

Original Article

ARCHIVOS DE Bronconeumología



www.archbronconeumol.org

Clinical Characteristics and Outcomes of the Phenotypes of COPD-Bronchiectasis Association



Cui-xia Pan^{a,1}, Zhen-feng He^{a,1}, Sheng-zhu Lin^{a,1}, Jun-qing Yue^{a,1}, Zhao-ming Chen^{a,1}, Wei-jie Guan^{a,b,*}

^a State Key Laboratory of Respiratory Disease, National Clinical Research Center for Respiratory Disease, National Center for Respiratory Medicine, Department of Respiratory and Critical Care Medicine, Guangzhou Institute of Respiratory Health, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou, Guangdong, China
^b Guangzhou National Laboratory, Guangzhou, Guangdong, China

ARTICLE INFO

Article history: Received 19 January 2024 Accepted 3 April 2024 Available online 23 April 2024

Keywords: Bronchiectasis COPD Clinical characteristics Two-step clustering

ABSTRACT

Introduction: Although COPD may frequently co-exist with bronchiectasis [COPD-bronchiectasis associated (CBA)], little is known regarding the clinical heterogeneity. We aimed to identify the phenotypes and compare the clinical characteristics and prognosis of CBA.

Methods: We conducted a retrospective cohort study involving 2928 bronchiectasis patients, 5158 COPD patients, and 1219 patients with CBA hospitalized between July 2017 and December 2020. We phenotyped CBA with a two-step clustering approach and validated in an independent retrospective cohort with decision-tree algorithms.

Results: Compared with patients with COPD or bronchiectasis alone, patients with CBA had significantly longer disease duration, greater lung function impairment, and increased use of intravenous antibiotics during hospitalization. We identified five clusters of CBA. Cluster 1 (N = 120, CBA-MS) had predominantly moderate–severe bronchiectasis, Cluster 2 (N = 108, CBA-FH) was characterized by frequent hospitalization within the previous year, Cluster 3 (N = 163, CBA-BI) had bacterial infection, Cluster 4 (N = 143, CBA-NB) had infrequent hospitalization but no bacterial infection, and Cluster 5 (N = 113, CBA-NHB) had no hospitalization or bacterial infection in the past year. The decision-tree model predicted the cluster assignment in the validation cohort with 91.8% accuracy. CBA-MS, CBA-BI, and CBA-FH exhibited higher risks of hospital re-admission and intensive care unit admission compared with CBA-NHB during follow-up (all P < 0.05). Of the five clusters, CBA-FH conferred the worst clinical prognosis.

Conclusion: Bronchiectasis severity, recent hospitalizations and sputum culture findings are three defining variables accounting for most heterogeneity of CBA, the characterization of which will help refine personalized clinical management.

© 2024 SEPAR. Published by Elsevier España, S.L.U. All rights reserved.

Abbreviations: COPD, chronic obstructive pulmonary disease; CBA, COPDbronchiectasis association; CBA-MS, COPD-bronchiectasis association with predominantly moderate-severe bronchiectasis; CBA-FH, COPD-bronchiectasis association characterized by frequent hospitalization within the previous year; CBA-BI, COPDbronchiectasis association with bacterial infection; CBA-NB, COPD-bronchiectasis association with infrequent hospitalization but no bacterial infection; CBA-NHB, COPD-bronchiectasis association with no hospitalization or bacterial infection in the past year; FVC% pred, forced vital capacity percentage predicted; FEV₁% pred, forced expiratory volume in one second percentage predicted; ICU, intensive care unit; CCI, Charlson Comorbidities Index; BSI, Bronchiectasis Severity Index; HRCT, high-resolution computed tomography.

* Corresponding author.

E-mail address: battery203@163.com (W.-j. Guan).

¹ These authors contributed equally to the work.

https://doi.org/10.1016/j.arbres.2024.04.003 0300-2896/© 2024 SEPAR. Published by Elsevier España, S.L.U. All rights reserved.

Introduction

Both bronchiectasis and chronic obstructive pulmonary disease (COPD) are common chronic airway inflammatory diseases with many shared manifestations such as chronic cough, sputum production and/or dyspnea.^{1,2} Bronchiectasis is defined as the persistent and irreversible dilation of bronchi, while COPD denotes the incompletely reversible airflow limitation associated with pollutant exposures.¹ Despite the notable heterogeneity of manifestations and pathophysiology, some patients reportedly had co-existing COPD and bronchiectasis – the COPD-bronchiectasis association (CBA).³ The prevalence of bronchiectasis ranged from 4.0% to 75.0% among COPD patients, and COPD from 8.8% to 32.0% among patients with bronchiectasis.³ This has been attributable to the inconsistency in diagnostic criteria and disease severity measures.

Addressing the knowledge gap in the clinical characteristics of CBA compared with bronchiectasis and COPD alone represents an important first step to improve the clinical management of CBA. Clinically, CBA could have originated from the progression of either bronchiectasis or COPD, thus representing the "hub" linking the aetiologically distinct disease entities.^{4,5} Apart from the differences in the underlying etiology, CBA may encompass distinct subgroups with diverse risk factor exposures, radiologic manifestations, lung function impairment and airway infections.^{6–8} However, there remains a paucity of medical evidence depicting the landscape of the diverse clinical characteristics of CBA among in-hospital adult patients.

We conducted a cross-sectional study of CBA, COPD, and bronchiectasis alone, comparing the clinical characteristics among the three groups from a large cohort in the hospital information system. Specifically, we explored the heterogeneity of CBA with unsupervised clustering approach^{9,10} via retrospective electronic medical record review spanning from 2.5 to 6.0 years, and documented the hospitalization outcomes. We also included an independent validation set of CBA for testing the robustness of the classifications. Unveiling the heterogeneity regarding the clinical characteristics and outcomes may prompt better personalized treatment and prognostic assessment of CBA.

Methods

Study participants

We included anonymized electronic medical records (EMRs) from inpatients aged 18-80 years hospitalized in The First Affiliated Hospital of Guangzhou Medical University (a tertiary hospital specialized in respiratory and critical care medicine) between July 2017 and December 2020 (training set), and between January 2021 and December 2022 (validation set) because of exacerbation onset. All hospitalized patients had a primary discharge physician-diagnosis of COPD (ICD-10-CM code: J44.), bronchiectasis (ICD-10-CM code: J47.) or both. Apart from extracting ICD-10 codes for diagnosis, we integrated chest high-resolution computed tomography (HRCT) and symptoms for confirming the diagnosis of bronchiectasis.¹ Furthermore, we integrated spirometric findings and clinical symptoms to jointly diagnose COPD.² We excluded patients with CBA without HRCT and spirometric results within 1 year (before and after hospitalization). The study was approved by ethics committee (ES-2023-K008-03) and additional informed consent was waived for the retrospective cohort.

Study design and data extraction

We retrospectively extracted EMRs from hospital information system, including demographics [age, gender, body-mass index (BMI), residency], medical history (disease duration, smoking history, number of exacerbations in the past year, etc.), major comorbidities (asthma, cor pulmonale, hypertension, diabetes, etc.), symptoms and signs (cough, phlegm, shortness of breath, chest pain, hemoptysis), HRCT findings (modified Reiff score, emphysema, etc.), spirometry [forced vital capacity percentage predicted (FVC% pred), forced expiratory volume in one second percentage predicted (FEV₁% pred) and FEV₁/FVC%], laboratory findings (blood routine test, sputum culture, C-reactive protein, total IgE, etc.), medications, clinical outcomes [length of stay, non-invasive ventilation, invasive ventilation, intensive care unit (ICU) admissions, re-admission, etc.]. We adopted the Charlson Comorbidities Index (CCI) for assessment of comorbidities, and the modified Reiff scores to evaluate the radiologic severity of bronchiectasis (mild: 1–6; moderate: 7–12; severe: 13–18 points). The Bronchiectasis Severity Index (BSI) was computed for patients with CBA. Because the modified Medical Research Council dyspnea scale was not recorded from EMR,¹¹ we estimated this scale based on patient's description of dyspnea in the present history as a proxy.

For the cross-sectional data analysis, we only included the findings of the first examination soon after admission among patients who underwent multiple examinations during the same hospitalization. For patients with missing data of HRCT and spirometry during hospitalization, we replaced these with the data from our study site within 1 year before and after hospitalization for subsequent analysis. We included only the single most complete hospital examination record for patients with multiple admissions within 1 year. We also conducted a longitudinal follow-up of CBA patients by capturing any of their subsequent electronic health records, extending until June 2023.

Variable selection

We collected 95 variables from the retrospective study (Fig. 1, Table S1, online supplement). We excluded variables with missing data among >15% of patients (N=29), no direct relevance to CBA classification (N=36) and medications and outcomes (N=13). For the remaining 17 variables, we performed multiple imputation for missing data with multi-level generalized linear regression model.¹² Next, we performed an exploratory factor analysis¹³ for continuous variables (Table S2) or system hierarchical clustering analysis¹⁴ for categorical variables (Fig. S1) with the remaining 17 variables, thereby leaving six variables in the clustering analysis. These included the modified Reiff score category (mild vs. moderate-to-severe), hospitalizations in the past year (0 vs. 1–2 vs. \geq 3 times), sputum culture (positive vs. negative), blood neutrophil count, blood eosinophil count, and the CCI (Fig. 2).

Clustering schemes

We employed a two-step clustering approach for patient categorization, beginning with the data compression into small subclusters to manage complex data sets effectively. This method utilized the log-likelihood distance for categorical data and the Euclidean distance for continuous variables to create dense subclusters. Following this step, an agglomerative hierarchical clustering method merged these subclusters based on the Ward's method. This has not only facilitated the efficient analysis of mixed data types but also automatically identified the most statistically significant cluster configuration, optimizing the handling of complex datasets.¹⁵ We then applied a decision-tree algorithm to all variables in clustering analysis to predict the specific cluster for individual patients, which allowed for calculation of the rates of misclassification.¹⁶

Validation set

We further included an independent CBA validation cohort (hospitalized between January, 2021 and December, 2022 and 348 patients who had overlapped with the training set have been excluded). By comparing the clinical characteristics of CBA in the discovery cohort with those in validation cohort, we sought to confirm the robustness of clustering findings.

Statistical analysis

We summarized categorical variables as frequencies and proportions. We tested continuous variables for the Gaussian distribution (Kolmogorov–Smirnov test) and expressed as



Fig. 1. Flowchart of patient recruitment and data analysis. Bx, bronchiectasis; COPD, Chronic Obstructive Pulmonary Disease; CBA, COPD-bronchiectasis association.



Fig. 2. Ranking of the contribution of variables for clustering analysis and the core variables. (A) Importance ranking of the clustering variables. Comparison of the grading of hospitalization in the past year (B), bronchiectasis severity (modified Reiff score category) (C), sputum culture findings (D), blood neutrophil (E), Charlson comorbidity index (F), blood eosinophil (G).

mean \pm standard deviation (SD) or median (interquartile range, IQR) as appropriate. We evaluated the differences between clusters for each variable using analysis-of-variance or Kruskal–Wallis test for continuous variables and Chi-square test or Fisher's exact test for categorical variables. Unless otherwise specified, we applied Bonferroni correction to two-group comparisons. We adopted multivariate Logistic regression models to analyze the risks of clinically significant events (e.g., ICU admission) during longitudinal follow-up, with the hazards ratio (HR) and 95% confidence interval (95%CI) being demonstrated. We defined statistical significance as two-sided *P* < 0.05. We conducted statistical analyses using SPSS (Version 25.0; IBM Corp., Armonk, NY).

Results

Patient recruitment in training set

We initially identified 6635 patients with bronchiectasis alone, 15,356 with COPD alone, and 2109 with CBA. After exclusion of repeated inpatient records, 4511 and 9320 patients had a diagnosis of bronchiectasis and COPD, respectively. The corresponding proportion of CBA was 29.9% in bronchiectasis and 14.5% in COPD, respectively. Finally, after excluding patients with malignant tumor, we included 2928 patients with bronchiectasis, 5158 with COPD, and 1219 with CBA (Fig. 1).

C.-x. Pan, Z.-f. He, S.-z. Lin et al.

Table 1

Demographic characteristics, medication, laboratory tests, hospitalization outcomes of patients with CBA, COPD, and bronchiectasis in the training set.

Baseline characteristics	COPD	CBA	Bronchiectasis	COPD vs. CBA	Bx vs. CBA
Ν	5158	1219	2928	-	-
Male, <i>N</i> (%)	4523 (87.7%)	1011 (82.9%)	1228 (41.9%)	<0.001	<0.001
Age (y), median (IQR)	66.0 (12.0)	67.0 (12.0)	58.0 (18.0)	<0.001	<0.001
BMI (kg/m ²), median (IQR)	21.5 (5.9)	21.2 (5.7)	20.7 (5.3)	0.037	0.004
Duration of symptoms (yrs), median (IQR)	8.0 (7.0)	10.0 (15.0)	8.0 (17.0)	<0.001	<0.001
Smoking history, N (%)	4724	1164	2520	-	-
Never smokers	891 (18.9%)	309 (26.5%)	1899 (75.4%)	<0.001	<0.001
Former smokers	1938 (41.0%)	507 (43.6%)	298 (11.8%)		
Current smokers	1895 (40.1%)	348 (29.9%)	323 (12.8%)		
CCI, median (IQR)	2.0 (1.0)	1.0 (1.0)	1.0 (1.0)	<0.001	<0.001
Previous PTB, N (%)	302 (5.9%)	143 (11.7%)	330 (11.3%)	<0.001	0.671
Peripheral blood cell count, median (IQR)					
Leukocytes (×10 ⁹ /L)	7.8 (3.8)	8.0 (4.3)	7.1 (3.5)	0.068	<0.001
Neutrophils (×10 ⁹ /L)	5.2 (3.8)	5.6 (4.3)	4.5 (3.3)	0.001	<0.001
Neutrophils (%)	68.5 (19.9)	71.4 (20.1)	64.5 (17.5)	<0.001	<0.001
Eosinophils ($\times 10^9/L$)	0.1 (0.3)	0.1 (0.3)	0.1 (0.1)	0.385	0.110
Eosinophils (%)	1.7 (3.5)	1.5 (3.4)	1.8 (2.5)	0.090	<0.001
Modified Reiff score	NA	4.0 (4.0)	5.0 (6.0)	NA	<0.001
FEV1% predicted, median (IQR)	47.4 (42.5)	39.2 (27.0)	72.1 (39.8)	<0.001	<0.001
Sputum culture, N (%)	1462 (28.3%)	740(60.7%)	954 (32.6%)		
Any bacteria	960 (65.7%)	381 (51.5%)	657 (68.9%)	<0.001	<0.001
Pseudomonas aeruginosa	172 (11.8%)	156 (21.1%)	368 (38.6%)	<0.001	<0.001
Medications, N (%)					
Intravenous corticosteroids	2361 (45.8%)	585 (48.0%)	295 (10.1%)	0.163	<0.001
Hospital stay (d), median (IQR)	7.0 (4.0)	7.0 (5.0)	6.0 (3.0)	0.016	<0.001
ICU admission, N (%)	368 (9.0%)	104 (8.7%)	151 (5.6%)	0.705	<0.001

COPD, Chronic Obstructive Pulmonary Disease; CBA, COPD-bronchiectasis association; Bx, bronchiectasis; IQR, interquartile range; SD, standard deviation; BMI, body-mass index; CCI, Charlson Comorbidity Index; PTB, pulmonary tuberculosis; FEV₁: forced expiratory volume in one second; FVC: forced vital capacity; ICU, intensive care unit; NA, not applicable.

Clinical characteristics of CBA in training set

Compared with COPD or bronchiectasis alone, CBA was consistently associated with a markedly greater age, longer disease duration, a higher proportion of patients with emphysema, higher blood neutrophil percentage, higher detection rate of *Pseudomonas aeruginosa* and more pronounced lung function impairment, evidenced by lower FEV_1/FVC , $FEV_1\%$ pred and FVC% pred (all P < 0.001).

The proportion of intravenous antibiotic use was significantly higher in CBA (P<0.001). CBA yielded non-significantly higher proportion of patients with comorbidities (e.g. pulmonary aspergillosis, cor pulmonale), a longer length of hospital stay, and more frequent receipt of invasive ventilation and ICU admission compared with COPD or bronchiectasis alone (Table 1, Tables S3 and S4).

Phenotypes of CBA in training set

Of the 1219 hospitalized patients with CBA, 572 were excluded due to missing spirometry or HRCT records. The clinical characteristics of 647 patients who were included in and those excluded from clustering analysis were comparable in terms of gender distribution (males: 82.5% vs. 83.4%), age (median: 68 vs. 67 years), BMI (median: 21.3 vs. 21.0 kg/m²), CCI (median: 1.0 vs. 1.0), and the proportion of patients admitted to ICU (8.2% vs. 9.3%) (all P > 0.05, Table S5).

We identified five clusters of CBA (Table 2, Tables S6 and S7). Based on the stepwise discriminant analysis, the most influential variables were ordered as hospitalization in the past year (1.000), bronchiectasis severity (0.745), sputum culture (0.494), neutrophil count (0.023), eosinophils count (0.001) and CCI (0.001) (Fig. 2).

Cluster 1 (CBA-MS, N=120) consisted of patients with predominantly moderate-to-severe bronchiectasis. Overall, CBA-MS exhibited more pronounced clinical symptoms and severe radiologic features of bronchiectasis. Apart from the highest Reiff score (median: 10.0), CBA-MS was characterized by the highest rate of cor pulmonale (30.8%), history of hemoptysis (22.5%) and detection of *P. aeruginosa* (39.1%). Regarding clinical outcomes, patients in CBA-MS yielded the highest rate of ICU admissions during hospitalization (14.2%).

Patients in cluster 2 (CBA-FH, N = 108) had frequent hospitalization (more than 2 times) due to exacerbations of bronchiectasis in the past year. All patients had mild bronchiectasis and poorer lung function (83.8% had FEV₁% predicted <50%: adjusted P<0.001). Patients with CBA-AE had a highest proportion of corticosteroids use (oral corticosteroids accounting for 40.7%, adjusted P=0.042), and the re-admission rate was highest (67.6%, adjusted P<0.001) among the five clusters.

Patients in cluster 3 (CBA-BI, N = 163) consistently had bacterial infection as determined with sputum bacterial culture. By contrast, patients in cluster 4 (CBA-NB, N = 143) had on average 1–2 hospitalizations per year but no bacterial infection The clinical characteristics of CBA-BI were overall similar with those of CBA-NB, except that CBA-BI had a higher proportion of patients with C-reactive protein >0.6 mg/dl (86.5% vs. 64.2%) and antibiotics use (35.0% vs. 28.0%) (both adjusted P < 0.05).

Cluster 5 (CBA-NHB, N=113) comprised patients with no bacterial infection or hospitalization in the past year. CBA-NHB yielded the lowest proportion of a history of cor pulmonale (5.3%) and procalcitonin > 0.05 ng/mL (29.9%) and consistently had mild bronchiectasis (100%). Although the rate of ICU admission (1.8%) and readmission within the past year (1.2%) was lowest in CBA-NB, the use of non-invasive (25.7%) and invasive ventilation (4.4%) was nominally lower than CBA-MS only.

Decision-tree analysis of training and validation set

We employed a decision-tree model to predict the cluster assignment (Fig. 3). Based on the three cardinal variables, bronchiectasis severity in HRCT, hospitalization in the past year and sputum culture, 91.8% (95% CI, 89.70–93.92%) of patients could be assigned to the correct cluster. Our model demonstrated a strong

Table 2

Demographic and clinical characteristics of five clusters of inpatients with CBA in the training set.

Baseline characteristics	Total	CBA-MS	CBA-FH	CBA-BI	CBA-NB	CBA-NHB	P value
N Male, N (%) Age (y), median (IQR) BMI (kg/m ²), mean ± SD	$\begin{array}{c} 647 \\ 534(82.5\%) \\ 68.0(11.0) \\ 21.6 \pm 3.8 \end{array}$	120 91(75.8%) 64.0(10.0) 20.6±3.6	108 92(85.2%) 71.0(12.8) [#] 21.6 \pm 4.3	163 136(83.4%) 68.0(11.0) [#] 22.4 ± 3.8 [#]	143 119(83.2%) 68.0(11.0) 21.1±3.4 ^{&}	$11396(85.0%)68.0(11.5)21.9 \pm 3.9^{\#}$	- 0.302 <0.001 0.001
Smoking history, N (%) Former smokers Smokers	647 299(46.2%) 196(30.3%)	- 53(44.2%) 23(19.2%)	- 62(57.4%) 32(29.6%)	- 76(46.6%) 52(31.9%)	- 67 (46.9%) 45 (31.5%)	- 41 (36.3%) 44 (38.9%) [#]	<0.001
Duration of symptom (yrs), median (IQR)	10.0 (9.0)	13.0 (17.5)	10.0 (12.0)	10.0 (7.0)#	9.0 (8.0)#\$	9.0 (5.0) ^{#,\$}	<0.001
Hospitalization in the past year 1−2 times ≥3 times	316(49.1%) 131(20.4%)	69(57.5%) 23(19.2%)	0 (0%) [#] 108 (100%)	105 (65.2%) ^{\$} 0 (0%)	142 (100%) ^{#,\$} 0 (0%)	0 (0%) [#] 0 (0%)	<0.001
CCI, median (IQR)	1.0 (1.0)	1.0 (1.0)	1.0 (1.0)	1.0(1.0)	1.0 (1.0)	1.0 (1.0)	0.650
<i>Comorbidities, N (%)</i> Cor pulmonale Respiratory failure	109(16.8%) 137(21.2%)	37(30.8%) 41(34.2%)	18(16.7%) 22(20.4%)	27 (16.6%) [#] 37 (22.7%)	21 (14.7%) 22 (15.4%)	6 (5.3%) ^{#,&} 15 (13.3%)	- <0.001 0.001
Hemoptysis, N (%)	70(10.8%)	27(22.5%)	6 (6.1%)#	10 (6.1%)#	14 (9.8%)#	13(11.5%)	<0.001
Peripheral blood cell count Leukocytes ($\times 10^9$ /L), median (IQR) Neutrophils ($\times 10^9$ /L), median (IQR) Neutrophils (%), means \pm SD Eosinophils ($\times 10^9$ /L), median (IQR) Eosinophils (%), median (IQR)	$\begin{array}{c} 8.1 \ (4.2) \\ 5.5 \ (4.0) \\ 70.2 \pm 12.8 \\ 0.1 \ (0.3) \\ 1.7 \ (3.6) \end{array}$	8.8 (4.9) 6.7 (4.1) 72.7 ± 13.6 0.1 (0.2) 1.2 (3.1)	$\begin{array}{c} 8.6 \ (5.3) \\ 6.3 \ (5.0) \\ 73.0 \pm 12.6 \\ 0.1 \ (0.2) \\ 1.2 \ (3.3) \end{array}$	7.6 (4.3) 5.1 (4.3) [#] 68.6 ± 12.4 0.1 (0.2) 1.8 (2.9)	$7.9 (3.7) 5.2 (3.5) 70.0 \pm 12.5 0.2 (0.3) 2.1 (3.7)$	7.4 (2.9) ^{#,\$} 4.9 (2.5) ^{#,\$} 67.0 ± 12.0 ^{#,\$} 0.2 (0.2) 2.6 (3.5)	- 0.004 <0.001 0.002 0.705 0.222
FEV1% predicted <50%, N (%)	459(70.9%)	97(80.8%)	90(83.3%)	111(68.1%)	88 (61.5%) ^{#,\$}	73 (64.6%) ^{\$}	<0.001
Bronchiectasis severity category, N (%) Mild Moderate-severe	528(81.6%) 119(18.4%)	1(0.8%) 119(99.2%)	108 (100%) [#] 0 (0%)	163 (100%) [#] 0 (0%)	143 (100%) [#] 0 (0%)	113 (100%) [#] 0 (0%)	_ <0.001
Sputum culture positive, N (%)	239(42.8%)	68(61.8%)	60(57.7%)	111 (100%) ^{#,\$}	0 (0%) ^{#,\$,&}	0 (0%) ^{#,\$,&}	<0.001
Pseudomonas aeruginosa, N (%) Other PPMs, N (%)	92(16.5%) 171(30.6%)	43(39.1%) 35(31.8%)	17 (16.3%) [#] 50(48.1%)	32(28.8%) 86(77.5%) ^{#,\$}	0 (0%) ^{#,\$,&} 0 (0%) ^{#,\$,&}	0 (0%) ^{#,\$,&} 0 (0%) ^{#,\$,&}	<0.001 <0.001
BSI score, median (IQR)	10.0 (6.0)	11.0 (5.5)	14.0 (2.0)#	10.0 (6.0)#,\$	10.0 (3.0)#,\$	6.0 (2.0) ^{#,\$,&,*}	<0.001
In-hospital outcomes Hospital stay (d), median (IQR) ICU admission, N (%)	7.0(4.0) 53(8.2%)	7.0 (4.0) 17(14.2%)	7.0 (3.0) 9 (8.3%)	7.0 (4.0) 18 (11.0%)	7.0 (3.0) 7 (4.9%)	7.0 (3.0) 2 (1.8%) ^{#,&}	0.047 0.003
Re-admission within 1 year, N (%)	180(31.3%)	30(26.5%)	75 (71.4%)#	32 (21.9%) ^{\$}	42 (33.3%) ^{\$}	1 (1.2%)#,\$,&,*	<0.001

IQR, interquartile range; SD, standard deviation; BMI, body mass index; CCI, Charlson Comorbidity Index; PTB, pulmonary tuberculosis. FEV1: forced expiratory volume in one second; FVC, forced vital capacity; HRCT, high-resolution computed tomography; CRP, C-reactive protein; OCS, oral corticosteroids; LABA, long-acting beta 2 agonist; LAMA: long-acting muscarinic antagonist; BSI, bronchiectasis severity index; ICU, intensive care unit.

P<0.05 vs cluster 1, with Bonferroni correction.

P < 0.05 vs cluster 2, with Bonferroni correction.

& P < 0.05 vs cluster 3, with Bonferroni correction.

* P < 0.05 vs cluster 4. with Bonferroni correction.

predictive accuracy, effectively identifying the true positives, as indicated by the high sensitivity (97.5% [95% CI, 96.26–98.68%]) and positive predictive values (96.0% [95% CI, 94.50–97.52%]). However, the specificity (55.6% [95% CI, 51.73–59.39%]) and negative predictive values (66.7% [95% CI, 63.04–70.30%]) revealed some further room for improvement in classifying the true negatives accurately. The moderate Matthews correlation coefficient (0.58 [95% CI, 0.45–0.69]) has highlighted the model's overall balanced performance in classification tasks (Table S8).

Next, we included 218 inpatients with CBA (validation set) for validation of the clustering findings of the decision-tree model. Compared with the training set, there were no significant differences in the baseline distribution in gender, age, BMI, CCI and duration of symptoms (all P > 0.05) (Fig. 1, Table S9). Five clusters were classified by applying the discriminant analysis findings of the decision-tree model. Comparison of clusters of the training set and validation set revealed no significant differences in the CCI, proportion of emphysema, FEV₁% pred < 50%, neutrophil ratio, length of hospital stay and the use of invasive ventilation (all P > 0.05, Table S10).

Prognosis of CBA in training set

We analyzed the risks of oxygen therapy, non-invasive ventilation, mechanical ventilation and ICU admission with logistic regression after adjusting for the age, gender, BMI and CCI. Compared with CBA-NHB, CBA-MS (OR: 8.87, 95%CI: 1.98, 39.62, adjusted P=0.004), CBA-BI (OR: 6.87, 95%CI: 1.56, 30.38, adjusted P=0.011) and CBA-FH (OR: 4.81, 95%CI: 1.01, 22.91, adjusted P=0.048) consistently had a markedly higher risk of ICU admission (Fig. S2). We next employed Cox proportional regression analysis to investigate the risk of readmissions during longitudinal follow-up, which differed considerably among the clusters. Compared with CBA-NHB, CBA-FH (HR: 49.86, 95%CI: 18.24, 136.26, adjusted P<0.001), CBA-NB (HR: 19.18, 95%CI: 7.02, 52.42, adjusted P<0.001), CBA-MS (HR: 15.04, 95%CI: 5.43, 41.69, adjusted *P*<0.001) and CBA-BI (HR: 10.18, 95%CI: 3.70, 28.06, adjusted P < 0.001) consistently exhibited a significantly higher risk of readmission of follow-up (Fig. 4A, Table S11). We also conducted receiver operation characteristics analysis to assess the risk of readmission within one years and two years of follow-up, the



Fig. 3. Scheme of predicting clinical cluster assignment and the major clinical characteristics of the clusters. Patients were assigned to the five clusters by using three core variables: bronchiectasis severity, hospitalization in the past year, and sputum culture (upper panel). The lower panel demonstrates the gradings for the five major variables or dimensions – lung function impairment, blood neutrophil count, detection of *P. aeruginosa* from sputum, ICU admission, readmission within 1 year of follow-up. The downward arrows indicate the decreased trend as compared with the normal levels, the plus signs denote the magnitude of increase or the likelihood of positivity as compared with the absence, the normal levels, and the minus signs signal the absence of the findings. ICU; intensive care unit.



Fig. 4. Longitudinal risk assessment and cluster membership dynamics of CBA. Panel A: Cox proportional hazards regression analysis demonstrating the risk for readmission over a longitudinal follow-up of 2.5–6 years, using CBA-NHB as the reference cluster, with adjustments for the gender, age, BMI and CCI (Training set). Panel B: Sankey plot illustrating the changes in the cluster membership over time during longitudinal follow-up (Training and validation set). The left side of the diagram represents the former cluster of patients with clusters labeled as CBA-BI (bacterial infection), CBA-FH (frequent hospitalizations), CBA-MS (moderate to severe bronchiectasis), CBA-NB (non-bacterial), and CBA-NHB (no hospitalizations or bacterial infection). The right side represents the re-evaluated clustering as per the rules of the decision-tree analysis, with the addition of 'new' to the original cluster labels—CBA-BI-new, CBA-FH-new, CBA-MS-new, CBA-NB-new, and CBA-NHB-new—the reassessment and potential shift in each patient's cluster membership. BMI, body mass index; CCI, Charlson Comorbidity Index; CI, confidence interval; HR adj, hazard ratio after adjusting for gender, age, BMI and CCI.

clustering grading (1 = NHB, 2 = BI, 3 = MS, 4 = NB, 5 = FH) demonstrated a significant predictive capability (area under curve [95%CI]: 0.757 [0.716, 0.799] at 1 year; 0.764 [0.726, 0.802] at 2 years, both P < 0.001, Fig. S3).

Characteristics and outcomes of patients with readmission

We identified 170 patients with CBA who had repeated hospitalization records within our study site (mean interval: 623.0 days). Overall, the clinical characteristics remained stable – no significant differences in blood routine test items, lung function metrics, bacterial culture findings, blood inflammatory markers, use of corticosteroids, and hospitalization outcomes (including invasive ventilation, length of stay, re-admission within 1 year, etc.). (all P > 0.05, Table S12).

The decision-tree model for patients with re-admissions revealed that 31.8% of patients changed their cluster membership. Using the former hospitalization as the control event, Logistic regression analysis did not identify any significant differences in the risk of re-admissions between the former and the subsequent hospitalization. However, the Sankey diagram showed certain cluster membership changes which centered on CBA-FH, CBA-BI and CBA-NB (Fig. 4B). Regression analysis on these clusters revealed that CBA-FH exhibited a significant increase in the risk of hospitalization among patients whose cluster membership remained unchanged (HR: 10.57, 95%CI: 2.94, 38.09, P<0.001) (Table S13).

Discussion

Based on the hospital EMRs (span: 2.5–6.0 years), we have compared the clinical characteristics of a reasonably large cohort of hospitalized patients with CBA, and bronchiectasis and COPD alone. Apart from the greater symptom burden and poorer clinical outcomes of CBA, we have identified five phenotypes of CBA (mostly based on bronchiectasis severity, hospitalization in the past year and sputum culture findings) which differed in clinical manifestations and outcomes.

Our findings pertaining to the proportion of patients with CBA – 14.5% of COPD patients and 29.9% of bronchiectasis patients – mirrored the published reports.³ Compared with bronchiectasis alone, our study indicated a significantly greater disease severity (greater symptom burden, prior exacerbation frequency, lung function impairment and detection of *P. aeruginosa*) and poorer outcomes (more frequent use of ventilators and higher rate of ICU admission) of CBA, which echoed previous studies.^{6,17–21} This might be partly interpreted by the greater magnitude of systemic (heightened neutrophilic inflammation) and airway inflammation (increased levels of C-reactive protein) in CBA. Hence, the clinical management of CBA may benefit more intensive treatment to reduce the disease burden and improve clinical outcomes.

Having unraveled the unique features of CBA compared with bronchiectasis or COPD alone, we next sought to decipher the heterogeneity of CBA. We phenotyped CBA with unsupervised clustering analysis of core metrics that best defined the clinical characteristics. The importance ranking of indicators was based on their predictive capacity for disease severity and therapeutic response.²²⁻²⁵ Congruent with previous findings, the radiologic severity of bronchiectasis was a crucial indicator for predicting the disease severity and therapeutic intensity.^{23,26,27} The history of hospitalizations was a proxy for the frequency of bronchiectasis or COPD exacerbations, an important factor for assessing future exacerbation risks.^{18,28–30} Sputum culture might indicate current airway infections and partially correlate with recent antibiotic treatment.^{28,31,32} Here, we also evaluated the diagnostic efficacy of CBA phenotypes in predicting hospital readmission risks at longitudinal follow-up.¹¹ The CBA-FH cluster demonstrated a readmission risk that was nearly four-fold higher (390% increase) relative to the CBA-BI cluster. The CBA-MS and CBA-NB clusters also exhibited elevated risks (48% and 88% increase, respectively). These pronounced discrepancies provided crucial evidence for dissecting the heterogeneity of CBA and the identification of its distinct phenotypes. The algorithms for phenotyping CBA have underscored the importance of rapidly identifying different CBA phenotypes. For instance, both CBA-MS and CBA-FH might benefit from more intensified management strategies (including medication therapy, pulmonary rehabilitation training, intensified follow-up assessment, etc.). In light of the greatest disease burden, CBA-MS would warrant intensified airway clearance and bronchodilator therapy and other therapeutic measures to reduce the risk of future exacerbations. Conversely, patients with CBA-BI might require targeted antibiotic therapy for specific bacteria and management of the coexisting conditions. In clinical practice, our proposed clustering scheme aids in formulating more effective treatment plans for CBA. To illustrate these findings more intuitively, we have comprehensively presented the key characteristics of each phenotype in Fig. 3. This figure visualization provides the rationale for clarifying the heterogeneity of CBA so that clinicians could identify any treatable trait to improve the decision making and treatment planning within the previously neglected subgroups of patients.

Some limitations should also be considered. First, we only included hospitalized patients from a single tertiary medical center, nor did we validate findings with an independent out-patient cohort because it differed considerably in clinical characteristics (significantly more severe bronchiectasis in inpatient cohort vs. COPD as the primary diagnosis in outpatient cohort). This may limit the generalizability of our findings across the broader clinical settings, underscoring the need for cautious interpretation when applying our results to different patient populations. Second, we did not dissect asthma from our study population because the primary diagnoses of COPD and bronchiectasis were made based on both discharge physician diagnosis and verification of other compelling records, contingent on the clinical conditions fully satisfying the guideline-defined diagnostic criteria. Third, because some key indicators were missing in <15% of patients, we conducted multiple imputation for estimating the missing values. However, there were no significant differences in demographic characteristics and other important metrics associated with disease severity between patients excluded from and included in the final analysis. Fourth, we did not integrate the grading of mucus plugging because of the lack of a standardized quantitative imaging system, which will be explored in future studies. Finally, because of the nature of retrospective in-hospital EMR review, the use of medications (esp. corticosteroids, antibiotics) could both reflect physician's adherence to guidelines and the practical needs of individual patients.

Conclusions

We have demonstrated that bronchiectasis severity, hospitalizations and sputum culture are core indicators of clustering membership of CBA. These revealed some clusters with significantly greater symptom burden and poorer clinical outcomes, hence justifying personalized management of CBA among the targeted patient subpopulations.

Author's contributions

C.X.P., J.Q.Y. and W.J.G. drafted the manuscript; C.X.P., Z.F.H., Z.M.C. and W.J.G. contributed to conception and design; C.X.P., Z.F.H., S.Z.L., J.Q.Y. and Z.M.C. were responsible for patient recruitment and data collection; C.X.P. and S.Z.L. performed data analysis; W.J.G. critically revised the manuscript. All authors have approved the final submission.

Ethical approval

The study was approved by the Institutional Review Board of The First Affiliated Hospital of Guangzhou Medical University (ES-2023-K008-03), and additional informed consent was waived for the retrospective cohort.

Data availability statement

Data will be available from corresponding authors through reasonable requirement.

Funding

This study was funded by the National Natural Science Foundation – Outstanding Youth Foundation (No. 82222001), R&D Program of Guangzhou National Laboratory (GZNL2024A02003), Guangdong Natural Science Foundation (No. 2019A1515011634), and the Zhongnanshan Medical Foundation of Guangdong Province (No. ZNSA-2020013), Guangzhou Science and Technology Plans (No. 2023B03J0407 and 202102010372), Plan on Enhancing Scientific Research in Guangzhou Medical University (grant number not applicable) (to Prof. Guan).

Conflict of interest

None declared.

Acknowledgement

We thank Mei Jiang (The First Affiliated Hospital of Guangzhou Medical University) for her guidance on statistical methods used in this article. We also thank Wen-hua Jian and Yi-jun Chen (department of big data center, National Clinical Research Center for Respiratory Disease, The First Affiliated Hospital of Guangzhou Medical University) for their help with electronic medical record data extraction. Fig. 3 was created with BioRender.com (Agreement number: PP2673TD0X).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.arbres.2024.04.003.

References

- 1. Hill AT, Sullivan AL, Chalmers JD, De Soyza A, Elborn SJ, Floto AR, et al. British Thoracic Society Guideline for bronchiectasis in adults. Thorax. 2019;74 Suppl. 1:1–69, http://dx.doi.org/10.1136/thoraxjnl-2018-212463.
- Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease report 2023. Available from: https://goldcopd.org/2023-gold-report-2/ [accessed 10.12.23].
- Traversi L, Miravitlles M, Martinez-Garcia MA, Shteinberg M, Bossios A, Dimakou K, et al. ROSE: radiology, obstruction, symptoms and exposure – a Delphi consensus definition of the association of COPD and bronchiectasis by the EMBARC Airways Working Group. ERJ Open Res. 2021;7, http://dx.doi.org/10.1183/23120541.00399-2021.
- Du Q, Jin J, Liu X, Sun Y. Bronchiectasis as a comorbidity of chronic obstructive pulmonary disease: a systematic review and meta-analysis. PLOS ONE. 2016;11:e0150532, http://dx.doi.org/10.1371/journal.pone.0150532.
- Martinez-Garcia MA, Miravitlles M. Bronchiectasis in COPD patients: more than a comorbidity? Int J Chron Obstruct Pulmon Dis. 2017;12:1401–11, http://dx.doi.org/10.2147/COPD.S132961.
- Hsu C-Y, Poon Y-Y, Chen Y-W, Hsieh M-H, Lin H-C, Fang W-F. Chronic obstructive pulmonary disease as a phenotype of bronchiectasis for longterm clinical presentation and treatment. Medicina (Kaunas). 2021;57, http://dx.doi.org/10.3390/medicina57060579.
- Bohn O, Xi M, Woodruff NK, Hansen GL, McEvoy CE. Chronic bronchitis in COPD patients creates worse symptom burden regardless of the presence of bronchiectasis in the COPDGene cohort. Chronic Obstr Pulm Dis. 2021;8:350–9, http://dx.doi.org/10.15326/jcopdf.2021.0202.
- Polverino E, Dimakou K, Hurst J, Martinez-Garcia M-A, Miravitlles M, Paggiaro P, et al. The overlap between bronchiectasis and chronic airway diseases: state of the art and future directions. Eur Respir J. 2018;52, http://dx.doi.org/10.1183/13993003.00328-2018.
- Aliberti S, Lonni S, Dore S, McDonnell MJ, Goeminne PC, Dimakou K, et al. Clinical phenotypes in adult patients with bronchiectasis. Eur Respir J. 2016;47:1113–22, http://dx.doi.org/10.1183/13993003.01899-2015.
- Guan WJ, Jiang M, Gao YH, Li HM, Xu G, Zheng JP, et al. Unsupervised learning technique identifies bronchiectasis phenotypes with distinct clinical characteristics. Int J Tuberc Lung Dis. 2016;20:402–10, http://dx.doi.org/10.5588/ijtld.15.0500.
- Chalmers JD, Goeminne P, Aliberti S, McDonnell MJ, Lonni S, Davidson J, et al. The bronchiectasis severity index. An international derivation and validation study. Am J Respir Crit Care Med. 2014;189:576–85, http://dx.doi.org/10.1164/rccm.201309-1575OC.
- Qiu R, Xie J, Chung KF, Li N, Yang Z, He M, et al. Asthma phenotypes defined from parameters obtained during recovery from a hospitaltreated exacerbation. J Allergy Clin Immunol Pract. 2018;6:1960–7, http://dx.doi.org/10.1016/j.jaip.2018.02.012.
- Schatz M, Hsu JW, Zeiger RS, Chen W, Dorenbaum A, Chipps BE, et al. Phenotypes determined by cluster analysis in severe or difficult-to-treat asthma. J Allergy Clin Immunol. 2014;133:1549–56, http://dx.doi.org/10.1016/j.jaci.2013.10.006.
- Sendín-Hernández MP, Ávila-Zarza C, Sanz C, García-Sánchez A, Marcos-Vadillo E, Muñoz-Bellido FJ, et al. Cluster analysis identifies 3 phe-

notypes within allergic asthma. J Allergy Clin Immunol Pract. 2018;6, http://dx.doi.org/10.1016/j.jaip.2017.10.006.

- Zhang X, Zhang L, Wang G, Feng M, Liang R, McDonald VM, et al. Clinical phenotypes of patients hospitalized for an asthma exacerbation: prognostic implications. J Allergy Clin Immunol Pract. 2021;9, http://dx.doi.org/10.1016/j.jaip.2020.09.031.
- Pikoula M, Quint JK, Nissen F, Hemingway H, Smeeth L, Denaxas S. Identifying clinically important COPD sub-types using data-driven approaches in primary care population based electronic health records. BMC Med Inform Decis Mak. 2019;19:86, http://dx.doi.org/10.1186/s12911-019-0805-0.
- 17. Lei J, Yang T, Liang C, Huang K, Wu S, Wang C. Comparison of clinical characteristics and short-term prognoses within hospitalized chronic obstructive pulmonary disease patients comorbid with asthma, bronchiectasis, and their overlaps: findings from the ACURE registry. Front Med (Lausanne). 2022;9:817048, http://dx.doi.org/10.3389/fmed.2022.817048.
- Kim Y, Kim K, Rhee CK, Ra SW. Increased hospitalizations and economic burden in COPD with bronchiectasis: a nationwide representative study. Sci Rep. 2022;12:3829, http://dx.doi.org/10.1038/s41598-022-07772-6.
- Seifer FD, Hansen G, Weycker D. Health-care utilization and expenditures among patients with comorbid bronchiectasis and chronic obstructive pulmonary disease in US clinical practice. Chron Respir Dis. 2019;16, http://dx.doi.org/10.1177/1479973119839961, 1479973119839961.
- Sobala R, De Soyza A. Bronchiectasis and chronic obstructive pulmonary disease overlap syndrome. Clin Chest Med. 2022;43:61–70, http://dx.doi.org/10.1016/j.ccm.2021.11.005.
- Ni Y, Shi G, Yu Y, Hao J, Chen T, Song H. Clinical characteristics of patients with chronic obstructive pulmonary disease with comorbid bronchiectasis: a systemic review and meta-analysis. Int J Chron Obstruct Pulmon Dis. 2015;10:1465–75, http://dx.doi.org/10.2147/copd.S83910.
- Lau KY, Ng KS, Kwok KW, Tsia KK, Sin CF, Lam CW, et al. An unsupervised machine learning clustering and prediction of differential clinical phenotypes of COVID-19 patients based on blood tests – a Hong Kong population study. Front Med (Lausanne). 2021;8:764934, http://dx.doi.org/10.3389/fmed.2021.764934.
- Abo-Leyah H, Browne A, Carreto L, Finch S, Chalmers J. The relationship between radiological severity and clinical phenotype in bronchiectasis. Eur Respir J. 2019;54 Suppl. 63:PA2878, http://dx.doi.org/10.1183/13993003.congress-2019.PA2878.
- Diaz AA, Young TP, Maselli DJ, Martinez CH, Gill R, Nardelli P, et al. Quantitative CT measures of bronchiectasis in smokers. Chest. 2017;151:1255–62, http://dx.doi.org/10.1016/j.chest.2016.11.024.
- Pakzad A, Jacob J. Radiology of bronchiectasis. Clin Chest Med. 2022;43:47–60, http://dx.doi.org/10.1016/j.ccm.2021.11.004.
 Habesoglu MA, Tercan F, Ozkan U, Fusun EO. Effect of radiological extent
- Habesoglu MA, Tercan F, Ozkan U, Fusun EO. Effect of radiological extent and severity of bronchiectasis on pulmonary function. Multidiscip Respir Med. 2011;6:284–90, http://dx.doi.org/10.1186/2049-6958-6-5-284.
- Kim SH, Yang B, Yoo JY, Cho JY, Kang H, Shin YM, et al. Clinical characteristics, radiological features, and disease severity of bronchiectasis according to the spirometric pattern. Sci Rep. 2022;12:13167, http://dx.doi.org/10.1038/s41598-022-17085-3.
- Zhang X, Pang L, Lv X, Zhang H. Risk factors for bronchiectasis in patients with chronic obstructive pulmonary disease: a systematic review and meta-analysis. Clinics (Sao Paulo). 2021;76:e2420, http://dx.doi.org/10.6061/clinics/2021/e2420.
- Yu Q, Peng H, Li B, Qian H, Zhang H. Characteristics and related factors of bronchiectasis in chronic obstructive pulmonary disease. Medicine (Baltimore). 2019;98:e17893, http://dx.doi.org/10.1097/md.000000000017893.
- Henkle E, Chan B, Curtis JR, Aksamit TR, Daley CL, Winthrop KL. Characteristics and health-care utilization history of patients with bronchiectasis in US medicare enrollees with prescription drug plans, 2006–2014. Chest. 2018;154:1311–20, http://dx.doi.org/10.1016/j.chest2018.07.014.
- Dicker AJ, Lonergan M, Keir HR, Smith AH, Pollock J, Finch S, et al. The sputum microbiome and clinical outcomes in patients with bronchiectasis: a prospective observational study. Lancet Respir Med. 2021;9:885–96, http://dx.doi.org/10.1016/s2213-2600(20)30557-9.
- Kaehne A, Milan SJ, Felix LM, Sheridan E, Marsden PA, Spencer S. Headto-head trials of antibiotics for bronchiectasis. Cochrane Database Syst Rev. 2018;9:Cd012590, http://dx.doi.org/10.1002/14651858.CD012590.pub2.