

Comparing Heterogenous Phenotypes of Chronic Obstructive Pulmonary Disease: Network Analysis and Penalized Generalized Linear Model

Hyeon-Kyoung Koo¹, Sung Jun Chung¹, Dongil Park², Ho Cheol Kim³, Hyewon Seo⁴, Hyun Jung Kim⁵, Hyoung Kyu Yoon⁶, Chin Kook Rhee⁷, Kwang Ha Yoo⁸, Deog Kyeom Kim⁹

¹Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Ilsan Paik Hospital, Inje University College of Medicine, Goyang, Republic of Korea; ²Department of Internal Medicine, Division of Pulmonology, College of Medicine, Chungnam National University, Daejeon, Republic of Korea; ³Department of Internal Medicine, Gyeongsang National University Changwon Hospital, Gyeongsang National University School of Medicine, Changwon, Republic of Korea; ⁴Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, School of Medicine, Kyungpook National University, Kyungpook National University Hospital, Daegu, Republic of Korea; ⁵Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Keimyung University School of Medicine, Dongsan Hospital, Daegu, Republic of Korea; ⁶Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, College of Medicine, Yeouido St Mary's Hospital, The Catholic University of Korea, Seoul, Republic of Korea; ⁷Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea; ⁸Division of Pulmonary, Allergy and Critical Care Medicine, Department of Internal Medicine, Konkuk University School of Medicine, Seoul, Republic of Korea; ⁹Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Seoul Metropolitan Government-Seoul National University Boramae Medical Center, Seoul National University College of Medicine, Seoul, Republic of Korea

Correspondence: Deog Kyeom Kim, Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Seoul Metropolitan Government-Seoul National University Boramae Medical Center, 20 Boramaero-5-Gil, Dongjak-Gu, Seoul, 07061, Republic of Korea, Tel +82-2-870-2228, Fax +82-2-831-2826, Email kimdkmd@snu.ac.kr

Background and Objective: Chronic obstructive pulmonary disease (COPD) is a heterogeneous disease, with chronic bronchitis (CB) and emphysema phenotypes. The aim of our study was to compare the distinct patterns of correlation networks for respiratory symptoms and predictors of future exacerbations of different COPD phenotypes.

Methods: CB and emphysema were identified using a questionnaire and computed tomography images, respectively, and also included patients with preserved ratio impaired spirometry (PRISm). We constructed separate correlation networks for each subgroup using Spearman correlation coefficients. Predictors of future exacerbations were selected via least absolute shrinkage and selection operation regression analyses in multivariable analysis.

Results: Among the 3436 patients, 2232 were non-CB, 1131 were CB, 1116 were emphysema, and 73 were PRISm groups. The forced expiratory volume in one second (FEV1) and respiratory symptoms worsened in the following order: PRISm, non-CB, emphysema, and CB groups. During the 1-year follow-up, 17.3%, 21.3%, and 18.9% of patients in the non-CB, CB, and emphysema groups, respectively, experienced exacerbation. Each group showed a distinct correlation pattern between demographic characteristics, comorbidities, pulmonary function, blood biomarkers, respiratory symptoms, and exercise capacity. Across all groups, lower FEV1 (%), higher white blood cell count, higher erythrocyte sedimentation rate, and worse Saint George's Respiratory Questionnaire symptom and total scores were identified as common risk factors for future exacerbations. However, each group showed distinct predictors for future exacerbations.

Conclusion: The correlation network patterns and predictors of future exacerbations varied significantly depending on the COPD phenotype. Further research is required to understand the heterogeneous COPD pathophysiology and facilitate personalized medicine.

Plain Language Summary: COPD has various subtypes, including chronic bronchitis, emphysema, and PRISm phenotypes. This study compared symptom patterns and predictors of future exacerbation in these groups. We analyzed data from over 3400 patients and observed that lung function and symptoms worsened in the following order: PRISm, non-chronic bronchitis, emphysema, and chronic bronchitis. Each group showed distinct patterns of relationships between demographics, lung function, biomarkers, and respiratory symptoms. Although some risk factors overlapped, each group had its own predictors for future exacerbation. Understanding these differences among subtypes could lead to better personalized treatments for COPD patients.

Keywords: COPD, exacerbation, heterogeneous, network analysis

Introduction

Chronic obstructive pulmonary disease (COPD) is a heterogeneous disease characterized by diverse phenotypes,^{1,2} which poses significant clinical challenges in disease management and prognosis. Despite significant advances in the diagnosis and treatment of COPD, this heterogeneity makes it difficult to understand the pathogenesis and predict the disease prognosis. Although several subtypes of COPD are recognized clinically, including the frequent exacerbator subtype, asthma-COPD overlap, and upper lobe-predominant emphysema,³ it remains challenging to define, classify, and explain the various phenotypic manifestations. Traditionally, chronic bronchitis (CB)^{4,5} and emphysema have been regarded as the two primary phenotypes of COPD because of their distinct and well-characterized clinical and radiological features. Recently, preserved ratio impaired spirometry (PRISm) has also been proposed as an important phenotype within the COPD spectrum, warranting further investigation.

This study aimed to compare the characteristics of different phenotypes of COPD using the Korean COPD Subtype Study (KOCOSS) cohort database.⁶ To understand the pathogenesis underlying the respiratory symptoms and functional capacity, we compared the distinct patterns of correlation networks for demographics, lung function, blood biomarkers, dyspnea, COPD-specific quality of life, and exercise capacity for each COPD phenotype. We also identified the predictors of future exacerbations for each phenotype, addressing the other axis of COPD severity and prognosis. By highlighting these differences, our study underscores the need for phenotype-specific approaches to managing this complex disease.

Methods

Study Population and Data Collection

Participants in the KOCOSS cohort (as of August 29, 2023) were included in this analysis. The KOCOSS study design has been previously described.⁷ Briefly, the KOCOSS is a prospective, multicenter, consecutive cohort designed to identify COPD phenotypes and investigate the prognosis of COPD progression for each phenotype. Since April 2012, 54 medical centers in the Republic of Korea have participated in the KOCOSS. Patients aged ≥ 40 years with fixed airflow limitation who confirmed using pulmonary function tests were included in the KOCOSS. Fixed airflow limitation was defined as a post-bronchodilator forced expiratory volume in one second (FEV1)/forced vital capacity (FVC) $< 70\%$. COPD was diagnosed and classified according to the Global Initiative for Chronic Obstructive Lung Disease (GOLD) guidelines.⁸ Patients with PRISm, defined as an FEV1/FVC $> 70\%$ and an FEV1 $< 80\%$, were also included in the study.⁹

Clinical Parameters and Subgroup Definitions

Patients' baseline characteristics, including age, sex, smoking history, and body mass index (BMI), were recorded at the initial patient visit. Pulmonary function tests, such as FEV1, FVC, FEV1/FVC, and diffusion capacity of the lungs for carbon monoxide (DLCO), were performed at baseline and annually, following the manufacturer's instructions. Patient symptoms, quality of life, and functional exercise capacity scores were also obtained using the modified Medical Research Council (mMRC) dyspnea scale, COPD Assessment Test (CAT), Saint George's Respiratory Questionnaire (SGRQ), and six-minute walk distance test. The BODE index was calculated using these variables.

CB was defined as frequent coughing and/or sputum production for > 3 months per year over 2 consecutive years, as reported on patient questionnaires.¹⁰ Emphysema was identified using computed tomography (CT) images interpreted by a chest radiologist. Additionally, patients with PRISm were classified separately in this analysis.

Exacerbations were defined as an acute worsening of respiratory symptoms beyond the usual day-to-day variation, which necessitates a change in regular treatment. Exacerbations were classified as mild if manageable with short-acting bronchodilators alone; moderate if they required short-acting bronchodilator combined with antibiotics and/or oral corticosteroids; and severe if they needed hospitalization or an emergency department visit due to acute respiratory deterioration, potentially requiring intravenous therapies or advanced respiratory support. The occurrence of moderate or severe exacerbations was recorded during the 1-year follow-up. However, due to the low incidence of severe exacerbations, we aggregated moderate and severe exacerbations into combined outcome for exacerbation in our analyses.

Ethical Approval

The study was conducted in accordance with the principles of the Declaration of Helsinki. All participants provided written informed consent, and the confidentiality of their personal information was protected. The ethics committee of each participating center approved the study protocol (Institutional Review Board of Ilsan Paik Hospital, Inje University IRB No. 2014-05-118).

Statistical Analysis

Continuous data are presented as mean \pm standard deviation or median and interquartile range. Categorical data are presented as number and percentage. The *t*-test or Wilcoxon rank-sum test was used to compare continuous variables, and the chi-square test or Fisher's exact test was used to compare categorical variables. Spearman's rank correlation was applied to the correlation matrix to generate the correlation network. In the correlation network, each variable was represented by a node, and links between the nodes indicated the existence of a significant association with an absolute coefficient value of >0.3 .¹¹ Stronger associations (absolute coefficient of >0.4) were indicated by darker colors, with the thickness and darkness of the links representing the strength of the correlation (Spearman's coefficient), and the color denoting the direction of the association (blue for positive and pink for negative). The igraph package was used to visualize the correlation networks. Zou's method with an alpha level of 0.05 and a confidence level of 0.95 was used to compare the strength of the correlation between each variable.¹² The cocor package was used to compare the correlation strengths between paired variables in each patient subgroup. The predictors of future exacerbations in the multivariable analysis were identified based on least absolute shrinkage and selection operation (LASSO) regression analysis using the glmnet package.^{13–21} The accuracy of each model was evaluated by calculating the area under the curve (AUC) of the receiver operating characteristic (ROC) curve using the pROC package. A ten-fold cross-validation was performed using the boot package to assess predictive validity. All statistical analyses were performed using R software (version 4.3.1; The R Foundation for Computing, Vienna, Austria).

Results

Patient Characteristics

This study included 3436 patients with a mean age of 68.5 ± 8.0 years. Most patients (3168; 92.2%) were male. A total of 3363 patients were diagnosed with COPD, and 73 (3.1%) patients were included in the PRISm group. Among the COPD patients, 1131 (32.9%) met the criteria for CB, therefore, the non-CB group included 2232 (65.0%) patients. Emphysema was observed in 1116 (32.5%) patients on CT scan. The PRISm group had the lowest mean age, highest proportion of females, and highest mean BMI (Table 1). The emphysema group had the lowest mean BMI. The CB group had the highest proportion of current smokers, although the smoking amount was greater in the emphysema group. The prevalence of hypertension and heart disease was higher in the PRISm group and lower in the emphysema group. Gastro-esophageal reflux disease (GERD), asthma, allergic rhinitis, atopic dermatitis, and tuberculosis were the most prevalent in the CB group, and asthma and tuberculosis were the least common in the emphysema group. The FEV1 (%) was highest in the PRISm group, followed by the non-CB, emphysema, and CB groups. The FVC (%) was the highest in the emphysema group and lowest in the PRISm group. The FEV1/FVC (%), DLCO, DLCO (%), DLCO/VA (alveolar volume), and DLCO/VA (%) were the highest in the PRISm group, followed by the non-CB, CB, and emphysema groups. The mMRC and SGRQ-symptom scores increased in the following order: PRISm, non-CB, emphysema, and CB. The white blood cell (WBC) count was the highest in the CB and emphysema groups, although the eosinophil count and percentage were not different between the patient subgroups. The erythrocyte sedimentation rate (ESR) was the highest in the CB group. The BODE index was the highest in the CB group and the lowest in the PRISm group. The number of individuals with overlapping features (e-Figure 1) and their baseline characteristics are summarized in e-Table 1.

Correlations Between Variables

The correlation network between demographic, comorbidity, pulmonary function, laboratory, and quality of life characteristics was constructed (e-Figure 2). The DLCO decreased with age, and the DLCO, DLCO (%), DLCO/VA, and DLCO/VA (%) increased with BMI. The FVC was lower in women than in men. Comorbidities, including hypertension, diabetes, heart disease, GERD, asthma, history of tuberculosis, allergic rhinitis, and atopic dermatitis, were not significantly associated with any variable. FEV1 was negatively correlated with the mMRC, CAT, and SGRQ-symptom/activity/impact scores. The

Table 1 Patient Demographic Characteristics

	Total Population (N=3436)	Non-CB (N=2232)	CB (N=1131)	Emphysema (N=1116)	PRISm (N=73)
Demographics					
Age	68.5 ± 8.0	69.0 ± 7.9*	67.7 ± 8.1*	68.6 ± 8.0	64.8 ± 9.1*
Female sex	268 (7.8%)	179 (8.0%)	79 (7.0%)	45 (4.0%)*	10 (13.7%)
Body mass index	23.1 ± 3.4	23.2 ± 3.4	22.9 ± 3.4*	22.6 ± 3.4*	24.2 ± 3.7*
Smoking status					
Never smoker	316 (9.2%)	200 (9.0%)	101 (8.9%)	32 (2.9%)*	15 (20.8%)*
Ex-smoker	2210 (64.5%)	1506 (67.6%)	664 (58.8%)	751 (67.4%)	40 (55.6%)
Current smoker	903 (26.3%)	521 (23.4%)*	365 (32.3%)*	331 (29.7%)	17 (23.6%)
Smoking pack-year	38.2 ± 27.0	37.8 ± 26.42	39.3 ± 28.0	43.5 ± 26.4*	34.0 ± 31.4
Comorbidities					
Hypertension	1298 (37.8%)	862 (38.6%)	399 (35.3%)*	374 (33.5%)*	37 (50.7%)*
Diabetes	595 (17.3%)	391 (17.5%)	190 (16.8%)	191 (17.1%)	14 (19.2%)
Heart disease	1511 (44.1%)	984 (44.2%)	488 (43.2%)	445 (39.9%)*	39 (53.4%)
GERD	270 (7.9%)	151 (6.8%)*	114 (10.1%)*	78 (7.0%)	5 (6.8%)
Asthma	869 (25.4%)	511 (23.0%)*	345 (30.6%)*	249 (22.4%)*	13 (19.1%)
Tuberculosis	792 (23.1%)	487 (21.9%)*	286 (25.4%)*	219 (19.7%)*	19 (26.4%)
Allergic rhinitis	271 (7.9%)	141 (6.3%)*	122 (10.8%)*	70 (6.3%)	8 (11.1%)
Atopic dermatitis	66 (1.9%)	31 (1.4%)*	31 (2.8%)*	10 (6.3%)*	4 (5.6%)
Pulmonary function					
FEV1 (L)	1.66 ± 0.61	1.67 ± 0.62	1.61 ± 0.61*	1.63 ± 0.63*	1.99 ± 0.46*
FEV1 (%)	57.0 ± 18.9	57.8 ± 19.0*	54.9 ± 18.7*	55.4 ± 19.1*	65.3 ± 9.5*
FVC (L)	3.23 ± 0.84	3.23 ± 0.85	3.25 ± 0.81	3.36 ± 0.83*	2.72 ± 0.66*
FVC (%)	78.8 ± 16.7	79.1 ± 16.8	79.1 ± 16.5	81.1 ± 17.2*	65.2 ± 11.3*
FEV1/FVC (%)	51.3 ± 12.9	51.5 ± 12.4	49.5 ± 12.9*	48.2 ± 12.3*	75.5 ± 12.3*
DLCO	13.42 ± 4.81	13.46 ± 4.77	13.26 ± 4.86	12.5 ± 4.7*	14.9 ± 5.2*
DLCO (%)	65.1 ± 21.0	65.7 ± 20.9*	63.7 ± 21.0*	59.9 ± 20.7*	69.2 ± 21.5
DLCO/VA	3.12 ± 0.97	3.13 ± 0.97	3.06 ± 0.95*	2.79 ± 0.91*	3.82 ± 1.03*
DLCO/VA (%)	76.1 ± 23.0	76.5 ± 23.1	74.5 ± 22.5*	69.0 ± 22.4*	89.3 ± 23.5*
Laboratory findings					
WBC (/ μ L*1000)	7.28 ± 2.43	7.16 ± 2.33*	7.53 ± 2.62*	7.51 ± 2.63*	6.78 ± 2.09
Eosinophil (%)	3.27 ± 3.19	3.22 ± 3.04	3.34 ± 3.42	3.11 ± 2.89	3.70 ± 3.81
BEC (/ μ L)	227.4 ± 247.5	221.1 ± 226.3	239.8 ± 285.1	223.8 ± 250.7	232.7 ± 242.0
Hb (g/dL)	14.2 ± 1.6	14.2 ± 1.6	14.2 ± 1.5	14.2 ± 1.5	14.2 ± 1.6
ESR (mm/h)	16.2 ± 16.7	15.6 ± 16.3*	17.3 ± 17.5*	16.7 ± 16.9	16.4 ± 16.0
Quality of life					
mMRC	1.30 ± 0.91	1.20 ± 0.85*	1.50 ± 1.00*	1.38 ± 0.93*	1.08 ± 0.89*
CAT	14.0 ± 8.0	12.4 ± 7.3*	17.1 ± 8.5*	14.2 ± 8.2	11.3 ± 7.2*
SGRQ-symptom	38.4 ± 21.4	33.4 ± 18.3*	51.7 ± 22.0*	39.1 ± 21.9	32.6 ± 20.0*
SGRQ-activity	38.6 ± 26.7	35.7 ± 25.3*	44.8 ± 28.3*	41.0 ± 27.3*	29.8 ± 23.1*
SGRQ-impact	20.6 ± 21.4	16.6 ± 18.2*	28.8 ± 25.0*	21.5 ± 22.5*	14.3 ± 17.2*
SGRQ-total	29.5 ± 20.7	25.6 ± 18.2*	37.7 ± 22.9*	30.7 ± 21.6*	22.3 ± 16.8*
6MWD	381.5 ± 116.2	383.1 ± 118.4	381.0 ± 111.7	384.9 ± 114.1	340.3 ± 117.8*
BODE index	2.27 ± 1.89	2.12 ± 1.84*	2.60 ± 1.97*	2.45 ± 2.01*	1.62 ± 1.54*

(Continued)

Table 1 (Continued).

	Total Population (N=3436)	Non-CB (N=2232)	CB (N=1131)	Emphysema (N=1116)	PRISm (N=73)
Exacerbation (1Y)					
Moderate	568 (16.5%)	345 (15.5%)*	223 (19.7%)*	191 (17.1%)*	0 (0%)*
Severe	163 (4.7%)	102 (4.6%)	61 (5.4%)	66 (5.9%)*	0 (0%)
Total	626 (18.2%)	385 (17.3%)*	241 (21.3%)*	211 (18.9%)*	0 (0%)*

Note: *indicates a statistically significant difference compared to the total population.

Abbreviations: CB, chronic bronchitis; PRISm, preserved ratio impaired spirometry; PY, pack-year; GERD, gastro-esophageal reflux disease; FEV1, forced expiratory volume in 1 second; FVC, forced vital capacity; DLCO, diffusion capacity; DLCO/VA, diffusion capacity/alveolar volume; WBC, white blood cell; BEC, blood eosinophil count; Hb, hemoglobin; ESR, erythrocyte sedimentation rate; mMRC, modified Medical Research Council; CAT, COPD assessment test; SGRQ, St. George Respiratory Questionnaire; 6MWD, 6-minute walk distance.

mMRC score was negatively associated with FEV1, FEV1 (%), FVC, DLCO, and DLCO (%). The SGRQ-activity score was negatively correlated with FEV1, FEV1 (%), FVC, FEV1/FVC, DLCO, and DLCO (%). The association between pulmonary function and six-minute walk distance (6MWD) was not significant.

Separate correlation matrices (e-Figure 3) and networks (Figure 1) were constructed for each patient subgroup to compare the correlation patterns. The correlation networks revealed distinct patterns among the groups. In the non-CB

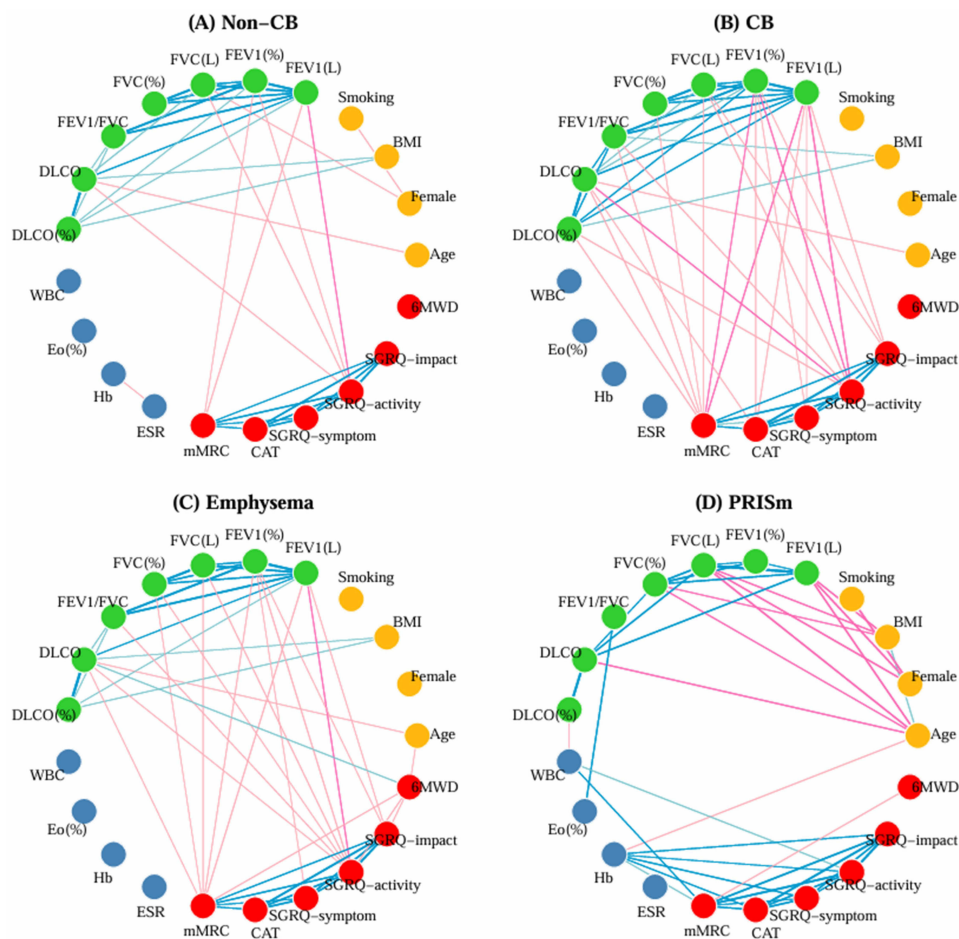


Figure 1 Correlation networks. The correlation networks show the different patterns between the chronic obstructive pulmonary disease phenotype groups. (A) Non-chronic bronchitis, (B) chronic bronchitis, (C) emphysema, and (D) preserved ratio impaired spirometry.

Abbreviations: CB, chronic bronchitis; PRISm, preserved ratio impaired spirometry; FEV1, forced expiratory volume in 1 second; FVC, forced vital capacity; DLCO, diffusion capacity of the lungs for carbon monoxide; WBC, white blood cell count; Eo, eosinophil; Hb, hemoglobin; ESR, erythrocyte sedimentation rate; mMRC, modified Medical Research Council; CAT, COPD assessment test; SGRQ, Saint George's Respiratory Questionnaire; 6MWD, 6-minute walk distance.

group, the mMRC score was negatively associated with FEV1 and FEV1 (%), and the SGRQ-activity score was negatively correlated with FEV1, FEV1 (%), FVC, and DLCO. No significant associations between pulmonary function, blood biomarkers, CAT, SGRQ-symptom, SGRQ-impact scores, and the 6MWD were identified. In the CB group, smoking status or FVC did not significantly differ between men and women. FEV1, FEV1 (%), FVC, FEV1/FVC, DLCO, and DLCO (%) were closely and strongly associated with mMRC, CAT, and all domains of the SGRQ scores. No significant association between pulmonary function and the 6MWD was identified. In the emphysema group, no difference in smoking status or FVC was also observed between men and women. The 6MWD was negatively correlated with age, mMRC, SGRQ-activity score, and SGRQ-impact score and positively correlated with DLCO. In the PRISm group, patient age was negatively correlated with FEV1, FEV1 (%), FVC (%), and DLCO. Additionally, BMI was negatively correlated with FVC and FVC (%), but not with DLCO or DLCO (%). The WBC count was positively associated with the mMRC and SGRQ-activity scores. Increased hemoglobin levels were associated with higher mMRC and SGRQ-activity scores. However, there was no significant association between pulmonary function and the mMRC, CAT, SGRQ-symptom, SGRQ-activity, SGRQ-impact scores, or the 6MWD. The mMRC score was negatively associated with the 6MWD. [e-Figure 4](#) provides a summary of the *P*-values for the comparison of Spearman correlation coefficients between each group and the total population. The correlation network identified 37 (17.6%) connections in the non-CB group, 50 (23.8%) in the CB group, 45 (21.4%) in the emphysema group, and 42 (20.0%) in the PRISm group out of a total of 210 possible connections.

Predictors of Exacerbations

During the 1-year follow-up, 568 (16.5%) moderate and 163 (4.7%) severe exacerbations occurred. Since 105 (3.1%) patients experienced both moderate and severe exacerbations, a total of 626 (18.2%) moderate-to-severe exacerbations were reported. The frequency of moderate exacerbations was increased in the CB and emphysema groups, and that of severe exacerbations was elevated only in the emphysema group ([Table 1](#)). No patient in the PRISm group experienced exacerbations. Significant variables were selected for multivariable analysis using LASSO regression, and the procedures for selecting variables are depicted in [e-Figure 5](#). Overall, exacerbations were associated with advanced age, lower BMI, current smoking status, smoking amount, presence of asthma, history of tuberculosis, GERD, lower FEV1 (%), FVC (%), and FEV1/FVC, and higher WBC count, eosinophil percentage, and ESR. In addition, higher SGRQ-symptom and SGRQ total scores, lower 6MWD, and higher BODE index were predictors of future exacerbations. The characteristics of patients with history of tuberculosis are compared in [e-Table 2](#).

In the non-CB group, exacerbations were associated with advanced age, lower BMI, current smoking status, smoking amount, presence of asthma or GERD, and history of tuberculosis. Lower FEV1 (%) and FVC (%), and higher WBC count, eosinophil percentage, ESR, SGRQ-symptom score, SGRQ total score, and BODE index were significantly associated with exacerbations in this group. In the CB group, lower BMI, smoking amount, history of asthma, tuberculosis, and GERD, lower FVC (%), FEV1/FVC, and DLCO (%), and higher WBC count, ESR, and SGRQ-symptom, SGRQ-activity, and total SGRQ scores were associated with exacerbations. In the emphysema group, advanced age, female sex, lower BMI, history of tuberculosis, lower FEV1 (%), and higher WBC count, ESR, SGRQ-symptom, SGRQ total scores, and lower 6MWD were significantly associated with exacerbations ([Table 2](#)). The AUC of the ROC curves was 0.729 (95% confidence interval [CI]: 0.677–0.781) for the total population, 0.725 (95% CI: 0.659–0.791) for the non-CB group, 0.731 (95% CI: 0.645–0.817) for the CB group, and 0.786 (95% CI: 0.719–0.853) for the emphysema group ([Figure 2](#)). The logistic regression models were validated using ten-fold cross-validation. The predictive validities for the total population, non-CB, CB, and emphysema groups were 0.790, 0.791, 0.755, and 0.805, respectively.

Discussion

In this study, we identified the complex relationship between COPD-related factors and their distinct patterns in different phenotypes. The non-CB group showed fewer correlations in the correlation network. In the non-CB group, FEV1, FEV1 (%), FVC, and DLCO were associated with a lower SGRQ-activity score. In the CB group, FEV1, FEV1 (%), FVC, FEV1/FVC, DLCO, and DLCO (%) were strongly associated with mMRC, CAT, and all domains of the SGRQ score.

Table 2 Predictors of Future Exacerbations

	Total	Non-CB	CB	Emphysema
Demographics				
Age	1.005	1.005	–	1.002
Female sex	–	–	–	1.052
BMI (kg/m ²)	0.987	0.990	0.999	0.985
Smoking status	1.040	1.054	–	–
Pack-year	1.003	1.001	1.005	–
Asthma	1.157	1.212	1.013	–
Tuberculosis	1.310	1.323	1.281	1.672
GERD	1.313	1.277	1.249	–
Pulmonary function				
FEV1 (%)	0.987	0.990	–	0.976
FVC (%)	0.996	0.995	0.990	–
FEV1/FVC	0.994	–	0.980	–
DLCO (%)	–	–	0.990	–
Laboratory findings				
WBC (/μL*1000)	1.097	1.141	1.029	1.213
Eosinophil (%)	1.004	1.016	–	–
Hb (g/dL)	–	–	0.998	–
ESR (mm/h)	1.010	1.011	1.007	1.003
Quality of life				
mMRC	–	–	–	–
CAT	–	–	–	–
SGRQ-symptom	1.012	1.011	1.018	1.019
SGRQ-activity	–	–	1.001	–
SGRQ-impact	–	–	–	–
SGRQ-total	1.008	1.009	1.005	1.019
6MWD	0.999	–	–	0.998
BODE index	1.003	1.089	–	–

Abbreviations: CB, chronic bronchitis; BMI, body mass index; FEV1, forced expiratory volume in 1 second; FVC, forced vital capacity; DLCO, diffusion capacity; DLCO/VA, diffusion capacity/alveolar volume; WBC, white blood cell; Hb, hemoglobin; ESR, erythrocyte sedimentation rate; CAT, COPD assessment test; SGRQ, St. George Respiratory Questionnaire; 6MWD, 6-minute walk distance.

The emphysema group showed a positive correlation between DLCO and the 6MWD. In the PRISm group, age was inversely correlated with pulmonary function, including FEV1, FEV1 (%), FVC, FVC (%), and DLCO; however, the associations between pulmonary function and respiratory symptoms or exercise capacity were weak. In the multivariable analyses for future exacerbations, lower BMI, history of tuberculosis, higher WBC count, ESR, SGRQ-symptom score, and total SGRQ score were identified as common risk factors for all groups, while specific risk factors were identified for each patient group. A lower FEV1 was not significant as a predictor of exacerbations in the CB group.

COPD is suspected in any patient with dyspnea, chronic cough, or sputum production and a history of exposure to risk factors for the disease. To establish a diagnosis of COPD, a post-bronchodilator FEV1/FVC of <0.7 on spirometry is required.^{8,22} However, pulmonary function tests do not fully capture the heterogeneity of COPD and do not adequately represent patient health status.²³ Therefore, the severity of airflow obstruction was removed from the combined assessment scheme due to its insufficient power to predict outcomes and guide treatment, and a multidimensional staging system that includes quality of life and exacerbation history has been adopted for the clinical assessment of COPD.⁸

In previous review on predicting future exacerbations, distinct variables have been identified across different studies.²⁴ These discrepancies may be attributed to variations in study populations or differences in the definition of exacerbations. However, we found that differences in COPD phenotypes are responsible for the poor precision of lung function in terms of symptoms, exercise capacity, and future exacerbations. Different patient selection strategies may

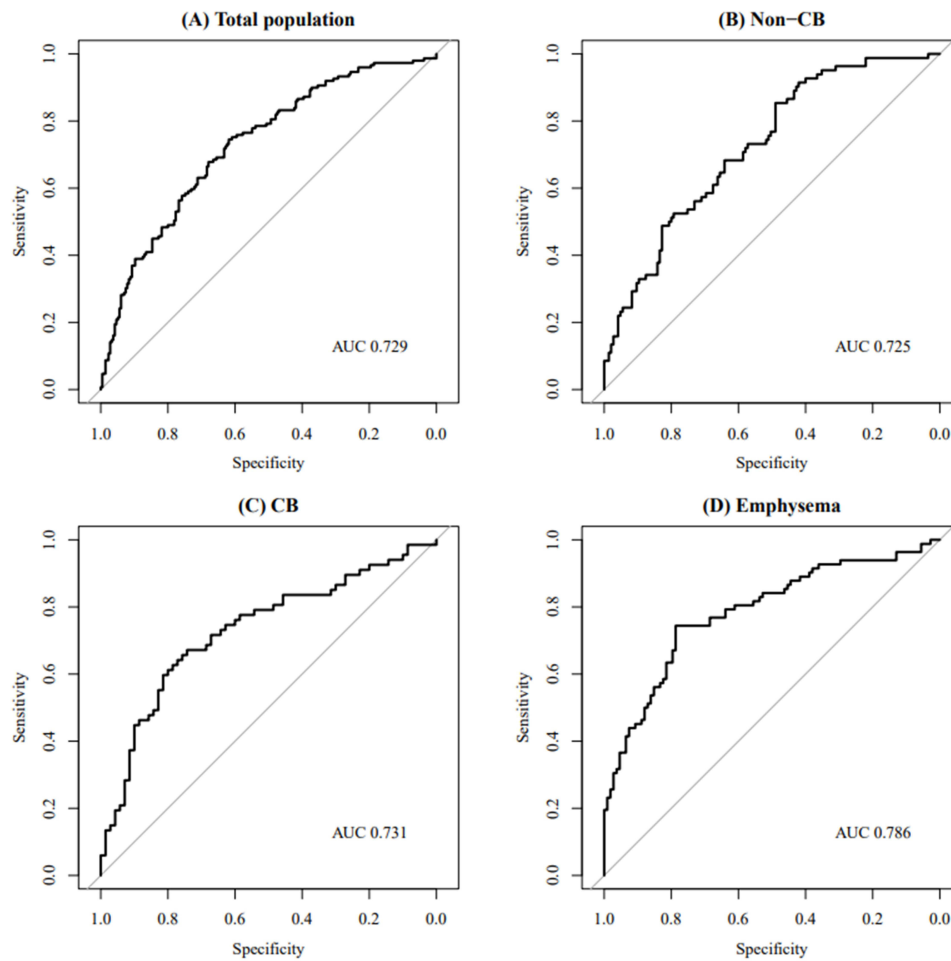


Figure 2 Receiver operating characteristic curves. The receiver operating characteristic curves representing the risk of future exacerbations in the (A) total population, (B) non-chronic bronchitis group, (C) chronic bronchitis group, and (D) emphysema group are shown.

Abbreviations: CB, chronic bronchitis; AUC, area under the curve of the receiver operating characteristic curve.

account for the inconsistent results of previous association studies. Therefore, subtype-based personalized approach would be necessary to better predict the characteristics and prognosis in COPD patients.

Consistent with previous studies,^{23,25–27} our findings demonstrated that the CB group exhibited worse symptoms, lower quality of life, and higher rate of exacerbations. Regarding the predictors of exacerbations, lower BMI, history of tuberculosis, and elevated WBC and ESR were identified as common risk factors for future exacerbation. The association between lower BMI and exacerbations suggests that malnutrition and muscle wasting might contribute to impaired immune function and respiratory muscle weakness, consequently leading to higher mortality in COPD. These factors could also affect airway clearance, exercise tolerance, and overall exacerbation risk. The link between elevated WBC, ESR, and exacerbation supports the hypothesis that chronic inflammation contributes to airway instability, mucus hypersecretion, and increased susceptibility to infection.²⁴ Additionally, since inflammatory markers and frequency of exacerbations were elevated in patients with history of tuberculosis, consistent with previous study,²⁸ tuberculosis may be linked to local airway inflammation, further predisposing patients to exacerbations.

In our study, pulmonary function was sufficient to predict respiratory symptoms in the CB group, but it was less accurate in the non-CB and PRISm groups. In contrast, exercise capacity was significantly associated with DLCO only in the emphysema group. Our results indicate that DLCO must be evaluated to predict exercise capacity in the emphysema group and future exacerbations in the CB group. Furthermore, while the SGRQ-activity score accurately reflected pulmonary function in the non-CB, CB, and emphysema groups, the SGRQ-symptom score should be carefully evaluated to predict future exacerbations.

Recently, the significance of blood eosinophil count as a biomarker of COPD has been extensively documented. An elevated blood eosinophil count serves as an indicator of a higher risk of future exacerbations and a more favorable response to inhaled corticosteroid (ICS) treatment.^{29–32} In the current study, the proportion of patients with history of asthma was higher in the CB group, potentially representing the asthma-COPD overlap phenotype. However, blood eosinophil counts did not differ significantly between groups. History of asthma was a significant predictor for future exacerbations in both the non-CB and CB group, while blood eosinophil counts were more relevant in the non-CB group than in the CB or emphysema group. Further research is needed to refine the indication for ICS containing regimen and advance personalized treatment strategies.

According to the GOLD guidelines, additional tests are needed when there is a marked discrepancy between the level of airflow obstruction and perceived symptoms,⁸ including lung volume measurements, DLCO, arterial blood gas analysis, and exercise testing. In the current study, the mMRC, CAT, and SGRQ scores were attributed to various pulmonary functions in the CB group; in contrast, DLCO was associated with the 6MWD only in the emphysema group. Our study results were consistent with the traditional emphasis on DLCO in patients with emphysema. Interestingly, DLCO was associated with dyspnea and quality of life in the CB group, but not with exercise capacity; however, reduced DLCO was associated with future exacerbations in the CB group. On the contrary, DLCO was associated with exercise capacity, but not with future exacerbations in the emphysema group. Therefore, more research regarding the importance of DLCO is necessary. In the PRISm group, no correlation was observed between lung function and respiratory symptoms or exercise capacity. The PRISm group still remains a novel phenotype that requires a deeper understanding of its clinical manifestations. Further studies are required to better understand the disease pathogenesis in patients with PRISm.

The key novelty of our study lies in the network-based approach, which is increasingly popular for the comprehensive analysis of big data. Furthermore, the inclusion of PRISm patients also offers a novel perspective, as the clinical implications for this group remain poorly understood. However, our study also has several limitations. We recorded only whether an exacerbation occurred, without capturing the precise timing during follow-up. Consequently, time-based analyses, such as survival analysis, were not feasible. We also did not compare mortality between the patient subgroups. In addition, there was insufficient power to compare the rate of lung function decline between the patient groups. Lastly, although we were able to identify characteristic differences between COPD subtypes, the underlying mechanisms driving these differences could not be determined through our study design. Further research should incorporate more detailed data on exacerbation timing and patient outcomes, enabling a more comprehensive analysis of exacerbation pattern, disease progression, and mortality. Additional efforts are required to generalize and understand the pathophysiology underlying each COPD phenotype.

In conclusion, the correlation patterns between variables and risk factors for exacerbation are significantly different among the COPD phenotypes. To enable personalized medicine, separate diagnostic and prognostic strategies must be developed for each heterogeneous disease phenotype.

Abbreviations

6MWD, six-minute walk distance; BMI, body mass index; CAT, chronic obstructive pulmonary disease assessment test; CB, chronic bronchitis; COPD, chronic obstructive pulmonary disease; DLCO, diffusion capacity of the lungs for carbon monoxide; ESR, erythrocyte sedimentation rate; FEV1, forced expiratory volume in one second; FVC, forced vital capacity; KOCOSS, Korean COPD Subtype Study; mMRC, modified Medical Research Council; PRISm, preserved ratio impaired spirometry; SGRQ, Saint George's Respiratory Questionnaire; WBC, white blood cell.

Data Sharing Statement

The data that support the findings of this study are available from the KOrea COPd Subgroup Study (KOCOSS) team but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of the KOCOSS team.

Human Ethics Approval Declaration

This study was conducted in accordance with the Declaration of Helsinki, approved by the Institutional Review Board at each hospital (Ilsan Paik Hospital, Inje University: No. 2014-05-118). Written informed consent was obtained from all participants.

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, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This work was supported by the Research Program of the Korea National Institute of Health (Fund CODE 2016ER670100, 2016ER670101, 2016ER670102, 2018ER67100, 2018ER67101, 2018ER67102, 2021ER120500, 2021ER120501, 2021ER120502, 2024ER120100 and 2024ER120101), as well as by a grant from the National Research Foundation of Korea (NIRF), funded by the Korean government (MSIT: grant no. RS-2024-00359875).

Disclosure

The authors report no competing interests in this work.

References

- Agusti A. Phenotypes and disease characterization in chronic obstructive pulmonary disease. Toward the extinction of phenotypes? *Ann Am Thorac Soc.* 2013;10 Suppl:S125–30. doi:10.1513/AnnalsATS.201303-055AW
- Miravittles M, Soler-Cataluna JJ, Calle M, Soriano JB. Treatment of COPD by clinical phenotypes: putting old evidence into clinical practice. *Eur Respir J.* 2013;41(6):1252–1256. doi:10.1183/09031936.00118912
- Corlateanu A, Mendez Y, Wang Y, Garnica RJA, Botnaru V, Siafakas N. Chronic obstructive pulmonary disease and phenotypes: a state-of-the art. *Pulmonology.* 2020;26(2):95–100. doi:10.1016/j.pulmoe.2019.10.006
- Fletcher C, Peto R. The natural history of chronic airflow obstruction. *BMJ.* 1977;1(6077):1645–1648. doi:10.1136/bmj.1.6077.1645
- Dotan Y, So JY, Kim V. Chronic bronchitis: where are we now? *Chronic Obstruct Pulm Dis.* 2019;6(2):178–192.
- The Korea COPD Subgroup Study. Available from: <http://www.kocoss.kr/>. Accessed April 7, 2025.
- Lee J, Chon GR, Rhee CK, et al. Characteristics of patients with chronic obstructive pulmonary disease at the first visit to a pulmonary medical center in Korea: the Korea COPd Subgroup Study Team Cohort. *J Korean Med Sci.* 2016;31(4):553–560. doi:10.3346/jkms.2016.31.4.553
- Global initiative for chronic obstructive lung disease (GOLD) guidelines, global strategy for the diagnosis, management and prevention of chronic obstructive lung disease. 2023. Available from: <https://goldcopd.org/2021-gold-reports/>. Accessed April 7, 2025.
- Wan ES, Castaldi PJ, Cho MH, et al.; COPDGene Investigators. Epidemiology, genetics, and subtyping of preserved ratio impaired spirometry (PRISm) in COPDGene. *Respir Res.* 2014;15(1):89. doi:10.1186/s12931-014-0089-y
- Ferris BG. Epidemiology standardization project (American Thoracic Society). *Am Rev Respir Dis.* 1978;118(6 Pt 2):1–120.
- Liang Y, Abbott D, Howard N, Lim K, Ward R, Elgendi M. How effective is pulse arrival time for evaluating blood pressure? Challenges and recommendations from a study using the MIMIC database. *J Clin Med.* 2019;8(3):337. doi:10.3390/jcm8030337
- Zou GY. Toward using confidence intervals to compare correlations. *Psychol Methods.* 2007;12:399–413. doi:10.1037/1082-989X.12.4.399
- Tibshirani R. The Lasso method for variable selection in the Cox model. *Stat Med.* 1997;16(4):385–395. doi:10.1002/(SICI)1097-0258(19970228)16:4<385::AID-SIM380>3.0.CO;2-3
- Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw.* 2010;33(1):1–22. doi:10.18637/jss.v033.i01
- Daneshvar A, Mousa G. Regression shrinkage and selection via least quantile shrinkage and selection operator. *PLoS One.* 2023;18(2):e0266267. doi:10.1371/journal.pone.0266267
- Zhao X, Li W, Zhang J, et al. Radiomics analysis of CT imaging improves preoperative prediction of cervical lymph node metastasis in laryngeal squamous cell carcinoma. *Eur Radiol.* 2023;33(2):1121–1131. doi:10.1007/s00330-022-09051-4
- Kang J, Choi YJ, Kim IK, et al. LASSO-based machine learning algorithm for prediction of lymph node metastasis in T1 colorectal cancer. *Cancer Res Treat.* 2021;53(3):773–783. doi:10.4143/crt.2020.974
- Ribbing J, Jonsson EN. Power, selection bias and predictive performance of the population pharmacokinetic covariate model. *J Pharmacokinet Pharmacodyn.* 2004;31(2):109–134. doi:10.1023/B:JOPA.0000034404.86036.72
- Mundry R, Nunn CL. Stepwise model fitting and statistical inference: turning noise into signal pollution. *Am Nat.* 2009;173(1):119–123. doi:10.1086/593303
- Wiegand RE. Performance of using multiple stepwise algorithms for variable selection. *Stat Med.* 2010;29(15):1647–1659. doi:10.1002/sim.3943
- Smith G. Step away from stepwise. *J Big Data.* 2018;5:32. doi:10.1186/s40537-018-0143-6

22. Buist AS, McBurnie MA, Vollmer WM, et al. International variation in the prevalence of COPD (the BOLD Study): a population-based prevalence study. *Lancet*. 2007;370(9589):741–750. doi:10.1016/S0140-6736(07)61377-4
23. Agusti A, Calverley PM, Celli B, et al. Characterisation of COPD heterogeneity in the ECLIPSE cohort. *Respir Res*. 2010;11:122. doi:10.1186/1465-9921-11-122
24. Hurst JR, Han MK, Singh B, et al. Prognostic risk factors for moderate-to-severe exacerbations in patients with chronic obstructive pulmonary disease: a systematic literature review. *Respir Res*. 2022;23(1):213. doi:10.1186/s12931-022-02123-5
25. Choi JY, Yoon HK, Park SJ, et al. Chronic bronchitis is an independently associated factor for more symptom and high-risk groups. *Int J Chron Obstruct Pulmon Dis*. 2016;11:1335–1341. doi:10.2147/COPD.S105516
26. de Oca MM, Halbert RJ, Lopez MV, et al. The chronic bronchitis phenotype in subjects with and without COPD: the PLATINO study. *Eur Respir J*. 2012;40(1):28–36. doi:10.1183/09031936.00141611
27. Kim V, Han MK, Vance GB, et al. The chronic bronchitic phenotype of COPD: an analysis of the COPDGene Study. *Chest*. 2011;140(3):626–633. doi:10.1378/chest.10-2948
28. Oh JY, Lee YS, Min KH, et al. Difference in systemic inflammation and predictors of acute exacerbation between smoking-associated COPD and tuberculosis-associated COPD. *Int J Chron Obstruct Pulmon Dis*. 2018;13:3381–3387. doi:10.2147/COPD.S177371
29. Brusselle G, Pavord ID, Landis S, et al. Blood eosinophil levels as a biomarker in COPD. *Respir Med*. 2018;138:21–31. doi:10.1016/j.rmed.2018.03.016
30. Hastie AT, Martinez FJ, Curtis JL, et al. Association of sputum and blood eosinophil concentrations with clinical measures of COPD severity: an analysis of the SPIROMICS cohort. *Lancet Respir Med*. 2017;5(12):956–967. doi:10.1016/S2213-2600(17)30432-0
31. Singh D, Agusti A, Martinez FJ, et al. Blood eosinophils and chronic obstructive pulmonary disease: a global initiative for chronic obstructive lung disease science committee 2022 review. *Am J Respir Crit Care Med*. 2022;206(1):17–24. doi:10.1164/rccm.202201-0209PP
32. Kang HS, Kim SK, Kim YH, et al. The association between eosinophilic exacerbation and eosinophilic levels in stable COPD. *BMC Pulm Med*. 2021;21(1):74. doi:10.1186/s12890-021-01443-4

International Journal of Chronic Obstructive Pulmonary Disease

Publish your work in this journal

The International Journal of COPD is an international, peer-reviewed journal of therapeutics and pharmacology focusing on concise rapid reporting of clinical studies and reviews in COPD. Special focus is given to the pathophysiological processes underlying the disease, intervention programs, patient focused education, and self management protocols. This journal is indexed on PubMed Central, MedLine and CAS. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/international-journal-of-chronic-obstructive-pulmonary-disease-journal>

Dovepress
Taylor & Francis Group