Of course, our study is also subject to some limitations: first, the patient cohort size is relatively small, although some interesting results have been found. However, the model still needs to be further validated in a large number of cohorts from multiple clinical centers. Second, future studies will be able to consider incorporating a wider range of clinical features to provide better outcomes. For example, other clinical information: patient medical images, ECG signals, lung function, etc.

https://doi.org/10.2147/JIR.S511092

In summary, a model for predicting the superaverage length of stay in COPD patients with HRF was developed and validated using clinical variables using 10 machine learnings. Among them, the Catboost model has the most advantages and can be used for clinical evaluation and patient monitoring. This model could predict the average length of hospital stay for HRF patients, assist clinicians in selecting appropriate treatment plans, and prevent the unnecessary expenditure of clinical resources.

Data Sharing Statement

All data generated or analysed during this study are included in this published article.

Ethics Approval and Consent to Participate

This study was approved by the ethics committees of the First People's Hospital of Yancheng (No.2020-K062) and the People's Hospital of Jiangsu Province (No. 2021-SR-346).

Consent for Publication

All authors approved the final manuscript and the submission to this journal.

Acknowledgments

Bingqing Zuo and Lin Jin are co-first authors for this work.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation. During the submission process, all authors were actively involved in drafting, revising, and providing final approval of the publication version of the article. All authors have unanimously agreed on the selection of the journal to which the article was submitted and have consented to assume responsibility for the content of the article.

Funding

This study was supported by the Yancheng key research and development plan guiding project (YCBE202344).

Disclosure

The authors declare that they have no competing interests in this work.

References

1. Comellini V, Pacilli AMG, Nava S. Benefits of non-invasive ventilation in acute hypercapnic respiratory failure. *Respirology*. 2019;24:308–317. doi:10.1111/resp.13469
2. Roussos C, Koutsoukou A. Respiratory failure. *Europ Resp J*. 2003;22(47 suppl):3s–14s. doi:10.1183/09031936.03.00038503
3. van der Leest S, Duiverman ML. High-intensity non-invasive ventilation in stable hypercapnic COPD: evidence of efficacy and practical advice. *Respirology*. 2019;24:318–328. doi:10.1111/resp.13450
4. Chung Y, Garden FL, Marks GB, Vedam H. Causes of hypercapnic respiratory failure and associated in-hospital mortality. *Respirology*. 2023;28:176–182. doi:10.1111/resp.14388
5. Luthi-Corridori G, Boesing M, Ottensarendt N, et al. Predictors of length of stay, mortality and rehospitalization in COPD patients: a Retrospective Cohort Study. *J Clin Med*. 2023;12. doi:10.3390/jcm12165322
6. Stefan MS, Shieh M-S, Pekow PS, et al. Epidemiology and outcomes of acute respiratory failure in the United States, 2001 to 2009: a national survey. *J Hospital Med*. 2013;8:76–82. doi:10.1002/jhm.2004
7. Chen WL, Chen C-M, Kung S-C, et al. The outcomes and prognostic factors of acute respiratory failure in the patients 90 years old and over. *Oncotarget*. 2018;9:7197–7203. doi:10.18632/oncotarget.24051
8. Hope AA, Adeoye O, Chuang EH, et al. Pre-hospital frailty and hospital outcomes in adults with acute respiratory failure requiring mechanical ventilation. *J Crit Care*. 2018;44:212–216. doi:10.1016/j.jcrc.2017.11.017
9. Roberts RR, Frutos PW, Ciavarella GG, et al. Distribution of variable vs fixed costs of hospital care. *JAMA*. 1999;281:644–649. doi:10.1001/jama.281.7.644
10. Faddy M, Graves N, Pettitt A. Modeling length of stay in hospital and other right skewed data: comparison of phase-type, gamma and log-normal distributions. *Value Health*. 2009;12:309–314. doi:10.1111/j.1524-4733.2008.00421.x

11. Dehouche N, Viravan S, Santawat U, et al. Hospital length of stay: a cross-specialty analysis and Beta-geometric model. *PLoS One*. 2023;18:e0288239. doi:10.1371/journal.pone.0288239
12. Naurzvai N, Mammadova A, Gursel G. Risk factors for prolonged intensive care unit stay in patients with hypercapnic respiratory failure. *Rev Recent Clinical Trials*. 2023;18:129–139. doi:10.2174/1574887118666230320163229
13. Tyler K, Stevenson D. Respiratory emergencies in geriatric patients. *Emerg Med Clin North Am*. 2016;34:39–49. doi:10.1016/j.emc.2015.08.003
14. Hou N, Li M, He L, et al. Predicting 30-days mortality for MIMIC-III patients with sepsis-3: a machine learning approach using XGboost. *J Transl Med*. 2020;18:462. doi:10.1186/s12967-020-02620-5
15. Du M, Haag DG, Lynch JW, Mittinty MN. Comparison of the tree-based machine learning algorithms to cox regression in predicting the survival of oral and pharyngeal cancers: analyses based on SEER database. *Cancers*. 2020;12:2802. doi:10.3390/cancers12102802
16. Vogelmeier CF, Criner GJ, Martínez FJ, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive lung disease 2017 report: GOLD executive summary. *Arch Bronconeumol*. 2017;53:128–149. doi:10.1016/j.arbres.2017.02.001
17. Zakynthinos S, Roussos C. Hypercapnic respiratory failure. *Respir Med*. 1993;87:409–411. doi:10.1016/0954-6111(93)90065-8
18. Kempker JA, Abril MK, Chen Y, et al. The epidemiology of respiratory failure in the United States 2002–2017: a serial cross-sectional study. *Crit Care Explor*. 2020;2:e0128. doi:10.1097/CCE.0000000000000128
19. Kwizera A, Nakibuuka J, Nakiyingi L, et al. Acute hypoxaemic respiratory failure in a low-income country: a prospective observational study of hospital prevalence and mortality. *BMJ Open Respir Res*. 2020;7. doi:10.1136/bmjresp-2020-000719
20. Hancock JT, Khoshgoftaar TM. CatBoost for big data: an interdisciplinary review. *J Big Data*. 2020;7(94). doi:10.1186/s40537-020-00369-8
21. Lyu Y, Wu H-M, Yan H-X, et al. Classification of coronary artery disease using radial artery pulse wave analysis via machine learning. *BMC Med Inf Decis Making*. 2024;24:256. doi:10.1186/s12911-024-02666-1
22. Zhao QY, Liu L-P, Luo J-C, et al. A machine-learning approach for dynamic prediction of sepsis-induced coagulopathy in critically ill patients with sepsis. *Front Med*. 2020;7:637434. doi:10.3389/fmed.2020.637434
23. Kong SH, Ahn D, Kim B, et al. A novel fracture prediction model using machine learning in a community-based cohort. *JBMR Plus*. 2020;4:e10337. doi:10.1002/jbm4.10337
24. Yang Y, Yuan Y, Han Z, Liu G. Interpretability analysis for thermal sensation machine learning models: an exploration based on the SHAP approach. *Indoor Air*. 2022;32:e12984. doi:10.1111/ina.12984
25. Luo H, Xiang C, Zeng L, et al. SHAP based predictive modeling for 1 year all-cause readmission risk in elderly heart failure patients: feature selection and model interpretation. *Sci Rep*. 2024;14:17728. doi:10.1038/s41598-024-67844-7
26. Scavuzzo CM, Scavuzzo JM, Campero MN, et al. Feature importance: opening a soil-transmitted helminth machine learning model via SHAP. *Infect Dis Model*. 2022;7:262–276. doi:10.1016/j.idm.2022.01.004
27. Hadaya J, Verma A, Sanaiha Y, et al. Machine learning-based modeling of acute respiratory failure following emergency general surgery operations. *PLoS One*. 2022;17:e0267733. doi:10.1371/journal.pone.0267733
28. Fimognari FL, De Vincentis A, Arone A, et al. Clinical outcomes and phenotypes of respiratory failure in older subjects admitted to an acute care geriatric hospital ward. *Int Emerg Med*. 2024;19:1359–1367. doi:10.1007/s11739-024-03625-4
29. Nowiński A, Kamiński D, Korzybski D, Stokłosa A, Górecka D. The impact of comorbidities on the length of hospital treatment in patients with chronic obstructive pulmonary disease. *Pneumonologia i alergologia polska*. 2011;79:388–396.
30. Madani Y, Saigal A, Sunny J, Morris L, Johns RH. Characterization of chronic obstructive pulmonary disease patients with a long length of stay: a retrospective observational cohort study. *Turk Thorac J*. 2017;18:119–124. doi:10.5152/TurkThoracJ.2017.17026
31. Wang Y, Stavem K, Dahl FA, Humerfelt S, Haugen T. Factors associated with a prolonged length of stay after acute exacerbation of chronic obstructive pulmonary disease (AECOPD). *Int J Chron Obstruct Pulmon Dis*. 2014;9:99–105. doi:10.2147/COPD.S51467
32. Hu G, Wu Y, Zhou Y, et al. Prognostic role of D-dimer for in-hospital and 1-year mortality in exacerbations of COPD. *Int J Chron Obstruct Pulmon Dis*. 2016;11:2729–2736. doi:10.2147/COPD.S112882
33. Ramaraju K, Kaza AM, Balasubramanian N, Chandrasekaran S. Predicting healthcare utilization by patients admitted for COPD exacerbation. *J Clin Diagn Res*. 2016;10:OC13–17. doi:10.7860/JCDR/2016/17721.7216
34. White H, Bird R, Sosnowski K, Jones M. An in vitro analysis of the effect of acidosis on coagulation in chronic disease states - a thromboelastograph study. *Clin Med*. 2016;16:230–234. doi:10.7861/clinmedicine.16-3-230
35. Malerba M, Nardin M, Radaeli A, et al. The potential role of endothelial dysfunction and platelet activation in the development of thrombotic risk in COPD patients. *Exp Rev Hematol*. 2017;10:821–832. doi:10.1080/17474086.2017.1353416
36. Husebø GR, Gabazza EC, D'alessandro gabazza C, et al. Coagulation markers as predictors for clinical events in COPD. *Respirology*. 2021;26:342–351. doi:10.1111/resp.13971
37. Ashoor TM, Hasseb AM, Esmat IM. Nebulized heparin and salbutamol versus salbutamol alone in acute exacerbations of chronic obstructive pulmonary disease requiring mechanical ventilation: a double-blind randomized controlled trial. *Korean J Anesthesiol*. 2020;73:509–517. doi:10.4097/kja.19418
38. Fruchter O, Yigla M, Kramer MR. D-dimer as a prognostic biomarker for mortality in chronic obstructive pulmonary disease exacerbation. *Am J Med Sci*. 2015;349:29–35. doi:10.1097/maj.0000000000000332
39. Tang N, Li D, Wang X, Sun Z. Abnormal coagulation parameters are associated with poor prognosis in patients with novel coronavirus pneumonia. *J Thromb Haemost*. 2020;18:844–847. doi:10.1111/jth.14768
40. Zhu J, Pang J, Ji P, et al. Coagulation dysfunction is associated with severity of COVID-19: a meta-analysis. *J med virol*. 2021;93:962–972. doi:10.1002/jmv.26336
41. Pang H, Wang L, Liu J, et al. The prevalence and risk factors of venous thromboembolism in hospitalized patients with acute exacerbation of chronic obstructive pulmonary disease. *Clin Respir J*. 2018;12:2573–2580. doi:10.1111/crj.12959
42. Liu M, Hu R, Jiang X, Mei X. Coagulation dysfunction in patients with AECOPD and its relation to infection and hypercapnia. *J Clin Lab Analysis*. 2021;35:e23733. doi:10.1002/jcla.23733
43. Zheng LL, Wang S, Li Z-G, et al. Correlation of coagulation dysfunction with infection and hypercapnia in acute exacerbation of COPD patients. *Infect Drug Resist*. 2023;16:5387–5394. doi:10.2147/idr.s421925
44. Bergamaschi G, Barteselli C, Del Rio V, et al. Impaired respiratory function reduces haemoglobin oxygen affinity in COVID-19. *Br J Haematol*. 2023;200:e44–e47. doi:10.1111/bjh.18620

45. Neukirch F, Verdier F, Chastang C, et al. Changes of P50 in hypoxaemia, hypercapnia and polycythaemia: multivariate analysis. *Biomedicine*. 1977;26:409–415.
46. Torbati D, Totapally BR, Camacho MT, Wolfsdorf J. Experimental critical care in ventilated rats: effect of hypercapnia on arterial oxygen-carrying capacity. *J Crit Care*. 1999;14:191–197. doi:10.1016/s0883-9441(99)90034-5
47. Trudzinski FC, Jörres RA, Alter P, et al. Associations of oxygenated hemoglobin with disease burden and prognosis in stable COPD: results from COSYCONET. *Sci Rep*. 2020;10:10544. doi:10.1038/s41598-020-67197-x
48. Hinke CF, Jörres RA, Alter P, et al. Prognostic value of oxygenated hemoglobin assessed during acute exacerbations of chronic pulmonary disease. *Respiration*. 2021;100:387–394. doi:10.1159/000513440

Journal of Inflammation Research

Dovepress
Taylor & Francis Group

Publish your work in this journal

The Journal of Inflammation Research is an international, peer-reviewed open-access journal that welcomes laboratory and clinical findings on the molecular basis, cell biology and pharmacology of inflammation including original research, reviews, symposium reports, hypothesis formation and commentaries on: acute/chronic inflammation; mediators of inflammation; cellular processes; molecular mechanisms; pharmacology and novel anti-inflammatory drugs; clinical conditions involving inflammation. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit http://www.dovepress.com/testimonials.php to read real quotes from published authors.

Submit your manuscript here: https://www.dovepress.com/journal-of-inflammation-research-journal