# Prediction of GOLD Stage in Patients Hospitalized with COPD Exacerbations Using Blood Neutrophils and Demographic Parameters as Risk Factors

**Jing Chen**
Chinese Academy of Science

**Yang Zhao**
Nanjing Medical University

**Yuan Qun**
Nanjing Medical University

**Guo Liquan** ( ✉ mangocj6@163.com )

**Xiong Daxi**
Chinese Academy of Sciences

---

**Research article**

---

# Abstract

**Background:** Patients hospitalized with chronic obstructive pulmonary disease (COPD) exacerbations are unable to complete the pulmonary function test reliably due to their poor health conditions. Creating an easy-to-use instrument to identify the Global Initiative for Chronic Obstructive Lung Disease (GOLD) stage will offer valuable information that assists clinicians to choose appropriate clinical care to decrease the mortality in these patients. The objective of this study was to develop a prediction model to identify the GOLD stage in the hospitalized exacerbation of chronic obstructive pulmonary disease (ECOPD) patients.

**Methods:** This study involved 155 patients hospitalized for ECOPD. All participants completed lung function tests and the collection of blood neutrophils and demographic parameters. Receiver operating characteristic (ROC) curve was plotted based on the data of 155 patients, and was used to analyze the disease severity predictive ability of blood neutrophils and demographic parameters. A support vector regression (SVR) based GOLD stage prediction model was built using the training data set (75%), whose accuracy was then verified by the testing data set (25%).

**Results:** The percentage of blood neutrophils (denoted as NEU%) combined with the demographic parameters was associated with a higher risk to severe episode of ECOPD. The area under the ROC curve was 0.84. The SVR model managed to predict the GOLD stage with an accuracy of 90.24%. The root-mean-square error (RMSE) of the forced expiratory volume in one second as the percentage of the predicted value (denoted as $FEV_1\%pred$) was 8.84%.

**Conclusions:** The NEU% and demographic parameters are associated with the pulmonary function of the hospitalized ECOPD patients. The established prediction model could assist clinicians in diagnosing GOLD stage and planning appropriate clinical care.

# Background

Chronic obstructive pulmonary disease (COPD) is a progressive lung condition and a leading cause of adult morbidity and mortality worldwide [1–3]. Exacerbation of chronic obstructive pulmonary disease (ECOPD) is an event characterized by a sustained worsening of the respiratory symptoms of a patient (including cough, phlegm production, and dyspnea), beyond normal day-to-day variations, which often necessitates additional therapies [4,5]. These episodes requiring hospitalization are associated with increased morbidity, mortality, and put enormous burden upon healthcare systems [6,7]. Inflammation is a key component in the pathogenesis of COPD [8]. It has previously been observed that COPD is not only associated with abnormal inflammatory response of the lung, but also with systemic inflammation, including systemic oxidative stress, activation of circulated immune cells and inflammatory cells, and the increased circulating levels of inflammatory cytokines [9]. It is generally considered that ECOPD reflects a flare-up of these underlying inflammatory processes [10], and is linked to a neutrophilic signature response [11].

A recent systematic literature review concluded that ECOPDs are extremely dangerous events. There is an urgent need to identify tolerable treatment guidelines and manage acute exacerbations in hospitalized ECOPD patients [12]. The Global Initiative for Chronic Obstructive Pulmonary Disease (GOLD) uses the ratio of forced expiratory volume in one second to forced vital capacity (denoted as $FEV_1/FVC$) as the diagnostic criteria for airflow obstruction, whose nominal value shall be smaller than 0.70; and classifies the airflow obstruction severity based on the value of forced expiratory volume in one second as percentage of predicted value (denoted as $FEV_1\%pred$) as shown in Table 1. GOLD stages 1–4 respectively represent mild, moderate, severe and very severe. Inadequate diagnosis of COPD and the lack of spirometric assessment can lead to inadequate treatment strategies, with health costs and risks for patients, leading to delays in diagnosing and treatment of the true cause of the symptoms [13–15]. Patients hospitalized with COPD exacerbations due to poor health status, are unable to complete the pulmonary function test reliably. As one longitudinal study indicated, 50% of pulmonary function test results are unacceptable [16].

Table 1
Classification of airflow obstruction severity according to GOLD

| In patients with $FEV_1/FVC < 0.70$ | | |
|---|---|---|
| GOLD 1 | Mild | $FEV_1\%pred \geq 80$ |
| GOLD 2 | Moderate | $50 \leq FEV_1\%pred < 80$ |
| GOLD 3 | Severe | $30 \leq FEV_1\%pred < 50$ |
| GOLD 4 | Very severe | $FEV_1\%pred < 30$ |

Many studies on evaluating the disease severity in COPD patients have been focused on the stable stage instead of exacerbating stage. Based on the needs of unified assessment criteria that can accurately reflect the pulmonary function of hospitalized ECOPD patients, we explored the predictive ability of the percentage of blood neutrophils (NEU%) and demographic parameters in GOLD stage and created a prediction model based on support vector regression (SVR) for predicting GOLD stage in Hospitalized ECOPD Patients [17].

# Methods

# Subjects selection

A total of 155 subjects (135 males and 20 females) were included in the study, all of whom were from the Respiratory Department of the Affiliated Suzhou Science and Technology Town Hospital of Nanjing Medical University in Suzhou, China. With the permission of the Ethics Committee of the Affiliated Suzhou Science and Technology Town Hospital of Nanjing Medical University, the study was conducted

between January 2018 and March 2020. 241 candidates were collected from initially selected patients hospitalized for ECOPD. These subjects were over 40 years old, clinically diagnosed as COPD, either with aggravating symptoms or with no history of pulmonary dysfunction. A total of 86 patients were excluded due to the following exclusion criteria: (1) 34 patients with noninfectious exacerbations; (2) 30 patients withdrew consent; (3) 22 patients with mechanical barrier or hearing disease. Ultimately, 155 patients were enrolled (Fig. 1).

# Clinical Data

Demographic information including gender, age, height and weight was recorded upon admission to the hospital. After interview and signing a written consent, 155 patients participated in the pulmonary function tests using equipment manufactured by CareFusion, USA. The tests were guided by the same professional doctor when the patients' health status allowed to do so. Three effective pulmonary function tests were performed with the same, regularly gauged spirometer to reduce measurement errors. The average $FEV_1$%pred value and the GOLD stage of the three tests were recorded. $FEV_1$ was reported in litres; and the Z-score to correctly characterize the cohorts was investigated. The $FEV_1$ Z-score of each patient was derived by using norms from the Global Lung Initiative (GLI) based specially on developed software [18]. Peripheral venous blood was drawn for peripheral blood examination at the same time period with pulmonary function tests. The automatic blood analyzer (SYSMEX Japan) was used to calculate the NEU%.

# Risk Stratification

As $FEV_1$ < 1L indicates that the impact of the disease on patients' daily life and health is increasing rapidly, most physicians, technologists and radiologists adopt 1L as a threshold for $FEV_1$ to evaluate pulmonary function [19,20]. All enrolled patients were assessed whether they have severe episodes of ECOPD using $FEV_1$ < 1L as the threshold, and were thus classified into four stages according to GOLD guidelines.

# Statistical analysis

The demographic information, blood NEU% and $FEV_1$ of all the participants were expressed as the mean (SD) for normally distributed data or median (IQR) for non-normally distributed data, and the percentages for the count data. Differences of continuous variables between GOLD stages were assessed using Student's t-tests or Mann-Whitney U tests. ANOVA or Kruskal-Wallis test was used to compare more than two groups. Univariate logistic regression models were developed to assess the correlation between NEU%, demographic parameters and severe episode of ECOPD. All variables correlated with the severe episode of ECOPD were considered in a multivariable model. Receiver operating characteristic (ROC)

curves were constructed to evaluate the discrimination of models. An area under the ROC curve of 0.8 or greater is generally considered to be a good predictor [21].

The outcome variables were defined as the $FEV_1$%pred value and the GOLD stage of patients hospitalized for ECOPD. The patients were randomly divided into a training set (75%) and a testing set (25%). The training set was used to develop the SVR based prediction model; whereas the testing set was used to validate the predictive performance of $FEV_1$%pred value and GOLD stage. Pearson correlation coefficient was adopted to evaluate the linear correlation of the predicted and the measured values of $FEV_1$%pred. If the P-value is less than 0.05, the difference was considered significant. The root mean squared errors (RMSE) and correlation coefficient (r) were used to quantitatively describe the strength of the relationship between the predicted and the measured values of $FEV_1$%pred. The GOLD stage was classified by the predicted values of $FEV_1$%pred according to the GOLD guidelines, and compared with the GOLD stage classified by the measured value of $FEV_1$%pred. The GOLD stage prediction accuracy was also calculated to assess the discrimination ability of the model.

# Results

# Demographic information, blood NEU% and $FEV_1$ in all the enrolled patients

Ultimately, 155 subjects were enrolled. A total of 93 subjects were defined as non-severe episode pattern; and 62 subjects were defined as severe episode pattern. The height, weight, BMI, $FEV_1$, and $FEV_1$%Pred in the non-severe episode group were higher than those in the severe episode group, with significance ($p < 0.05$). Blood NEU% in the non-severe episode group was lower than that in the severe episode group, with significance ($p < 0.001$). There was no significant difference in the gender and age between the two groups (P = 0.587, P = 0.202) (Table 2).

Table 2
Comparison of parameters between non-severe episode group and severe episode group in all enrolled subjects

| Category | Variable | Units | FEV$_1$ < 1 L as the threshold of severe episode of ECOPD | | |
| --- | --- | --- | --- | --- | --- |
| | | | Non-severe episode (N = 93) | Severe episode (N = 62) | P value |
| Demographics | Male | N (%) | 82(88.17) | 53(85.48) | 0.587 |
| | Age | Years | 74.00(68.00, 79.00) | 76.50(69.00, 80.00) | 0.202 |
| | Height | cm | 165.50(162.00, 170.00) | 163.50(157.50, 168.00) | 0.018 |
| | Weight | kg | 66.57(9.94) | 59.79(10.31) | < 0.001 |
| | BMI | | 24.38(3.73) | 22.65(3.36) | 0.008 |
| Blood count | NEU | % | 72.05(65.08, 80.95) | 82.40(71.80, 86.40) | < 0.001 |
| Pulmonary function | FEV$_1$ | L | 1.37(1.15, 1.72) | 0.78(0.67, 0.86) | < 0.001 |
| | FEV$_1$ Z-score | | -2.29(-2.70, -1.83) | -3.53(-3.77, -3.18) | < 0.001 |
| | FEV$_1$%Pred | | 57.34(49.50, 67.95) | 33.46(28.83, 37.46) | < 0.001 |
| Data expressed as mean (SD) or median (IQR); N: numbers of patients | | | | | |

All enrolled patients were classified into GOLD stages 1–4 based on the GOLD. The factors associated with GOLD stage are shown in Table 3. Univariate analysis demonstrated that gender, age, weight, BMI and NEU % are the risk factors of different GOLD stages (Table 3). On the basis of the univariate analysis, the univariable and multivariable models were used to discriminate a severe episode of ECOPD.

Table 3
Characteristic distribution comparison among different COPD GOLD stage groups

| Characteristic | COPD GOLD stage | | | | P value |
|---|---|---|---|---|---|
| | 1 (N = 13) | 2 (N = 60) | 3 (N = 66) | 4 (N = 16) | |
| Gender | | | | | 0.032 |
| Male | 10(7.41%) | 50(37.04%) | 60(44.44%) | 15(11.11%) | |
| Female | 3(15%) | 11(55%) | 5(25%) | 1(5%) | |
| Age | | | | | 0.001 |
| < 60 | 2(15.38%) | 10(76.92%) | 1(7.69%) | 0(0%) | |
| 60−69 | 4(7.84%) | 24(47.06%) | 18(35.29%) | 5(9.80%) | |
| 70−79 | 5(8.62%) | 21(36.21%) | 28(48.28%) | 4(6.90%) | |
| ≥ 80 | 2(6.06%) | 6(18.18%) | 18(54.55%) | 7(21.21%) | |
| Height | | | | | 0.101 |
| < 155 | 2(16.67%) | 7(58.33%) | 3(25%) | 0(0%) | |
| 155−164 | 4(6.67%) | 21(35%) | 27(45%) | 8(13.33%) | |
| 165−174 | 7(9.46%) | 30(40.54%) | 30(40.54%) | 7(9.46%) | |
| ≥ 175 | 0(0%) | 3(33.33%) | 5(55.56%) | 1(11.11%) | |
| Weight | | | | | 0.035 |
| < 50 | 0(0%) | 5(26.32%) | 13(68.42%) | 1(5.26%) | |
| 50−59 | 1(3.13%) | 11(34.38%) | 19(59.38%) | 1(3.13%) | |
| 60−69 | 3(4.84%) | 28(45.16%) | 20(32.26%) | 11(17.74%) | |
| ≥ 70 | 9(21.43%) | 17(40.48%) | 13(30.95%) | 3(7.14%) | |
| BMI | | | | | 0.038 |
| < 18.5 | 0(0%) | 6(30%) | 13(65%) | 1(5%) | |
| 18.5−24.5 | 3(4%) | 28(37.33%) | 35(46.67%) | 9(12%) | |
| 24.5−30 | 9(17.31%) | 23(44.23%) | 14(26.92%) | 6(11.54%) | |
| > 30 | 1(12.50%) | 4(50%) | 3(37.50%) | 0(0%) | |
| NEU% | | | | | < 0.001 |
| < 70 | 9(15.25%) | 30(50.85%) | 19(32.20%) | 1(1.69%) | |
| N: numbers of patients | | | | | |

| Characteristic | COPD GOLD stage | | | | P value |
|---|---|---|---|---|---|
| | 1 (N = 13) | 2 (N = 60) | 3 (N = 66) | 4 (N = 16) | |
| 70−80 | 3(6.98%) | 18(41.86%) | 17(39.53%) | 5(11.63%) | |
| > 80 | 1(1.89%) | 13(24.53%) | 29(54.72%) | 10(18.87%) | |
| N: numbers of patients | | | | | |

# Discrimination Of A Severe Episode Of Ecopd

ROC plots and the areas under the ROC curves of the various models to discriminate a severe episode of ECOPD are shown in Fig. 2 and Table 4. A model incorporating NEU% had an area under the ROC curve of 0.68. The final model including all proved risk factors (Model 7) identified in the multivariable analysis showed good discrimination with an area under the ROC curve of 0.84. The area under the ROC curve of the final model without NEU% (Model 6) was 0.77, indicating that NEU% improved the discrimination accuracy as a useful biomarker.

Table 4
The areas under the ROC curves of univariable and multivariable models

| Model | The areas under the ROC (95% CI) | P value |
|---|---|---|
| Model 1 | 0.52 (0.41 to 0.62) | 0.760 |
| Model 2 | 0.57 (0.47 to 0.67) | 0.202 |
| Model 3 | 0.65 (0.56 to 0.75) | < 0.001 |
| Model 4 | 0.63 (0.53 to 0.73) | 0.013 |
| Model 5 | 0.68 (0.59 to 0.78) | < 0.001 |
| Model 6 | 0.77 (0.68 to 0.85) | < 0.001 |
| Model 7 | 0.84 (0.77 to 0.89) | < 0.001 |

# Gold Stage Prediction

The characteristics of the training set and testing set are included in Table 5. There was no significant difference in all the involved factors, i.e., demographics, blood count, pulmonary function and COPD GOLD stage of subjects between the two groups. Figure 3 and Fig. 4 show the predictive capability in

FEV$_1$%pred value of the SVR based prediction model. The association between the predicted and the measured FEV$_1$%pred value was strong with r = 0.92; and the difference was not significant (P > 0.05).

The total sample size of GOLD stage 1 was only 13, since most hospitalized ECOPD patients tended to have higher GOLD stage. As the degree of airflow limitation of patients with GOLD stage 1 and 2 is moderate, we combined GOLD stage 1 and GOLD stage 2 as the moderate group. The predictive performance on the FEV$_1$%pred value and GOLD stage are shown in Table 6. Figure 4 indicated that in the case of FEV$_1$%pred exceeding 70%, the model could bring pessimistic prediction results. Analysis on the GOLD stage predictive performance showed that, under the circumstance of predicted FEV$_1$%pred exceeding 70%, the algorithm would overestimate the GOLD stage. To be more specific, patients of GOLD stage 1 may be classified to GOLD stage 2. GOLD defines GOLD stage 1 and GOLD stage 2 as moderate airflow obstruction. Their treatment plan will not be confused with GOLD stage 3 and GOLD stage 4, which stand for severe airflow obstruction. Figure 5 shows the predictive performance in different GOLD stage. The overall COPD GOLD stage prediction accuracy was 90.24%.

Table 5
Characteristic of the training set and the testing set

| Category | Variable | Units | Training set (N = 114) | Testing set (N = 41) | P value |
|---|---|---|---|---|---|
| Demographics | Male | N (%) | 101(88.60) | 34(82.93) | 0.428 |
| | Age | years | 76.00(68.00, 80.00) | 71.00(69.00, 78.00) | 0.673 |
| | Weight | kg | 63.46(10.15) | 63.23(12.12) | 0.917 |
| | BMI | kg/m2 | 23.54(3.68) | 23.66(3.63) | 0.869 |
| Blood count | NEU | % | 76.50(66.60, 84.25) | 78.40(69.90, 84.30) | 0.385 |
| Pulmonary function | FEV$_1$%Pred | % | 46.59(34.57, 59.34) | 47.66(32.64, 59.30) | 0.822 |
| COPD GOLD Stage | 1 | N (%) | 9(7.89) | 4(9.76) | 0.603 |
| | 2 | N (%) | 49(42.98) | 13(31.72) | |
| | 3 | N (%) | 43(37.72) | 21(51.22) | |
| | 4 | N (%) | 13(11.40) | 3(7.32) | |
| N: numbers of patients | | | | | |

Table 6
Results of the prediction model in the testing sets

| Model evaluation | | RMSE of FEV$_1$%Pred (%) | Accuracy of COPD GOLD stage prediction (%) |
|---|---|---|---|
| Total population of test set (N = 41) | | 8.84 | 90.24 |
| COPD GOLD Stage | 1 and 2 (N = 17) | 12.51 | 94.11 |
| | 3 (N = 21) | 4.98 | 90.47 |
| | 4 (N = 3) | 2.83 | 66.67 |
| N: numbers of patients | | | |

# Discussion

ECOPD is a kind of acute attack process, where the patients' respiratory symptoms continue to worsen over their daily status. The frequent episodes of ECOPD resulted in an accelerated decline in FEV$_1$. Meanwhile, the rapid decline of FEV$_1$ performs as an independent hazard factor for ECOPD. The vicious circle between the decline of FEV$_1$ and the frequent attack of ECOPD affect the prognosis and mortality of the patients [22]. In this analysis, we focused on the discrimination value of blood NEU% as a biomarker for a severe episode of ECOPD, and the GOLD stage prediction in hospitalized ECOPD patients. We attempted to create an easy-to-use measure to estimate the value of FEV$_1$%Pred and to identify the GOLD stage that could assist clinicians in choosing appropriate measures of medical care to decrease future hospitalization rates and mortality in hospitalized ECOPD patients.

In line with previous studies, the outcome of pulmonary function test relied on the cooperation of ECOPD patients, most likely due to the limitation by force-velocity characteristics of expiratory muscles [16, 23]. Biomarkers were required for effective risk stratification and making individualized treatment decision.

The pathophysiological mechanism of most cases of ECOPD is an acute burst of local or systemic inflammatory mediators following respiratory bacterial or virus infection. Usually, high levels of non-specific inflammatory biomarkers are expected [24]. Neutrophils are the most abundant inflammatory cells in blood and sputum. As neutrophil proteases can generalize many of the characteristics of ECOPD including emphysema and mucus hypersecretion [25], ECOPD is characterized as a neutrophil inflammatory disorder in most cases. A study on peripheral blood neutrophils from ECOPD patients conducted by Milara et. al. showed that compared with healthy control group, the release of the neutrophil activation marker neutrophil elastase (NE) and reactive oxygen species (ROS) increased by 2 times and 30% respectively [26]. Jones et. al. observed that compared with the healthy controls, bacteria

stimulated neutrophil degranulation was greater in the ECOPD group [27]. Corhay et. al. focused on exacerbation whichever its trigger, and found that neutrophil inflammatory markers declined after treatment [28]. We designated a statistically significant difference in the NEU% between ECOPD patients with different GOLD stages to extend these findings. ECOPD patients with higher blood NEU% had a higher tendency of severe episode of ECOPD, whose GOLD stage risk stratification could thus be higher. The differences between ECOPD patients with different GOLD stages are consistent with the results of Perera et. al. They found that there were significant differences in systemic markers of inflammation between patients with GOLD stages 3 and 4 vs. controls without COPD; while there was no significant difference between GOLD 2 patients and controls [29].

We sought for factors that would discriminate a severe episode of ECOPD in clinical cases. Although the multivariable demographic parameters or NEU% values reflected the relative risk of a severe episode of ECOPD, considering the moderate values of areas under the ROC curves, the overall prediction performance is still quite limited. No matter which cut-off level is chosen, the false positive rate is still very high, so the specificity for acceptable value of sensitivity is low. With increase in blood NEU%, the risk of a severe episode of ECOPD increased. The overall discrimination value of multivariable factors including demographic parameters and blood NEU% was encouraging with the area under the ROC curve of 0.84.

To further study the $FEV_1$%Pred prediction and the GOLD stage categorization capability of the blood NEU% and demographic parameters, we randomly divided the data collected from the ECOPD patients into a training data set to develop a prediction model and a testing data set to validate the predictive performance. The selected demographic parameters included gender, age, weight and BMI, which had demonstrated their relevance to the target values. We used supervised learning algorithm to evaluate the predictive ability of the risk factors, and classified the subjects to 4 different GOLD stages. Searching for the right subjects was one of the major difficulties of our study.

On the other hand, support vector machine (SVM) is a learning method based on the principle of structural risk minimization of statistical learning theory. It shows many unique advantages in solving the problem of small sample and nonlinearity [30]. SVR is a model dealing with the SVM regression problems, which showed acceptable regression capacity in estimating the value of $FEV_1$%Pred and identifying the GOLD stage.

To our knowledge, this is the first study in ECOPD patients to predict the value of $FEV_1$%Pred and identify the GOLD stage based on demographic parameters and blood NEU%. In the absence of a clear biomarker to categorize the GOLD stage of ECOPD patients, our research provides an auxiliary guidance value for the clinicians to diagnose GOLD stage and establish appropriate clinical care, since the demographic parameters and blood NEU% are easy to be obtained.

Limitations of our current study should also be noted. First, the relatively small number of subjects enrolled in this study could limit the predictive performance of the model, especially when comparing to

the previous work of Cristóbal et. al. [31] and Godtfredsen and coworkers [32]. The predictive performance of the prediction model was limited in the ECOPD patients with optimistic degree of airflow obstruction, which could also be resulted from the lower influence of inflammatory factors when the symptoms were moderate. To find proper ECOPD patients and guide them to complete the pulmonary function test turned out to be one of the biggest difficulties during our research. To overcome this limitation, we used the most widely accepted learning method SVM to establish the prediction model. The grouping strategy of the training set and testing set was able to tackle the problem of multiple covariates larger than the samples (patients) or "p > n problem". Importantly, the overall ECOPD GOLD stage prediction accuracy of the establish prediction model was 90.24%. Secondly, Sørheim and coworkers showed that pulmonary function injury may differ between genders. There was a gender imbalance in our study, as the ECOPD patients included were mostly male (135/155). The model's predictive performance on female patients could be limited. Therefore, our future work is to balance the gender composition and extend the observation time to carry out larger scale research to verify our findings. As an additional limitation of the study, the patient's general condition, comprehension and cooperative degree could also influence the accuracy of pulmonary function test results. Nevertheless, every enrolled patient was trained and guided by the same professional physician to minimize the impact of external factors on the measurement.

## Conclusions

In summary, a prediction model based on demographic parameters and blood NEU% has been established to predict the value of $FEV_1$%Pred and identify the GOLD stage of the patients hospitalized with ECOPD. This easy-to-use instrument can assist clinicians in diagnosing GOLD stage, and offers valuable information to determine the appropriate clinical care for hospitalized ECOPD patients.

## Abbreviations

COPD

Chronic Obstructive Pulmonary Disease

GOLD

The Global Initiative for Chronic Obstructive Lung Disease

ECOPD

Exacerbation of Chronic Obstructive Pulmonary Disease

ROC

The Receiver Operating Characteristic

SVR

Support Vector Regression

**NEU%**

The Percentage of Neutrophils

**RMSE**

The Root-mean-square Error

**FEV$_1$/FVC**

The Ratio of Forced Expiratory Volume in One Second to Forced Vital Capacity

**FEV$_1$%pred**

The Forced Expiratory Volume in One Second as the Percentage of the Predicted

**GLI**

The Global Lung Initiative

**ROS**

Reactive Oxygen Species

**NE**

Neutrophil Elastase

**SVM**

Support Vector Machine

# Declarations

material collection, data analysis, decision to publish, or preparation of the manuscript.

## Authors' information

Affiliations

**School of Biomedical Engineering (Suzhou), Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei 230026, P. R. China**

Jing Chen, Li-quan Guo, Da-xi Xiong

**Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou 215163, P. R. China**

Jing Chen, Li-quan Guo, Da-xi Xiong

**Respiratory Department, The Affiliated Suzhou Science and Technology Town Hospital of Nanjing Medical University, Suzhou 215163, P. R. China**

Zhao Yang, Qun Yuan

Corresponding author:

Correspondence to Li-quan Guo.

## Authors' contributions

J.C. conceived and designed the study, acquired and analyzed the data, and wrote the manuscript. Z.Y. and Q.Y. contributed to data analysis. D.X.X. provided support for statistical analysis, interpretation, and critical revision of the manuscript for important intellectual content. L.Q.G. helped to conceptualize and design the study. All authors have read and approved the manuscript.

## Ethics approval and consent to participate

The study was approved by the Ethics Committee of the Affiliated Suzhou Science and Technology Town Hospital of Nanjing Medical University. All patients agree to participate and signed a consent form.

## 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.

# References

1 GOLD. Global strategy for the diagnosis, management and prevention of COPD, global initiative for chronic obstructive lung disease (GOLD);2019. [Update 2019] Available from: http://goldcopd.org. Accessed February 5, 2019

2 Adeloye D, Chua S, Lee C, Basquill C, Papana A, Theodoratou E, et al. Global and regional estimates of COPD prevalence: Systematic review and meta-analysis. J Glob Health 2015; 5:020415.doi: 10.7189/jogh.05-020415

3 Projections of mortality and causes of death, 2015 and 2030. Geneva: World Health Organization; Available from: http://www.who.int/healthinfo/global_burden_disease/projections/en/ (Accessed on September 11, 2017).

4 Burge S, Wedzicha JA. COPD exacerbations: definitions and classifications. Eur Respir J Suppl. 2003;41:46s–53s.

5 Rubinsztajn R, Przybylowski T, Maskey-Warzechowska M, Karwat K, Chazan R. Exacerbations of chronic obstructive pulmonary disease and quality of life of patients. Adv Exp Med Biol. 2016;884:69–74.

6 Donaldson GC, Seemungal TAR, Bhowmik A, Wedzicha JA. Relationship between exacerbation frequency and lung function decline in chronic obstructive pulmonary disease. Thorax. 2002;57:847–852.

7 Schmidt SAJ, Johansen MB, Olsen M, Xu X, Parker JM, Molfino NA, Lash TL, Sørensen HT, Christiansen CF. The impact of exacerbation frequency on mortality following acute exacerbations of COPD: a registry-based cohort study. BMJ Open. 2014;4:e006720.

8 Rhee C K , Kim D K . Role of phosphodiesterase-4 inhibitors in chronic obstructive pulmonary disease[J]. Korean J Intern Med, 2020, 35(2):276-283.

9 Agusti A G N , Noguera A , Sauleda J , et al. Systemic effects of chronic obstructive pulmonary disease[J]. Eur Respir J, 2003, 21(2):347-360.

10 Jing Z , Chun C , Ning S , et al. Systemic Inflammatory Marker CRP Was Better Predictor of Readmission for AECOPD Than Sputum Inflammatory Markers[J]. Arch Bronconeumol (English Edition), 2016.

11 Gao J , Chen B , Wu S , et al. Blood cell for the differentiation of airway inflammatory phenotypes in COPD exacerbations[J]. BMC Pulm Med, 2020, 20.

12 Angela, M, Abbatecola, et al. Practical management problems of chronic obstructive pulmonary disease in the elderly: acute exacerbations.[J]. Curr Opin Pulm Med, 2011.

13 Fernandez-Villar A, Soriano JB, Lopez-Campos JL. Overdiagnosis of COPD: precise definitions and proposals for improvement. Br J Gen Pract. 2017;67:183–4.

14 Hartl S, Lopez-Campos JL, Pozo-Rodriguez F, Castro-Acosta A, Studnicka M, Kaiser B, et al. Risk of death and readmission of hospital-admitted COPD exacerbations: European COPD Audit. Eur Respir J. 2016;47:113–21.

15 Wu H, Wise RA, Medinger AE. Do Patients Hospitalized With COPD Have Airflow Obstruction? Chest. 2017;151:1263–71.

16 Ulmer WT. Lung function--clinical importance, problems, and new results.[J]. J PHYSIOL PHARMACOL, 2003, 54 Suppl 1(2):11.

17 Vapnik V N. The nature of statistical learning theory [M]. New York: Springer, 2000: 138-167.

18 GLI-2012 desktop software for individual calculations. http://www.ers-education.org/guidelines/global-lung-function-initiative/spirometry-tools/desktop-individual-calculator.aspx. Last access Jun 3rd 2018.

19 Donner C F , Carone M , Bertolotti G , et al. Methods of assessment of quality of life[J]. Eur Respir Rev, 1997, 7(42):43-45.

20 Manhire A , Charig M , Clelland C , et al. Guidelines for radiologically guided lung biopsy[J]. Thorax, 2003, 58(11):920-36.

21 Berge M V D , Hop W C , Molen T V D , et al. Prediction and course of symptoms and lung function around an exacerbation in chronic obstructive pulmonary disease[J]. Respir Res, 2012, 13(1):44.

22 Celli BR, Thomas NE, Anderson JA, et al. Effect of Pharmacotherapy on Rate of Decline of Lung Function in Chronic Obstructive Pulmonary Disease [J]. Am J Respir Crit Care Med, 2008, 178(4):332-338.

23 Torun E , Cakir E , Özgüç, Fatma, et al. The Effect of Obesity Degree on Childhood Pulmonary Function Tests[J]. Balkan Med J, 2014, 31(3):235-238.

24 Jing C,Yang Z, Qun Y et al. Prediction models for pulmonary function during acute exacerbation of chronic obstructive pulmonary disease [J]. Physiol Meas, 2020, 41, 125010.

25 Lonergan, M., Dicker, A.J., Crichton, M.L. et al. Blood neutrophil counts are associated with exacerbation frequency and mortality in COPD[J]. Respir Res, 2020, 21: 166.

26 Milara J, Juan G, Peiro T, Serrano A, Cortijo J. Neutrophil activation in severe, early-onset COPD patients versus healthy non-smoker subjects in vitro: effects of antioxidant therapy. Respiration. 2012;83(2):147–58.

27 Jones AW, Robinson R, Mohamed P, Davison G, Izzat HJ, Lewis KE. Impaired blood neutrophil function in the frequent Exacerbator of chronic obstructive pulmonary disease: a proof-of-concept study. Lung. 2016;194(6):881–7.

28 Corhay J , Moermans C , Henket M , et al. Increased of exhaled breath condensate neutrophil chemotaxis in acute exacerbation of COPD[J]. Respir Res, 2014, 15(1):115.

29 Perera WR, Hurst JR, Wilkinson TM, et al. Inflammatory changes, recovery and recurrence at COPD exacerbation. Eur Respir J. 2007;29(3):527–34.

30 Vapnik V N. The nature of statistical learning theory [M]. New York: Springer, 2000: 138-167.

31 Cristóbal, Esteban, Inmaculada, et al. A decision tree to assess short-term mortality after an emergency department visit for an exacerbation of COPD: a cohort study[J]. Respir Res, 2015.

32 Godtfredsen N S , Jrgensen D V , Marsaa K , et al. Soluble urokinase plasminogen activator receptor predicts mortality in exacerbated COPD[J]. Respir Res, 2018, 19(1):97.
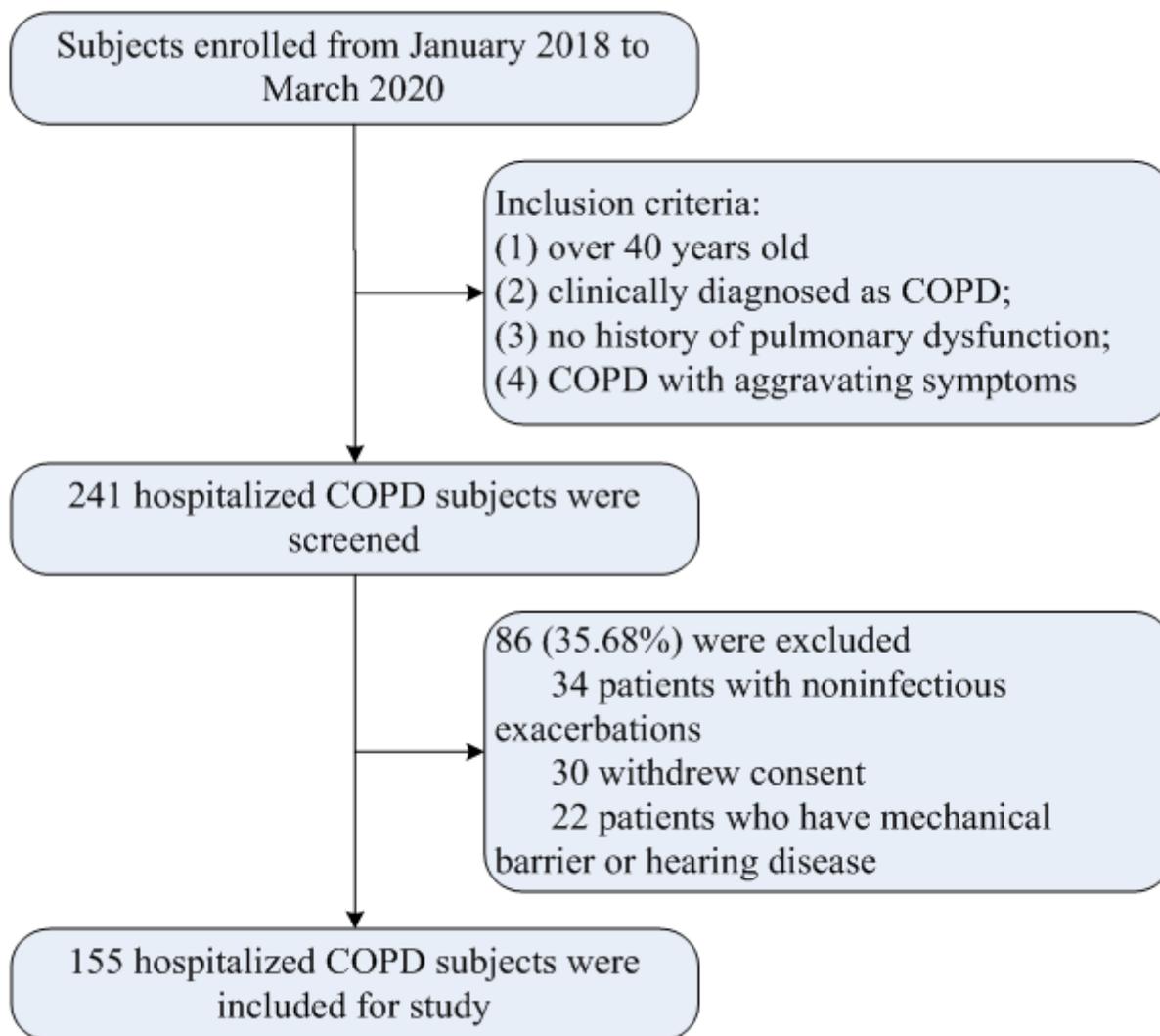
# Figures

**Figure 1**

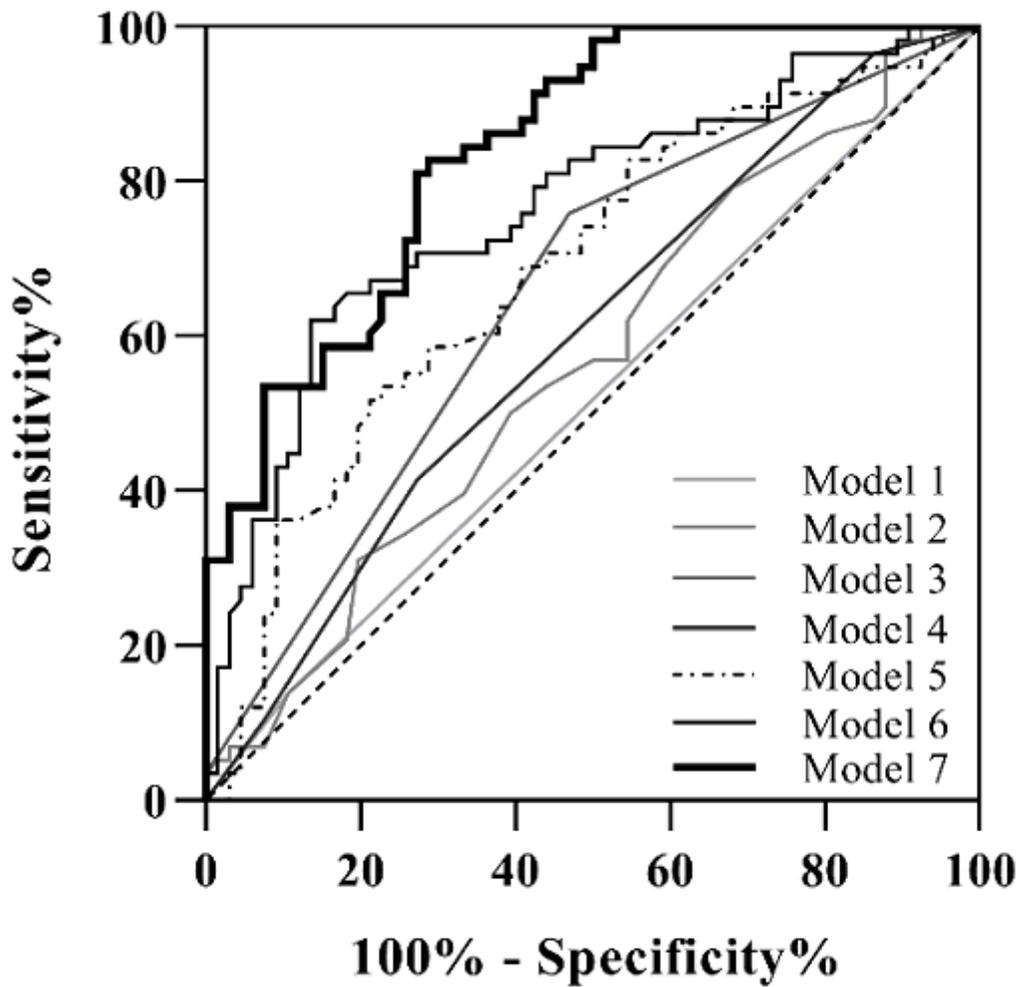Flowchart of inclusion and exclusion criteria.

**Figure 2**

ROC curves for predicting severe episode of ECOPD. The areas under the ROC demonstrate the ability of univariable and multivariable models to predict severe episode of ECOPD. Model 1=gender; Model 2=age; Model 3=weight; Model 4=BMI; Model 5=NEU %; Model 6= multivariable model without NEU%; Model 7= multivariable model.
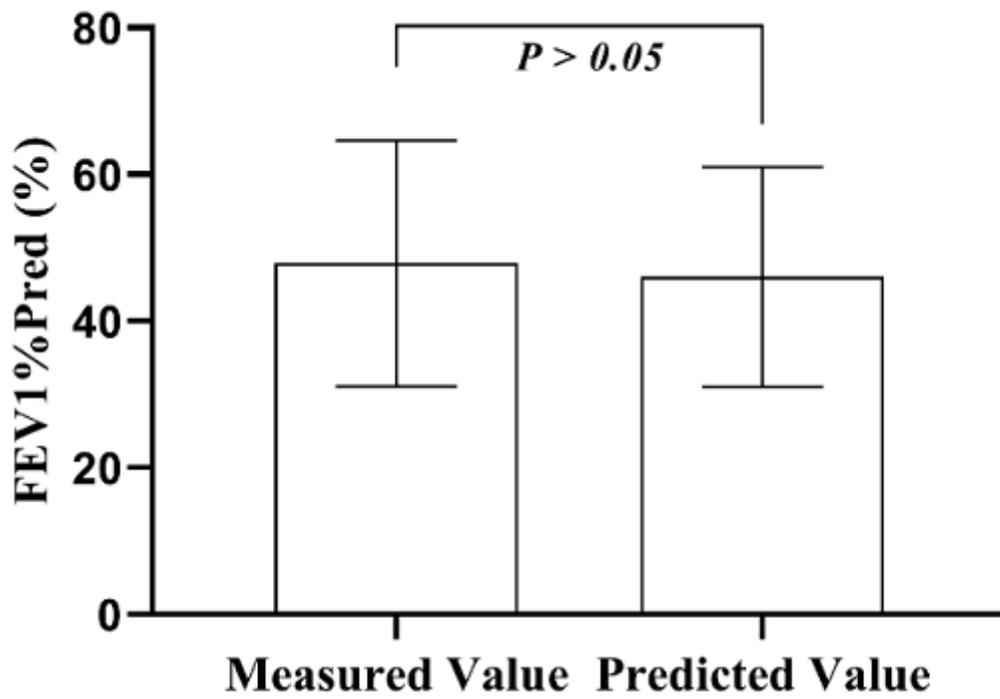
**Figure 3**

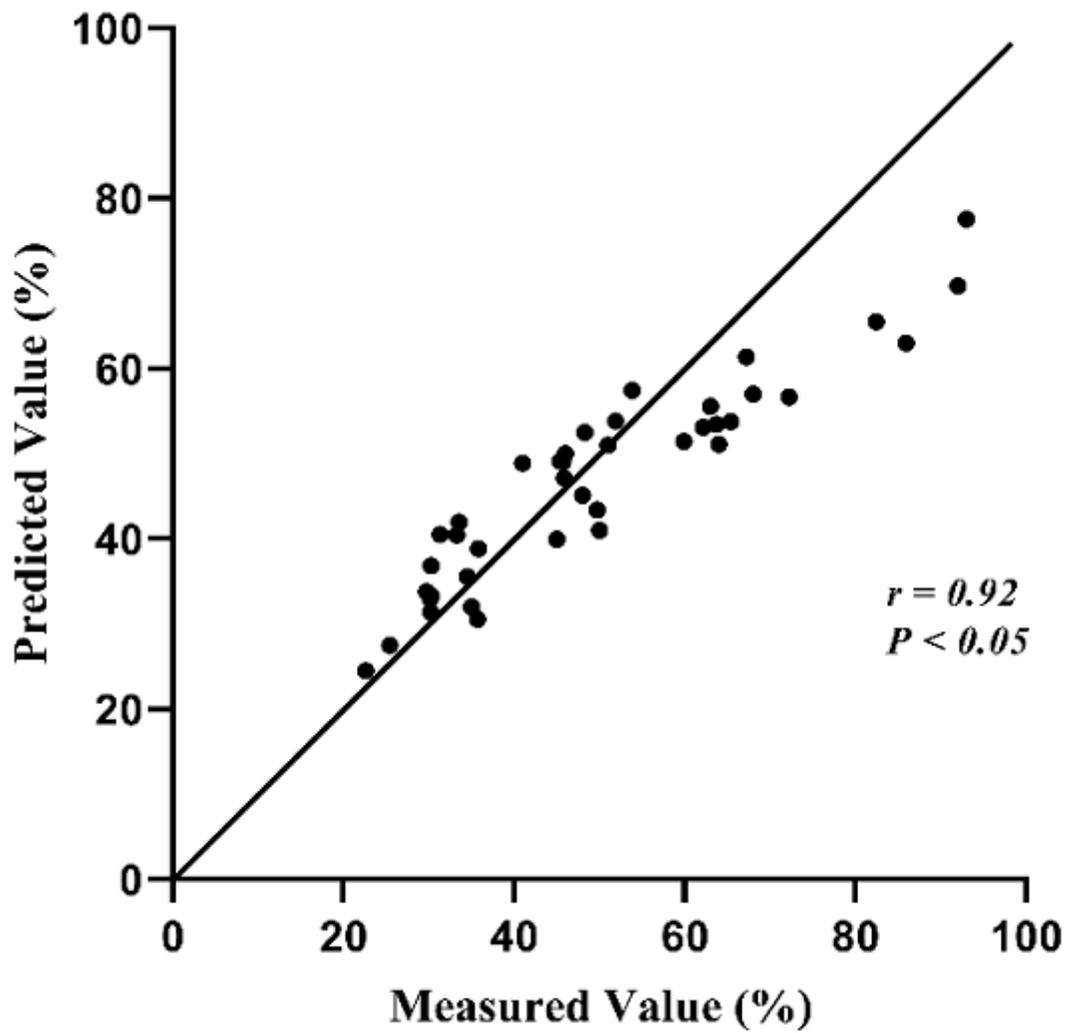Comparison between the predicted and the measured FEV1%pred values

**Figure 4**

Linear relationship between the predicted and the measured FEV1%pred values
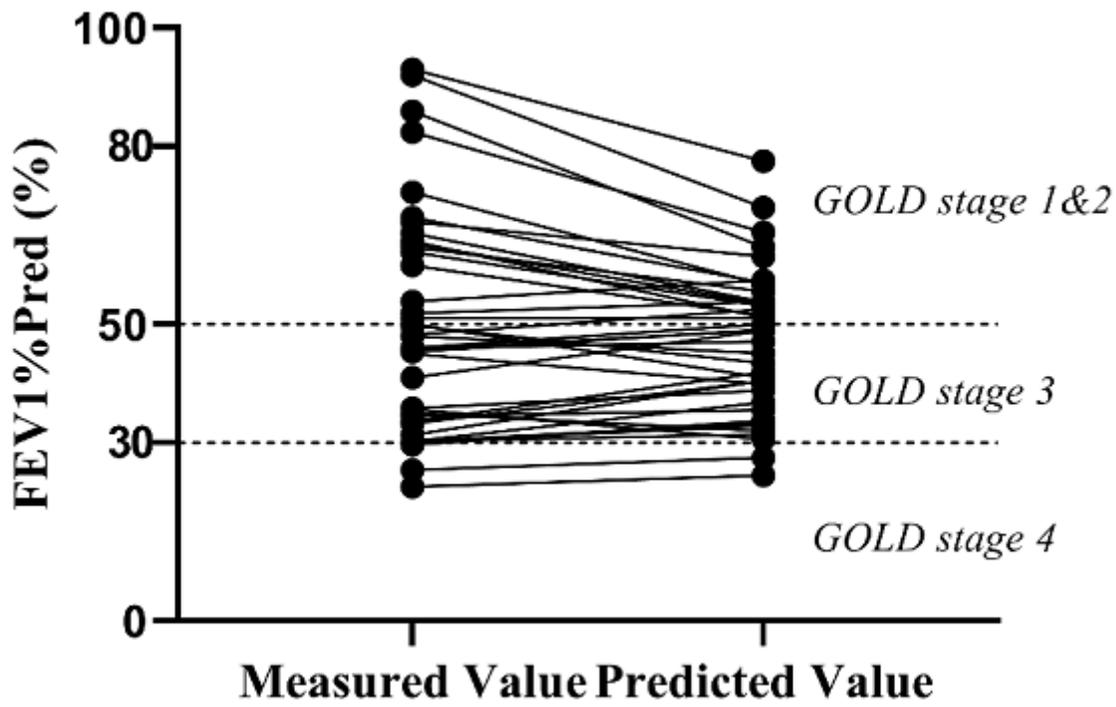
**Figure 5**

Comparison between the predicted and the measured GOLD stage. GOLD stage 1 and GOLD stage 2 were combined as the moderate group