Clinical Subtypes of Neutrophilic Asthma: A Cluster Analysis From Australasian Severe Asthma Network

Background
Clinical heterogeneity may exist within asthma subtypes defined by inflammatory markers. However, the heterogeneity of neutrophilic asthma (NA) remains largely unexplored.

Objective
To explore potential clusters and the stability of NA.

Methods
Participants with NA from the Australasian Severe Asthma Network underwent a multidimensional assessment. They were then asked to participate in a 12-month longitudinal cohort study. We explored potential clusters using a hierarchical cluster analysis and validated the differential future risk of asthma exacerbations in the identified clusters. A decision tree analysis was developed to predict cluster assignments. Finally, the stability of prespecified clusters was examined within 1 month.

Results
Three clusters were identified in 149 patients with NA. Cluster 1 (n = 99; 66.4%) was characterized by female-predominant nonsmokers with well-controlled NA, cluster 2 (n = 16; 10.7%) by individuals with comorbid anxiety/depressive symptoms with poorly controlled NA, and cluster 3 by older male smokers with late-onset NA. Cluster 2 had a greater proportion of participants with severe exacerbations (P = .005), hospitalization (P = .010), and unscheduled visits (P = .013) and a higher number of emergency room visits (P = .039) than that of the other two clusters. The decision tree assigned 92.6% of participants correctly. Most participants (87.5%; n = 7) in cluster 2 had a stable NA phenotype, whereas participants of clusters 1 and 3 had variable phenotypes.

Conclusions
We identified three clinical clusters of NA, in which cluster 2 represents an uncontrolled and stable NA subtype with an elevated risk of exacerbations. These findings have clinical implications for the management of NA.

Key words
Asthma
Neutrophilic asthma
Heterogeneity
Phenotypes
Cluster analysis
Phenotype variability

Abbreviations used
ACO
Asthma–chronic obstructive pulmonary disease overlap
ACT
Asthma Control Test
COPD
Chronic obstructive pulmonary disease
EA

Eosinophilic asthma
GINA
Global Initiative for Asthma
ICS
Inhaled corticosteroid
MGA
Mixed granulocytic asthma
NA
Neutrophilic asthma
PGA
Paucigranulocytic asthma

Introduction

Asthma is a heterogeneous disease with multiple phenotypes and diverse clinical characteristics.1,2 The identification of different phenotypes has improved our understanding of asthma pathogenesis and led to more targeted and personalized approaches to asthma management.3 Asthma can be classified into four inflammatory phenotypes based on the proportions of eosinophils and neutrophils in the sputum: eosinophilic asthma (EA), neutrophilic asthma (NA), mixed granulocytic asthma (MGA), and paucigranulocytic asthma (PGA).4Eosinophilic asthma has been demonstrated to be heterogeneous, with allergic and nonallergic subtypes or clusters, such as individuals with the lowest blood eosinophil levels and modest severe asthma, male smokers with severe persistent airway obstruction and moderate eosinophilia, and individuals with the highest blood eosinophil levels and good treatment response.5, 6, 7, 8Furthermore, our recently published study9 demonstrated that PGA is also a heterogeneous inflammatory phenotype and can be further divided into three subtypes: mild PGA, PGA with psychological dysfunction and rhinoconjunctivitisand other allergic diseases, and smoking-associated PGA. These clinical asthma clusters based on a primary inflammatory subtype are useful for understanding the heterogeneity and complexity of asthma and raise the question of whether NA is similarly heterogeneous. Neutrophilic asthma accounts for 20% to 30% of all asthma cases and may represent up to 50% of severe asthma cases resistant to corticosteroid therapy.10,11 Neutrophilic asthma is characterized by cigarette smoking, increased airflow obstruction, frequent exacerbations, glucocorticoid resistance, and increased sputum levels of IL-1β and IL-8.12 Furthermore, numerous demographic characteristics and clinical features such as age, cigarette smoking, obesity, and comorbidities such as respiratory infection and insulin resistance are associated with airway neutrophilia.10,13, 14, 15 These studies imply that NA may not be a homogeneous phenotype and may be composed of different subtypes. However, the heterogeneity of NA remains largely unexplored. Therefore, insights into the heterogeneity of the NA phenotype can result in a better understanding of NA, leading to improved asthma outcomes.

In this study, we hypothesized that patients with NA could be classified into several clusters based on distinct physiologic and clinical features. To investigate the heterogeneity of NA in well-characterized patients with asthma and identify unique

clusters of patients with NA, we performed unsupervised hierarchical cluster analysis. We then observed the stability of NA in these identified clusters, which may have potential implications for the individualized management of asthma in clinical practice.

Methods

Study design

This study was based on a prospective cohort selected from a Chinese population from the Australasian Severe Asthma Network16,17 and consisted of four individual components (Figure 1). Part I was a cross-sectional study evaluating patients with NA at baseline to identify clusters using unsupervised hierarchical clustering analysis. Part II was a cohort study in which patients in part I were longitudinally observed for 12 months to assess the future risk of asthma exacerbation. Part III was a cross-sectional study that tested the identified clusters using decision tree analysis from an independent population. Part IV was a post hoc analysis to investigate the stability of airway neutrophilic inflammation over 1 month among the identified clusters in part I (Figure 1). Based on the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis Statement,18,19 patients with NA were divided into training (part I) and validation sets (part III) in a split ratio of 7:3 using the random split-sample method,20, 21, 22 with the training set used to explore the heterogeneity of NA and the independent validation set used to evaluate its robustness. The case report form used in this study was based on the Australasian Severe Asthma Network.17 This study was approved by the Institutional Review Board of West China Hospital, Sichuan University, Chengdu, China (No. 2014-30). All participants provided written informed consent.

Figure 1. Flowchart of participants with asthma in the current study. ASAN, Australasian Severe Asthma Network; NA, neutrophilic asthma.

Participants

Participants with asthma, which was confirmed by diagnostic features of a history of both variable respiratory symptoms and airflow obstruction or airway hyperresponsiveness, who were aged 18 years or more, were consecutively recruited from the asthma clinic of West China Hospital, Sichuan University from March 2014 to June 2020. Detailed inclusion and exclusion criteria for participants can be found in our previously published study.17 All participants included in part I were asked to participate in a longitudinal cohort study and were followed up at 1, 3, 6, 9, and 12 months. For part IV, patients who completed visits 1 (baseline visit) and 2 (1-month visit) and received 1-month stable antiasthma therapy were enrolled.

Definition of inflammatory phenotypes

All enrolled participants with stable asthma at baseline were classified into the following four inflammatory phenotypes according to inflammatory cell counts in induced sputum, as described in our previous studies: NA (neutrophils of 61% or greater and eosinophils less than 3%), EA (neutrophils less than 61% and eosinophils of 3% or greater), PGA (neutrophils less than 61% and eosinophilsless than 3%), and MGA (neutrophils of 61% or greater and eosinophils of 3% or greater).9,23 Patients with NA were included in this study.

Data collection and assessments

At baseline, all participants underwent a multidimensional assessment,24including the collection of sociodemographic information; history of asthma, medication use, and previous exacerbations; comorbidities; spirometry data; anxiety/depressive symptoms; atopy; asthma control (as measured by the Asthma Control Test [ACT],25 Asthma Control Questionnaire,26 and Asthma Quality of Life Questionnaire27); and blood sample results, the details of which are described in our previously published study.17 Before the multidimensional assessment, all participants were required to withhold controller and rescue therapies (see Supplemental Methods in this article's Online Repository at www.jaci-inpractice.org for a description of the withholding time). Spirometry (MedGraphics CPES/D USB, St Paul, Minn) was performed according to American Thoracic Society and European Respiratory Society standards.28 Comorbid anxiety/depressive symptoms were defined as a score of at least 8 on the Anxiety or Depression subscale of the 14-item Hospital Anxiety and Depression Scale.29 Atopy was identified by at least one positive skin prick test for common allergen extracts. The definition of asthma–chronic obstructive pulmonary disease (COPD) overlap (ACO) and further information are provided in the Supplemental Methods.

Sputum induction and analysis

All enrolled patients underwent sputum induction based on standard methods, as described in our previous study30 (see the Supplemental Methods).

Management of asthma

All patients with asthma who were enrolled in parts II and IV of the study received antiasthma treatment, as determined by their treating physicians based on Global Initiative for Asthma (GINA) guidelines in a real-world setting.31 The inhaled corticosteroid (ICS) formoterol as a single-inhaler maintenance and reliever therapy has been recommended for asthma management in patients prescribed budesonide-formoterol or beclometasone-formoterol.32,33 Patients who experienced an exacerbation received standard therapy for acute asthma from their physicians based on GINA guidelines.31,34 Details of patient management are described in Supplemental Methods.

Variable selection and assessment for cluster analysis

We used a standardized assessment of asthma characteristics (see Table E1 in this article's Online Repository at www.jaci-inpractice.org). Initially, 27 clinical variables were selected for this study. Little's Missing Completely at Random test was used to investigate the mechanism of missing data, and multiple imputation methods were employed to impute missing variables (see Table E2 in this article's Online Repository at www.jaci-inpractice.org).16,35 Finally, we conducted exploratory factor analysis to select the final variables (see Supplemental Methods).

Cluster analysis

We performed hierarchical cluster analysis using Ward's method on squared Euclidian distances with the 11 variables selected using exploratory factor analysis (see Table E3 in this article's Online Repository at www.jaci-inpractice.org), and selected variables for the cluster model were standardized using z-scores.9,36,37 The optimal number of clusters was determined based on the majority rule of R-package NbClust (Charrad, 2014, Tunisia).37,38 The detailed rule is described in the Supplemental Methods. The

identified clusters are presented as a dendrogram to visualize the hierarchical structure of the data.

Validation using decision tree analysis

We conducted a decision tree analysis to predict the cluster assignment for each participant in parts I and III of the study to validate the accuracy of the clustering results9 (see the Supplemental Methods).

Future outcomes in prespecified subtypes of NA

Asthma outcomes included moderate and severe exacerbations in which patients had unscheduled visits, emergency room visits, hospitalizations, and intensive care unit admissions. Based on the American Thoracic Society/European Respiratory Society statements,39 a moderate asthma exacerbation is an event that, when recognized, should result in a temporary change in treatment to prevent the exacerbation from becoming severe, such as an unscheduled visit. Severe asthma exacerbations are events that require urgent action from the patient and physician to prevent serious outcomes, including emergency room visits, hospitalization or intensive care unit admission using systemic corticosteroids for at least 3 days. The occurrence and frequency of the asthma exacerbation events were recorded during the 12-month follow-up.9

Stability of NA phenotype

In part IV of the study, we first analyzed the variation ratio of the inflammatory phenotype among the clusters to explore the stability of the inflammatory phenotype assessed visually using a Sankey diagram.40 Changes in sputum cells in the identified clusters before and 1 month after stable antiasthma treatment were compared.

Statistical analysis

Continuous variables are expressed as means ± SDs or medians (interquartile ranges). Student t test, Mann-Whitney test, analysis of variance, or Kruskal-Wallis test was used for comparison. Categorical variables are expressed as frequencies and proportions. We used $\chi 2$ test or Fisher exact test for comparisons. Least significant difference was used for multiple comparisons as post hoc tests. Bonferroni correction was applied for multiple testing and a corrected significance level computed as 0.05 divided by 3 (that is, 0.017). Spearman's test was performed to evaluate correlations. We performed a paired t test to evaluate changes in sputum cells during the 1-month follow-up. Because the variation between ICS dose data was large, natural logarithm (lnICS) conversion was performed. Multiple imputation was performed using the mice package, the optimal number of clusters was determined based on NbClust, a cluster dendrogram was performed using packages factoextra and ggplot2, decision tree analysis was performed using rpart and rpart.plot packages, and a Sankey diagram was described based on ggalluvial package in R software (version 4.1.1). We performed statistical analysis using Statistical Package for the Social Sciences (version 22.0, IBM Corp, Chicago, Ill). Two-tailed P less than .05 was considered statistically significant.

Results

Variable selection

Eleven variables were selected for cluster analysis, including age, sex, body mass index, asthma duration, smoking status, ACT score, comorbid anxiety/depressive symptoms, prebronchodilator FEV1% predicted, peripheral blood eosinophil count, peripheral blood neutrophil count, and ICS dosage. The detailed steps and results of the variables selection are described in the Supplemental Results (in this article's Online Repository at www.jaci-inpractice.org).

Participants' characteristics

A total of 790 patients with the diagnosis of asthma were consecutively recruited, and 201 eligible patients with NA were included in this study. The 201 participants were randomly assigned to part I (n = 149) or III (n = 52) of the study (see Table E4 and the Supplemental Results in this article's Online Repository at www.jaci-inpractice.org for demographic and clinical characteristics of the included participants).

Cluster analysis

According to the Dindex criterion, three optimal clusters were identified (Figure 2, A). A hierarchical dendrogram of the three clusters is presented in Figure 2, B. Participants' demographics and clinical characteristics and the inflammatory cell profile of the sputum and blood of the three clusters are presented in Table I and Table E5 (in this article's Online Repository at www.jaci-inpractice.org). Cluster 1 was characterized by female-predominant nonsmokers with well-controlled asthma, cluster 2 by individuals with comorbid anxiety/depressive symptoms with poorly controlled asthma, and cluster 3 by older male smokers with late-onset asthma.

Figure 2. Three clusters identified in the study. (A) Number of clusters identified using

NbClust package. According to the majority rule, a significant knee in the plot of index values against the number of clusters and a significant peak in the plot of second differences values indicate that the optimal number of clusters is 3. (B) Hierarchical structure dendrogram of three clusters identified using Ward's hierarchical clustering. Blue indicates cluster 1; yellow, cluster 2; and green, cluster 3.

Table I. Demographic and clinical characteristics of three clusters in part I of study

Variable	Cluster 1	Cluster 2	Cluster 3	F/χ2/H	P
n	99	16	34		
Male/female	26/73	8/8	32/2‡,††	47.462	<.001
Age, y	48.89 ± 14.10	48.25 ± 12.24	60.71 ± 10.78||,¶	10.633	<.001
Body mass index, kg/m2	23.33 ± 3.30	21.02 ± 3.68§	22.86 ± 3.51	3.219	.043
Age at asthma onset, y	31.00 (12.00-47.00)	35.50 (20.50-48.50)	55.50 (46.00-64.25)‡,††	36.223	<.001
Asthma duration, y	11.00 (3.00-29.00)	10.00 (4.25-18.75)	4.00 (1.75-8.50)‡	12.575	.002
Smoking status, n (%)				107.514	<.001

Current smoker	1 (1)	3 (18.8)	18 (52.9)		
Ex-smoker	7 (7.1)	4 (25.0)	15 (44.1)		
Never	91 (91.9)	9 (56.3)	1 (2.9)		
Smoking pack-years	0.00	0.00 (0.00-16.96)	31.00 (10.83-42.750)‡,‡‡	95.086	<.001
Spirometry (prebronchodilator)					
FEV1, L	1.92 (1.39-2.63)	1.72 (1.28-2.27)	1.46 (1.17-1.91)†	9.064	.011
FEV1, % predicted	72.65 ± 20.98	59.50 ± 23.11§	51.72 ± 18.56‖	13.960	<.001
FVC, L	2.91 (2.29-3.49)	2.69 (2.36-3.56)	2.87 (2.39-3.18)	0.052	.974
FVC, % predicted	94.84 ± 13.43	87.56 ± 17.24	86.61 ± 13.79§	5.387	.006
FEV1/FVC	66.48 ± 12.22	58.49 ± 16.40§	52.50 ± 12.02‖	16.286	<.001
Change in FEV1 (%)	11.00 (5.20-19.40)	9.07 (–0.19 to 22.18)	12.90 (9.00-23.68)	3.854	.146
Asthma Control Test score	21.00 (17.00-24.00)	16.5 (14.25-19.50)†	18.50 (14.75-21.00)*	15.693	<.001
Atopy, n (%)	49 (49.5)	12 (75.0)	15 (44.1)	4.422	.111
Comorbidities, n (%)					
Comorbid anxiety/depressive symptoms	2 (2)	16 (100)‡	0‡‡	130.550	<.001
Rhinoconjunctivitis and other allergic diseases	65 (65.7)	13 (81.3)	13 (38.2)*,**	11.074	.003
Asthma–chronic obstructive pulmonary disease overlap	4 (4.0)	4 (25.0)*	23 (67.7)‡,††	57.144	<.001
Nasal polyps	7 (7.1)	1 (6.3)	1 (2.9)	0.701	.764
Gastroesophageal reflux disease	5 (5.1)	0	1 (2.9)	0.402	1.000
Bronchiectasis	5 (5.1)	0	3 (8.8)	1.300	.521
Diabetes	4 (4.0)	0	1 (2.9)	0.285	1.000
Osteoporosis	7 (7.1)	0	1 (2.9)	0.956	.622
Maintenance medications					
ICS dose (beclomethasone dipropionate equivalent), μg/d	555.30 ± 482.03	712.50 ± 556.03	735.29 ± 618.35	5.028	.081
ICS/long-acting β2-agonist, n (%)	57 (57.6)	7 (43.8)	18 (52.9)	1.142	.588
Theophylline, n (%)	13 (13.1)	3 (18.8)	5 (14.7)	0.654	.770
Leukotriene receptor antagonist, n (%)	27 (27.3)	5 (26.3)	7 (20.6)	0.875	.653

FeNO, ppb	26.50 (16.00-50.75)	16.50 (10.75-44.00)	23.00 (16.00-44.00)	2.458	.293
IgE, IU/mL	75.00 (27.70-232.82)	76.05 (24.59-426.59)	38.85 (21.40-121.19)	3.945	.139
Moderate to severe exacerbation in past year, n (%)	50 (50.5)	10 (62.5)	15 (44.1)	1.474	.487

ICS, inhaled corticosteroid; LSD, least significant difference.
Data are represented as means ± SDs; medians (quartiles 1-3), or frequency (%). Smoking pack-years = pack/d × smoking-years. ICS dose = equivalent dose of beclomethasone dipropionate/d for ICS/long-acting β2-agonist inhalers.
*
P < .017 vs cluster 1, with the Bonferroni correction.
†
P < .005 vs cluster 1, with the Bonferroni correction.
‡
P < .001 vs cluster 1, with the Bonferroni correction.
§
P < .05 vs. cluster 1, LSD was used as post hoc tests.
||
P < .001 vs cluster 1, LSD was used as post hoc tests.
¶
P < .05 vs cluster 2, LSD was used as post hoc tests.
**
P < .017 vs cluster 2, with the Bonferroni correction.
††
P < .005 vs cluster 2, with the Bonferroni correction.
‡‡
P < .001 vs cluster 2, with the Bonferroni correction.

Cluster 1: female-predominant nonsmokers with well-controlled asthma

Cluster 1 (n = 99) was the largest, comprising 66.4% of the total NA population, and had a female predominance (n = 73; 73.7%). Most participants were nonsmokers (n = 91; 91.9%). Participants in cluster 1 had a mean ACT score of 21.00 (17.00–24.00), indicating well-controlled asthma. Cluster 1 had a longer duration of asthma than did clusters 2 and 3 (11.00 [3.00-29.00] vs 10.00 [4.25-18.75] vs 4.00 [1.75-8.50] years, respectively; P = .002). Furthermore, cluster 1 exhibited less severe airway obstruction in prebronchodilator FEV1% predicted values (72.65% ± 20.98% vs 59.50% ± 23.11% vs 51.72% ± 18.5%; P < .001) and FEV1/FVC% values (66.48% ± 12.22% vs 58.49% ± 16.40% vs 52.50% ± 12.02%; P < .001) than did clusters 2 and 3, respectively.

Cluster 2: individuals with comorbid anxiety/depressive symptoms with poorly controlled asthma

Cluster 2 was the smallest (n = 16), comprising 10.7% of the total NA population. The proportion of patients with comorbid anxiety/depressive symptoms in cluster 2 was greater than that in clusters 1 and 3 (100.0% vs 2.0% vs 0.0%, respectively; P < .001). Moreover, cluster 2 had lower ACT scores than did clusters 1 and 3 (16.50 [14.25-19.50] vs 21.00 [17.00-24.00] vs 18.50 [14.75-21.00], respectively; P < .001). All participants were uncontrolled (n = 12; 75%) or partially controlled (n = 4; 25%) for asthma.

Furthermore, 81.3% (n = 13) presented with rhinoconjunctivitis and other allergic diseases.

Cluster 3: older male smokers with late-onset asthma

Cluster 3 (n = 34) was dominated by older adult males (n = 32, 94.12%, mean age 60.71 ± 10.78 years), with a later age at asthma onset (55.50 [46.00-64.25] vs 31.00 [12.00-47.00] vs 35.50 [20.50-48.50] years; P < .001) compared with clusters 1 and 2, respectively. A total of 97.1% of participants (n = 33) were smokers, with 31.00 median smoking pack-years (10.83-42.75). They had the worst airway obstruction compared with the other two clusters (Table I).

Future outcomes of three identified clusters

To validate the differential future risk of asthma within the prespecified clusters, we conducted a 12-month longitudinal cohort study. Of the participants, 89.2% (n = 133) completed the 12-month follow-up (part II of the study). The ICS adherence and medication use of the three clusters in part II are described in Table E6 and the Supplemental Results (in this article's Online Repository at www.jaci-inpractice.org). In cluster 2, compared with clusters 1 and 3, a greater proportion of patients experienced severe asthma exacerbations (35.7% vs 6.8% vs 19.4%, respectively; P = .005), more frequent hospitalizations (28.6% vs 4.5% vs 12.9%, respectively; P = .010), and more frequent unscheduled visits that were indicative of a higher risk of moderate exacerbations (50.0% vs 14.8% vs 22.6%, respectively; P = .013). Furthermore, patients in cluster 2 also had a higher number of emergency room visits (0.71 ± 2.40 vs 0.01 ± 0.11 vs 0.10 ± 0.40, respectively; P = .039) than those in clusters 1 and 3 (Table II). At the end of the 12-month follow-up, there were 11 (12.50%), 3 (21.43%), and 6 (19.35%) patients with severe asthma based on the 2020 GINA guideline in clusters 1, 2, and 3 (P = . 447), respectively.

Table II. Asthma outcomes within 12-month follow-up in neutrophilic asthma, grouped by cluster analysis in part II of study

Outcome	Cluster 1	Cluster 2	Cluster 3	F/χ2/H	P
n	88	14	31		
Severe exacerbation					
n (%)	6 (6.8)	5 (35.7)*	6 (19.4)	9.735	.005
Mean ± SD	0.07 ± 0.30	0.92 ± 2.57	0.17 ± 0.47	5.572	.062
Hospitalization					
n (%)	4 (4.5)	4 (28.6)*	4 (12.9)	8.181	.010
Mean ± SD	0.03 ± 0.18	0.17 ± 0.39	0.16 ± 0.45	5.127	.077
Intensive care unit admission, n (%)	0	1 (7.1)	0	4.963	.105
Emergency room visit					
n (%)	2 (2.3)	2 (14.3)	2 (6.5)	4.445	.071
Mean ± SD	0.01 ± 0.11	0.71 ± 2.40†	0.10 ± 0.40‡	6.473	.039
Unscheduled visit					
N (%)	13 (14.8)	7 (50.0)*	7 (22.6)	8.313	.013
Mean ± SD	0.20 ± 0.59	1.73 ± 3.55	0.38 ± 0.98	4.631	.099

*
P < .017 vs cluster 1, with the Bonferroni correction.
†
P < .005 vs cluster 1, LSD was used as post hoc tests.
‡
P < .05 vs cluster 2, LSD was used as post hoc tests.

Validation of identified clusters using decision tree analysis

We developed a tree diagram to predict the cluster assignments of the participants with NA (Figure 3). With only two variables such as smoking and comorbid anxiety/depressive symptoms, 92.6% of participants could be assigned to the correct cluster by decision tree analysis (Table III).

Figure 3. Decision tree analysis and correct classification rate for each cluster. Decision tree analysis using two variables of cigarette smoking and comorbid anxiety/depressive symptoms. Blue indicates cluster 1, female-predominant nonsmokers with well-controlled asthma. Red indicates cluster 2, individuals with comorbid anxiety/depressive symptoms with poorly controlled asthma. Green indicates cluster 3, older male smokers with late-onset asthma. The size of the circle represents the sample size.

Table III. Misclassification of predicted clinical cluster assignment by decision tree analysis

Clusters using 11 variables by Ward cluster analysis				Clusters using two variables by decision tree analysis
	Cluster 1, n	Cluster 2, n	Cluster 3, n	n (%)
Cluster 1, n	89	2	8	99 (66.4)
Cluster 2, n	0	16	0	16 (10.7)
Cluster 3, n	1	0	33	34 (22.9)
n (%)	90 (60.4)	18 (12.1)	41 (27.5)	149 (100)

Fifty-two patients with NA from an independent cohort (part III of the study) were used to validate the identified clusters using the two-variable tree diagram, and three clusters were identified (clusters I, II, and III). Although a small number of variables in clusters I, II, and III in part III of the study did not reach statistical significance, the differential features of clusters I, II, and III were similar to the features of clusters 1, 2, and 3 as shown in part I of the study (see Tables E7 and E8 in this article's Online Repository at www.jaci-inpractice.org).

Switching of inflammatory phenotype and NA stability

Of 149 patients, 91 (61.6%) from part I of the study who underwent successful sputum induction at both visits 1 and 2 were included in part IV. We excluded 58 patients because they underwent sputum induction only at visit 1. There were no significant differences in the demographic and clinical characteristics between the 91 included and 58 excluded patients (see Table E9 in this article's Online Repository at www.jaci-inpractice.org). The stability of the NA phenotype was significantly different among the three prespecified clusters (P = .010) after 1 month of stable treatment. At that time, 66.7% of patients in cluster 1 (n = 42) and 60.0% in cluster 3 (n = 12) shifted to other phenotypes, such as EA, PGA, or MGA. However, only 12.5% of patients in cluster 2 (n = 1) shifted to PGA, indicating that most patients (87.5%) had a persistent NA phenotype

between the two visits (Figure 4; also see Table E10 in this article's Online Repository at www.jaci-inpractice.org).

Figure 4. Sankey diagram of inflammatory phenotype variation among three clusters in part IV. After 1 month of stable treatment, 66.7% and 60% patients with neutrophilic asthma (NA) in clusters 1 and 3, respectively, shifted to other phenotypes, and only 12.5% of patients in cluster 2 shifted to other phenotypes. Pink indicates eosinophilic asthma (EA); green, mixed granulocytic asthma (MGA); blue, paucigranulocytic asthma (PGA); and gray, NA.

No differences were observed in the induced sputum total cell counts and the proportion of eosinophils, neutrophils, lymphocytes, and macrophages in the 91 patients among the three clusters at baseline (all $P > .05$). After 1 month of stable treatment, the induced sputum total cell counts ($P = .040$) and percentages of neutrophils ($P = .023$) were significantly different among the three clusters. The percentages of sputum neutrophils in the three identified clusters demonstrated different changes between baseline vs that at 1-month follow-up, with a significant decrease in clusters 1 (85.75% [72.50% to 96.00%] vs 47.20% [31.75% to 83.75%], respectively; $P < .001$) and 3 (79.50% [69.85% to 89.69%] vs 67.17% [38.69% to 96.38%], respectively; $P = .034$), but not in cluster 2 (78.13% [65.00% to 93.56%] vs 88.50% [76.00% to 97.38%], respectively; $P = .404$) (Figure 5). Furthermore, correlation analysis showed that the percentage of sputum neutrophils was correlated with age ($r = 0.231$; $P = .028$), smoking pack-years ($r = 0.216$; $P = .040$), and rhinoconjunctivitis and other allergic diseases ($r = –0.267$; $P = .011$). No significant correlations were found among the percentages of sputum neutrophils and sex, body mass index, ICS dose, comorbid anxiety/depressive symptoms, or atopy.

Figure 5. Changes in induced sputum cells in the three clusters after 1 month of stable treatment in part IV of the study. (A) Change in sputum neutrophils (%). (B) Change in sputum eosinophils (%). (C) Change in sputum lymphocytes (%). (D) Change in sputum macrophages (%). Visit 1 was the baseline visit. Visit 2 was at 1 month after baseline visit. Statistical significance was determined by paired t test.

The ICS doses in the three clusters are described in the Supplemental Results. To investigate the impact of ICS treatment on NA phenotype switching, we explored the correlation between ICS dose and non-neutrophilic inflammation at visit 2 within the three prespecified clusters. This analysis showed no statistically significant difference in ICS doses between patients with intermittent NA (neutrophilic inflammation presenting at visit 1 but not visit 2) and persistent NA (neutrophilic inflammation presenting at visits 1 and 2) in the three clusters (all $P > .05$) (see Table E11 in this article's Online Repository at www.jaci-inpractice.org).

Discussion

To the best of our knowledge, this is the first study to use cluster analysis to group patients with NA, validate the identified subtypes associated with differential outcomes, and further explore prespecified subtype stability. As a result, we identified three subtypes of NA in a real-world setting: female-predominant nonsmokers with well-controlled asthma, individuals with comorbid anxiety/depressive symptoms with a poorly controlled asthma, and older male smokers with late-onset asthma. Most of the burden of health careuse owing to an increased risk of severe exacerbation of asthma,

frequent hospitalizations, unscheduled visits, and emergency room visits was observed in cluster 2. Furthermore, the switch from one inflammatory phenotype to another implied that neutrophilic inflammation could convert to other cellular profiles of inflammation during the management of NA. However, cluster 2, characterized by individuals with comorbid anxiety/depressive symptoms with a poorly controlled level, appears to be a more stable subtype of NA. Cluster 3 included patients with features that resembled COPD, such as a smoking history, older age, and airflow limitation,41 and appeared to be an ACO subtype of NA.42 This study indicates that NA is heterogeneous and can be divided into different subtypes characterized by distinct features, and this has important implications in clinical practice. Neutrophilic asthma associated with comorbid anxiety/depressive symptoms and poor asthma control is a small but stable subtype with the greatest health care burden.

Asthma is a heterogeneous condition, and it has been well-classified in different ways in clinical practice.43,44 This study demonstrated different levels of asthma control across the three identified clusters, reflecting the heterogeneity of NA. Cluster 2 had poorer symptom control, with an ACT score of less than 20, and the worst asthma outcomes, indicating that cluster 2 may represent an uncontrolled NA subtype. Cluster 1 may represent a well-controlled NA subtype, with a mean ACT score of 21 and fewer asthma exacerbations. It has been shown that traits can be related to asthma control and exacerbation,45, 46, 47, 48, 49 and the heterogeneity of NA reflected in clusters with different levels of asthma control may be explained by several traits. First, adherence to treatment is essential in the management of asthma, and poor adherence to ICS leads to poorer asthma control and an increased number of asthma exacerbations.45,50,51 However, the overall ICS adherence in this study was good, with a mean medication possession ratio above 0.8, which did not differ across the three identified clusters. We believe that ICS adherence did not account for the heterogeneity of NA in the current study.

Second, mental health condition (anxiety/depression), as an important treatable trait, may affect asthma control and exacerbations.52,53 It has been demonstrated that anxiety/depressive symptoms correlate with non-T2 inflammation, leading to a poor response to antiasthma treatment.53, 54, 55 For example, our study54found that depressive symptom–associated IL-1β and TNF-α release correlated with impaired bronchodilator response and neutrophilic airway inflammation in asthma. Thus, cluster 2 (ie, individuals with comorbid anxiety/depressive symptoms) may represent a neuropsychological subtype in the heterogeneity of NA.54

Third, the different asthma control levels within the identified clusters may be explained by the traits of smoking and age.56, 57, 58 Smoking has been shown to be associated with worse clinical outcomes, a decline in lung function, and reduced sensitivity to corticosteroids in asthma,59,60 which could be explained by oxidative stress61,62 and a reduced glucocorticoid receptor α:β ratio63 in activated macrophages and neutrophils.64,65 Similar to smoking, age is positively associated with sputum neutrophils,58,66 an important trait in predicting worse asthma outcomes such as airway obstruction, a reduced response to antiasthma treatment, and increased exacerbations.58,67 Patients in cluster 3 were older, and almost all smokers with fixed airway obstruction required medium- to high-dose ICS to maintain asthma control, which may represent an ACO subtype in the heterogeneity of NA.

The clinical implications of this study support a treatable traits approach to asthma care. The study demonstrates that the clear heterogeneity in asthma is related to important clinical outcomes such as future exacerbation risk. Our study shows that heterogeneity in NA can be usably characterized by evaluating comorbidities (such as comorbid anxiety/depressive symptoms) and asthma risk factors (such as smoking), which are related to the future risk of asthma exacerbations. These variables or treatable traits provide a clinically useful method to assess the heterogeneity of the inflammatory subtypes of asthma, particularly NA.

Recent studies have shown variability in sputum inflammatory phenotypes during the course of asthma disease.68, 69, 70 In this study, we found that cluster 1 showed significant switching between inflammatory phenotypes; therefore, cluster 1 may be a transient NA subtype. Some issues associated with sputum neutrophils may explain the shifts in inflammatory phenotypes. First, respiratory pathogens (ie, bacteria and viruses) have an important effect on sputum neutrophilia in asthma,71, 72, 73 because infection resulting from pathogens is reported to be significantly associated with circulating blood leukocytes.74 Here, the percentages of neutrophils and lymphocytes in blood were both significantly different among the three clusters, which suggests that pathogen infection may differ among the three clusters, particularly for cluster 1, which had the highest percentage of blood neutrophils with transient sputum neutrophilia, possibly related to pathogen infection. Furthermore, because growing evidence suggests that the airway microbiome differs among different asthma phenotypes, we postulated that colonized pathogens in the airway may affect the phenotype. For example, NA was reported to be associated with lower richness and diversity of the airway bacterial community than was EA.75Furthermore, studies using a murine model showed that long-term exposure to low-dose Haemophilus influenzae can cause a switch in the inflammatory phenotype from EA to NA.76,77 It is possible that cluster 2, representing an NA phenotype that does not transition to other phenotypes with time, may have lower richness and diversity of the airway bacterial community compared with cluster 1, or it may have had long-term exposure to H influenzae. However, sputum pathogens were not detected in this study, which needs to be explored further.

Second, the effects of air pollution cannot be neglected. It was reported that airway neutrophilia was related to environmental exposures such as living in close proximity to a main road,78 and a neutrophilic endotype in asthma was more prevalent in countries with higher environmental pollutants.79,80 Traffic-related air pollution and diesel exhaust exposure were demonstrated to enhance ozone-induced airway neutrophilic inflammation and affect neutrophil function to favor more degranulation and local tissue damage.81,82 Participants in the current study were from different regions, urban and rural areas with different levels of air quality and pollution. Therefore, in the current study, it is possible that patients in cluster 2 with a stable NA phenotype were from areas with persistent air pollution. However, we did not investigate the effect of air pollution on the switching of inflammatory phenotypes in NA, which needs to be explored further.

Third, psychological stress has been strongly implicated in morbidity and mortality in asthma.83, 84, 85 Moreover, it is associated with increased neutrophilic and eosinophilic airway inflammatory response.86, 87, 88 Anxiety disorders or anxiety symptoms can be markers or a consequence of chronic stress.89 Therefore, the differences in stability of neutrophilic inflammation phenotype among the three

identified clusters with distinct anxiety/depressive symptoms may result from different levels of stress.
Fourth, ICS may have a role on sputum neutrophils in asthma because it promotes eosinophil apoptosis and prolongs the survival of functional neutrophils.90,91 For example, Brooks et al66 found that current use of ICS in adults with asthma was independently associated with increased sputum neutrophil levels. However, other studies showed that neutrophilic inflammation was observed in patients with asthma regardless of corticosteroid treatment.92,93 Moreover, the discontinuation or decrease of ICS could not significantly reduce the percentage of sputum neutrophils.94 We found no difference in corticosteroid exposure across the three identified clusters in the current study, and the ICS dose was not associated with percentages of sputum neutrophils. The effects of ICS treatment on airway neutrophils remain controversial, and further studies will be needed to confirm them.
Fifth, demographic characteristics such as age and obesity were related to airway neutrophils,14,66 but they did not significantly change during the 1-month follow-up, which could not explain the variability in this study.
This study had a few limitations. First, we could not avoid bias, although we used factor analysis and objective statistical methods such as the NbClust criterion to ensure as much objectivity as possible. Second, we investigated only a small Chinese population, which limited the generalizability of our findings; these need to be validated with future studies. Third, only two consecutive monthly sputum samples were measured to analyze the stability of the inflammatory phenotype, which may have underestimated the rate of inflammatory phenotype variability. Further studies are needed to explore the long-term stability of NA. Fourth, this study only suggested correlations between anxiety/depressive symptoms and asthma outcomes, and no causal analysis of asthma outcomes was performed. Fifth, there may have been a subjective bias in sputum differential cell counting. However, we performed sputum induction and processing in the study using the standard operating procedure17,30 and two well-trained laboratory researchers from Australia and China determined differential cell counts, with high agreement.54

Conclusion

To our knowledge, this is the first study to identify NA heterogeneity using cluster analysis in a real-world setting. Our findings demonstrated that NA can be divided into three clusters based on distinct physiologic and clinical features. Cluster 1 was characterized by a mild and transient NA subtype with significant shifts in granulocyte patterns and cluster 2 by an uncontrolled and stable NA subtype characterized by comorbid anxiety/depressive symptoms with worse asthma outcomes. Almost all patients in cluster 3 were smokers with fixed airway obstruction that may represent as an asthma–COPD overlap subtype. These findings have important implications for the individualized assessment and management of patients with NA in clinical practice.

Acknowledgments

We wish to thank Ms Zhi Lin and Ms Michelle Gleeson for their work in sputum processing. G. Wang, W.M. Li, P.G. Gibson, K. Deng, and L.X. He designed the framework for the study. J. Wang, X. Zhang, L. Wang, K. Deng, and L.X. He assisted with the acquisition and analysis of data. G. Wang, W.M. Li, L. Wang, H.P. Zhang, L. Zhang, H.J. Wan, D. Liu, F.M. Luo, K. Deng, and L.X. He contributed to the interpretation of

results. G. Wang, W.M. Li, K. Deng, and L. X. He contributed to the study design, data analysis, and interpretation of results. G. Wang, W.M. Li, M. Xie, Z.H. Chen, J. Zhang, A.C.-Y. Hsu, B.G. Oliver, P.A.B. Wark, L. Qin, P. Gao, K. Deng, and L.X. He contributed to the interpretation of results and drafted the manuscript.

Online Repository

Table E1. Missing rates of variables in dataset

Variables	Missing (%)
Inhaled corticosteroid dose	6.97
Atopy	6.47
Postbronchodilator MMEF	4.98
Change in FEV1 (%)	3.48
Prebronchodilator FVC (% predicted)	2.99
Prebronchodilator FEV1	2.99
Prebronchodilator FVC	2.99
Prebronchodilator MMEF	2.49
Age at asthma onset	1.00
Asthma duration	1.00
Smoking pack-years	1.00

MMEF, maximum midexpiratory flow.

Table E2. Univariate comparisons of missing data grouped by asthma exacerbation

Variable	AE (n = 75)	Non-AE (n = 74)	χ2	P
Inhaled corticosteroid dose (beclomethasone dipropionate equivalent), μg/d	4 (5.33)	5 (6.76)	0.133	.983
Age at asthma onset	0	2 (2.70)	2.055	.245
Asthma duration	0	2 (2.70)	2.055	.245
Change in FEV1 (%)	3 (4.00)	1 (1.35)	1.000	.622
Smoking pack-years	0	1 (1.35)	1.020	.497
Atopy	5 (6.67)	5 (6.76)	<0.001	1.00
Postbronchodilator MMEF	6 (8.0)	1 (1.35)	3.678	.116
Prebronchodilator FVC (% predicted)	3 (4.00)	1 (1.35)	1.000	.620
Prebronchodilator FEV1	3 (4.00)	1 (1.35)	1.000	.620
Prebronchodilator FVC	3 (4.00)	1 (1.35)	1.000	.620
Prebronchodilator MMEF	2 (2.67)	0	2.000	.497

AE, asthma exacerbation; MMEF, maximum midexpiratory flow.

Table E3. Factor analysis using orthogonal varimax rotation of 27 variables

Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
Prebronchodilator FEV1	0.976	0.004	0.172	–0.017	0.023	–0.001	–0.013	0.052
Postbronchodilator FEV1	0.965	–0.007	0.133	0.038	–0.046	–0.079	0.058	–0.054
Prebronchodilator MMEF	0.891	0.211	0.135	–0.006	0.045	0.056	–0.144	0.098
Postbronchodilator MMEF	0.875	0.24	0.092	0.066	0.014	–0.046	–0.131	0.01
Prebronchodilator FVC	0.863	–0.37	0.11	–0.011	–0.081	–0.035	0.069	–0.094
Postbronchodilator FVC	0.809	–0.451	0.064	0.023	–0.144	–0.04	0.153	–0.214
Prebronchodilator FEV1(% predicted)*	0.757	0.392	0.181	–0.218	0.235	–0.064	0.144	0.176
Prebronchodilator FEV1/FVC (%)	0.648	0.53	0.187	–0.079	0.184	0.033	–0.109	0.221

Prebronchodilator FVC (% predicted) 0.596 0.171 0.066 –0.303 0.066 –0.163 0.493 –0.168
Age*
–0.592 –0.24 0.056 –0.156 0.395 –0.115 0.395 0.167
Sex*
–0.157 0.897 0.054 –0.054 –0.034 0.012 –0.043 0.097
Smoking status*
0.138 0.828 0.062 0.081 –0.17 –0.008 0.081 0.038
Smoking pack-years –0.19 –0.763 –0.053 –0.067 0.203 –0.043 –0.08 0.146
Asthma Control Questionnaire score –0.247 –0.12 –0.877 0.083 0.022 –0.096 –0.064 –0.029
Asthma Control Test score*
0.222 0.111 0.863 –0.053 –0.093 0.065 0.115 –0.016
Asthma Quality of Life Questionnaire score 0.09 –0.025 0.854 –0.179 0.057 0.012 –0.005 –0.039
Peripheral blood eosinophil count*
0.003 0.008 –0.144 0.935 –0.002 –0.04 0.067 –0.029
Peripheral blood eosinophils (%) 0.03 0.072 –0.147 0.894 0.039 –0.275 0.044 –0.026
Age at asthma onset –0.202 –0.236 <0.001 –0.045 0.900 –0.11 0.158 0.013
Asthma duration*
–0.273 0.089 0.049 –0.084 –0.811 0.039 0.153 0.134
Peripheral blood neutrophil count*
–0.085 –0.057 0.055 –0.033 –0.12 0.904 0.117 0.043
Peripheral blood neutrophil count (%) –0.025 0.113 0.093 –0.246 –0.01 0.844 0.112 0.031
Body mass index*
0.037 0.031 –0.004 0.012 –0.044 0.154 0.712 0.04
Comorbid anxiety/depressive symptoms*
–0.081 0.028 0.219 0.235 0.056 0.099 0.524 0.177
lnICS*
–0.046 –0.07 –0.003 0.087 0.289 0.034 –0.132 –0.693
Change in FEV1 (%) –0.305 –0.002 –0.168 0.319 –0.169 –0.243 0.277 –0.477
Atopy –0.227 –0.029 –0.159 0.226 0.213 0.091 0.192 0.471

MMEF, maximum midexpiratory flow.

*
Final variables used in cluster analysis.

Table E4. Demographic and clinical characteristics of subjects with neutrophilic asthma included in parts I and III

Variable	Part I	Part III	P
n	149	52	
Male/female	66/83	19/33	.415
Age, y	51.52 ± 14.07	50.71 ± 14.46	.725
Body mass index, kg/m2	22.98 ± 3.44	23.54 ± 3.54	.316
Age at asthma onset, y	38.00 (24.50-54.00)	35.50 (21.00-46.00)	.331
Asthma duration, y	7.00 (3.00-22.00)	9.50 (4.00-25.75)	.300
Smoking status, n (%)			.098

Current smoker	22 (14.8)	2 (3.8)		
Ex-smoker	26 (17.4)	12 (23.1)		
Never	101 (67.8)	38 (73.1)		
Smoking pack-years	0.00 (0.00-9.88)	0.00 (0.00-1.70)	.422	
Spirometry (prebronchodilator)				
FEV1,L	1.71 (1.28-2.35)	1.90 (1.44-2.32)	.351	
FEV1 (% predicted)	66.46 ± 22.43	69.76 ± 20.06	.349	
FVC, L	2.87 (2.33-3.48)	2.97 (2.51-3.64)	.259	
FVC (% predicted)	92.18 ± 14.36	94.60 ± 17.63	.326	
FEV1/FVC	62.43 ± 13.92	62.56 ± 11.77	.955	
Change in FEV1 (%)	11.56 (5.74-19.9)	12.30 (4.66-17.20)	.514	
Asthma Control Test score	20.00 (16.00-22.00)	20.50 (16.25-23.00)	.486	
Atopy, n (%)	76 (51.0)	29 (55.8)	.629	
Comorbidities, n (%)				
Comorbid anxiety/depressive symptoms	18 (12.1)	3 (5.8)	.293	
Rhinoconjunctivitis and other allergic diseases	91 (61.6)	35 (67.3)	.506	
Asthma–chronic obstructive pulmonary disease overlap	31 (20.8)	7 (13.5)	.306	

Data are represented as means ± SDs, medians (quartiles 1-3), or frequencies (%).
Smoking pack-years = pack/d × smoking-years.

Table E5. Differential inflammatory cells in sputum and blood in patients of three clusters in part I

Variable	Cluster1	Cluster2	Cluster3	F/χ2/H	P
n	99	16	34		
Sputum					
Total cell counts, ×106/mL	5.54 (3.00-9.94)	4.14 (2.31-6.83)	3.83 (2.45-9.68)	2.968	.227
Eosinophils (%)	0.00 (0.00-0.50)	0.00 (0.00-0.44)	0.25 (0.00-1.25)	3.926	.140
Neutrophils (%)	83.25 (70.50-95.00)	79.50 (67.44-91.13)	78.69 (72.94-91.63)	0.863	.649
Lymphocytes (%)	0.25 (0.00-0.75)	0.50 (0.25-1.00)	0.50 (0.25-0.7)	2.750	.253
Macrophages (%)	13.25 (4.00-27.00)	18.88 (8.00-31.75)	19.38 (7.44-26.00)	1.246	.536
Blood					
Leukocytes, ×109/L	5.75 (35.04-7.18)	5.62 (4.86-6.72)	5.75 (5.01-6.83)	0.815	.665
Neutrophils, ×109/L	3.52 (3.05-4.73)	3.38 (2.68-4.08)	3.43 (2.87-4.03)	2.289	.318
Neutrophils (%)	62.22 (57.33-66.80)	58.27 (52.95-61.94)	61.14 (52.26-65.44)	6.668	.036
Eosinophils, ×109/L	0.19 (0.10-0.31)	0.14 (0.07-0.27)	0.15 (0.10-0.23)	2.576	.276
Eosinophils (%)	3.16 (1.63-5.20)	2.21 (1.16-5.72)	2.40 (1.74-4.55)	1.548	.461

Monocytes, ×109/L	0.32 (0.26-0.41)	0.32 (0.23-0.47)	0.42 (0.32-0.55)	8.299	.016
Monocytes (%)	5.47 (4.67-6.75)	5.93 (5.03-6.84)	7.03 (5.88-8.48)*	15.455	<.001
Lymphocytes, ×109/L	1.66 (1.40-1.93)	1.82 (1.46-2.21)	1.64 (1.41-2.32)	2.074	.354
Lymphocytes (%)	27.27 (23.14-32.92)	33.16 (26.94-37.47)	29.78 (25.16-35.23)	6.993	.030
Basophils, ×109/L	0.03 (0.02-0.05)	0.02 (0.01-0.04)	0.03 (0.02-0.04)	4.448	.108
Basophils (%)	0.53 (0.35-0.80)	0.35 (0.23-0.69)	0.54 (0.37-0.80)	2.938	.230

*
P < .001 vs cluster 1, with the Bonferroni correction.

Table E6. Medication use in three identified clusters after 1 year

Maintenance medications	Cluster 1	Cluster 2	Cluster 3	F/χ2/H	P
n	88	14	31		
Inhaled corticosteroid dose (beclomethasone dipropionate equivalent), μg/d	439.47 ± 498.94	600.00 ± 692.82	690.00 ± 619.76	4.992	.082
Inhaled corticosteroid/long-acting β2-agonist, n (%)	38 (43.2)	6 (42.3)	20 (64.5)	4.353	.119
Theophylline, n (%)	5 (5.7)	2 (14.3)	2 (6.5)	1.787	.407
Leukotriene receptor antagonist, n (%)	16 (18.2)	3 (21.4)	7 (22.6)	0.518	.793

Table E7. Demographic and clinical characteristics of three clusters in part III

Variable	Cluster I	Cluster II	Cluster III	F/χ2/H	P
n	34	5	13		
Male/female	5/29	2/3	13/0	30.714	<.001
Age, y	47.824 ± 13.325	41.800 ± 9.654	57.000 ± 14.463	3.148	.052
BMI, kg/m2	23.209 ± 2.927	21.244 ± 4.980	24.515 ± 4.654	1.546	.223
Age at asthma onset, y	31.00 (24.00-43.00)	18.00 (11.00-37.00)	40.00 (15.00-53.50)	2.170	.338
Asthma duration, y	7.50 (4.00-25.50)	16.00 (13.00-26.50)	10.00 (2.00-33.00)	1.630	.443
Smoking status, n (%)				49.370	<.001
Current smoker	0.00	0	2 (15.4)		
Ex-smoker	0.00	2 (40.00)	11 (84.6)		
Never	34 (100.00)	3 (60.00)	0		
Smoking pack-years	0.00	0.00 (0.00-17.75)	25.00 (3.80-39.25)	44.232	<.001
Spirometry (prebronchodilator)					
FEV1, L	2.00 (1.58-2.35)	1.83 (1.56-2.14)	2.01 (1.44-2.78)	0.330	.848
FEV1 (% predicted)	73.104 ± 21.715	62.316 ± 12.398	64.308 ± 16.685	1.310	.279
FVC, L	2.74 (2.47-3.33)	2.98 (2.38-3.54)	3.64 (2.82-4.78)	5.189	.075

FVC (% predicted)	94.930 ± 18.937	86.356 ± 4.478	93.769 ± 16.639	0.515	.601
FEV1/FVC	65.497 ± 11.508	63.684 ± 12.460	56.740 ± 9.864	2.875	.066
Change in FEV1 (%)	12.51 (5.54-17.18)	10.00 (5.57-16.73)	9.45 (5.40-18.70)	0.030	.958
Asthma Control Test score	21.00 (18.00-23.00)	16.00 (13.50-18.50)	21.00 (16.5-22.5)	5.931	.052
Atopy, n (%)	19 (55.9)	3 (60.00)	9 (69.2)	0.767	.824
Comorbidities, n (%)					
Comorbid anxiety/depressive symptoms	0	5 (100.00)	0	27.002	<.001
Rhinoconjunctivitis and other allergic diseases	23 (67.6)	5 (100.00)	7 (53.8)	3.203	.202
Asthma–chronic obstructive pulmonary disease overlap	0	1 (20.00)	6 (46.15)*	16.184	<0.001
ICS dose (beclomethasone dipropionate equivalent), μg/d	400.00 (400.00-500.00)	800.00 (400.0-1200.00)	400.00 (400.00-700.00)	2.253	.324
IgE, IU/mL	97.35 (32.08-97.35)	142.73 (5.61-553.04)	134.00 (46.83-568.19)	0.600	.741
FeNO, ppb	28.50 (17.75-43.00)	21.00 (10.00-77.00)	26.00 (17.00-119.00)	0.776	.678
Moderate to severe exacerbation in past y, n (%)	20 (58.8)	5 (100.00)	8 (61.5)	3.003	.239

ICS, inhaled corticosteroid.

Smoking pack-years = packs/d × smoking-years. ICS dose = equivalent dose of beclomethasone dipropionate/d for ICS/long-acting β2-agonist inhalers.

*

P < .001 vs cluster 1, with the Bonferroni correction.

Table E8. Differential inflammatory cells in sputum and blood in subjects of three clusters in part III

Variable	Cluster I	Cluster II	Cluster III	F/χ2/H	P
n	34	5	13		
Sputum					
Total cell counts, ×106/mL	5.67 (2.43-10.08)	4.14 (1.35-10.66)	4.37 (1.69-5.85)	1.289	.525
Eosinophils (%)	0.00 (0.00-0.25)	0.00 (0.00-0.50)	0.00 (0.00-0.63)	0.625	.732
Neutrophils (%)	81.50 (50.25-96.4)	77.00 (66.24-84.38)	76.50 (34.50-90.00)	0.592	.744
Lymphocytes (%)	0.25 (0.00-0.75)	0.00 (0.00-1.38)	0.50 (0.25-0.50)	0.725	.696
Macrophages (%)	11.25 (1.50-46.75)	21.25 (15.13-33.27)	23.00 (8.00-65.00)	0.963	.618
Blood					
Leukocytes, ×109/L	6.15 (4.76-7.53)	4.72 (4.19-5.43)	5.42 (4.57-8.31)	3.078	.215

Neutrophils, ×109/L	3.61 (2.60-4.88)	2.69 (1.89-3.35)	2.92 (2.85-6.22)	3.462 .177
Neutrophils (%)	63.16 (54.54-67.67)	56.28 (45.96-61.73)	60.94 (53.81-73.99)	1.297 .523
Eosinophils, ×109/L	0.16 (0.10-0.31)	0.23 (0.12-0.26)	0.16 (0.13-0.52)	0.689 .708
Eosinophils (%)	2.83 (1.75-4.50)	4.87 (2.39-5.62)	3.57 (1.76-6.47)	1.246 .536
Monocytes, ×109/L	0.34 (0.27-0.47)	0.26 (0.19-0.34)	0.43 (0.35-0.54)	7.199 .027
Monocytes (%)	5.69 (4.99-7.00)	5.51 (4.47-6.43)	7.37 (5.23-9.35)	5.836 .054
Lymphocytes, ×109/L	1.58 (1.35-2.08)	1.62 (1.39-2.00)	1.36 (1.16-1.70)	2.554 .279
Lymphocytes (%)	28.21 (22.18-34.08)	33.89 (28.40-43.87)	25.89 (17.69-30.34)	4.207 .122
Basophils, ×109/L	0.03 (0.02-0.05)	0.01 (0.01-0.03)	0.04 (0.03-0.06)	7.607 .022
Basophils (%)	0.59 (0.39-0.80)	0.22 (0.19-0.58)	0.67 (0.40-0.91)	5.076 .079

Table E9. Demographic and clinical characteristics of subjects with successful sputum induction at visits 1 and 2

Variable	Sputum induction at visits 1 and 2	Sputum induction only at visit 1	P
n	91	58	
Male/female	39/52	27/31	.736
Age, y	51.85 ± 14.33	51.00 ± 13.77	.719
Body mass index, kg/m2	22.91 ± 3.09	23.08 ± 3.96	.762
Age at asthma onset, y	38.00 (26.00-54.00)	35.50 (16.00-50.00)	.370
Asthma duration, y	7.00 (3.00-20.00)	9.00 (2.75-24.25)	.418
Smoking status, n (%)			.671
Current smoker	13 (14.3)	9 (15.5)	
Ex-smoker	14 (15.4)	12 (20.7)	
Never	64 (70.3)	37 (63.8)	
Smoking pack-years	0.00 (0.00-14.00)	0.00 (0.00-8.03)	.572
Spirometry (prebronchodilator)			
FEV1, L	1.70 (1.28-2.30)	1.77 (1.31-2.42)	.844
FEV1 (% predicted)	66.82 ± 22.55	65.89 ± 22.43	.805
FVC, L	2.87 (2.26-3.48)	2.88 (2.41-3.62)	.432
FVC (% predicted)	92.40 ± 13.94	91.85 ± 15.11	.822
FEV1/FVC	63.15 ± 13.90	61.32 ± 13.98	.436
Change in FEV1 (%)	11.00 (6.00-19.60)	12.29 (5.60-21.13)	.557
Asthma Control Test score	21.00 (17.00-23.00)	18.50 (16.00-22.00)	.113
Atopy, n (%)	48 (52.7)	29 (48.3)	.618
Comorbidities, n (%)			
Comorbid anxiety/depressive symptoms	10 (11.0)	8 (13.8)	.797
Rhinoconjunctivitis and other allergic diseases	61 (67.0)	30 (51.7)	.085

Data are represented as means ± SDs; medians (quartiles 1-3), or frequencies (%). Smoking pack-years = packs/d × smoking-years.

Table E10. Variations of inflammatory phenotype among three clusters within 1-month follow-up in part IV

Cluster	Neutrophilic asthma	Eosinophilic asthma	Paucigranulocytic asthma	Mixed granulocytic asthma	Total
1 (n [%])	21 (33.3)	8 (12.7)	29 (46.0)	5 (7.9)	63
2 (n [%])	7 (87.5)	0	1 (12.5)	0	8
3 (n [%])	8 (40.0)	1 (5.0)	10 (50.0)	1 (5.0)	20

Table E11. Inhaled corticosteroid doses among persistent and transient NA patients of three identified clusters in part IV

Cluster	Persistent NA	Intermittent NA	P
1	428.57 ± 247.27	571.42 ± 500.87	.223
2	771.42 ± 593.62	400.00.580	
3	560.00 ± 505.96	620.00 ± 520.26	.797
Total	526.32 ± 411.78	577.36 ± 495.61	.605

NA, neutrophilic asthma.

Persistent NA was defined as NA inflammation presenting both at visits 1 and 2. Intermittent NA was defined as NA inflammation presenting at visit 1 but not visit 2. Inhaled corticosteroid dose = equivalent dose of beclomethasone dipropionate/d for inhaled corticosteroid/long-acting β2-agonist inhalers (μg/d).

Supplemental Methods and Results

Methods

Data collection and assessments

Physiologic testing of lung function included pre- and postbronchodilator spirometry and was performed according to American Thoracic Society and European Respiratory Society standardsE1 using spirometry (MedGraphics CPES/D USB, St Paul, Minn). Comorbid anxiety/depressive symptoms were defined as having a score of 8 or more on the anxiety subscale or depression subscale using the 14-item Hospital Anxiety and Depression Scale.E2 Atopy was identified by at least one positive skin prick test to allergen extracts including house dust mites (Dermatophagoides pteronyssinus or Dermatophagoides farinae), mold (Alternaria tenuis or Aspergillus species), dog hair, cat hair, pollen (ragweed, birch, or London plane), and cockroach, in addition to positive (histamine) and negative (saline) controls.E3 Asthma–chronic obstructive pulmonary disease overlap was diagnosed as age 35 years or greater, a 10-pack-year or greater smoking history, and postbronchodilator FEV1/FVC value less than 0.70.E4Venous blood samples were collected in untreated tubes to obtain serum for the measurement of IgE levels by immunoassay (Beckman Image 800 immunoassay analyser, Beckman Coulter Inc, Brea, Calif), with a minimum detectable level of 5.0 IU/mL, or by ethylenediamine tetraacetic acid–treated tubes for total and differential blood cell counts (Sysmex XN-9000 83 hematology analyzer, Sysmex Corporation, Kobe, Japan).

Sputum induction and analysis

Before the pulmonary function test and sputum induction, short-acting inhaled drugs (eg, anticholinergic agent ipratropium bromide or β2-agonist albuterol/salbutamol) were withheld for at least 4 hours, long-acting β2-agonist bronchodilators for 24 to 36 hours, long-acting muscarinic antagonist for 36 to 48 hours, theophylline for 24 to 48

hours, inhaled corticosteroid for 72 h, and leukotriene receptor antagonist for 96 hours.E5, E6, E7 All enrolled patients underwent sputum induction based on standard methods.E8 Briefly, patients were pretreated with 400 μg salbutamol (GSK, Avda de Extremadura, Spain), and 15 minutes later, sputum was induced using 4.5% saline atomized with an ultrasonic nebulizer (Cumulus, Heyer Medical AG, Bad Ems, Germany). If pre- or post-FEV1 was less than 40% predicted, sputum was induced with 0.9% saline for safety. Then sputum plugs were selected and dispersed using dithiothreitol and the total cell count and viability were analyzed. Cytospins were prepared using centrifugation-smear (CytoPro 7620, Wescor, Inc, Logan, Utah) and stained (May-Grunwald Giemsa). Then, differential cell counts were performed by well-trained researchers from Australia and China.

Management of asthma

All subjects with asthma who were enrolled in parts II and IV received antiasthma treatment, determined by their treating physicians based on the Global Initiative for Asthma guidelines in a real-world setting.E9 Enrolled patients received inhaled corticosteroid (ICS)/long-acting β2-agonist, leukotriene receptor antagonist, or theophylline based on the stepwise of Global Initiative for Asthma treatment recommendations. Patients in part IV were required to maintain their controller medications or not change them. Patients were excluded from the analysis if they changed controller medications or had asthma exacerbation or a respiratory tract infection during the 1-month follow-up.

Variable selection and assessment for cluster analysis

A total of 286 variables were collected, 38 with 10% or more missing values were excluded.E10 Furthermore, 131 redundant and 90 irrelevant variables were excluded based on demographic, clinical relevance and the purpose of this study. Finally, after initial data selection, 27 variables were selected. There were 11 variables with missing data, with missing rates ranging from 0.5% to 6.97% (Table E1). Missing data are often categorized into three types: missing completely at random (MCAR), missing at random, and missing not at random. To diagnose the mechanism of missing data, we first used Little's MCAR test to check the missing values in the dataset.E11 The result showed that missing data in the study were not MCAR (P = .206). Second, to explore the mechanism for the missing variables, subjects were divided into two groups based on the incidence of acute exacerbations or not, according to the main outcome of this study, and we performed a $\chi 2$ test between the groups.E12,E13 No significant difference was found between the groups (Table E2). Therefore, the missing data of the variables may have resulted from missing at random. Then, we conducted the multiple imputation method (multilevel and generalized linear regression) to impute variables by the mice package in R software (version 4.1.1; RStudio, Boston, MA), which was described in detail in our previous study.E13 First, the 27 distinct variables were evaluated by factor analysis with orthogonal varimax rotation.E14 The appropriate number of factors was selected by analysis of the Scree plot with the requirements that retained factors explain more than 70% of data variance and each factor have an eigenvalue greater than 1.0. Variables with a loading greater than 0.5 were retained. Second, a correlation matrix of the variables was created using the correlation coefficient. Variables with a correlation coefficient greater than 0.6 were initially selected,E15,E16 and the final selected variables were based on a clinical perspective.E16

Cluster analysis

In the NbClust package, the optimal number of clusters is identified by a significant knee in the plot of index values against the number of clusters. A significant peak in the plot of second differences values indicates the relevant number of clusters.E17

Validation by decision tree analysis

We conducted decision tree analysis to predict the cluster assignment for each subject in part I using all 11 variables, and then calculated the misclassification rates. Decision tree analysis was performed using the algorithm based on classification and regression trees based on a cost-complexity pruning criterion.E18 Then, we used the main variables from the decision tree analysis to group subjects from another separate population enrolled in part III to validate the accuracy of the clustering results. Finally, demographic and clinical data of clusters were compared between parts I and III.

Results

Variable selection

The initial 27 selected variables were evaluated by exploratory factor analysis with orthogonal varimax rotation. First, the Kaiser-Meyer-Olkin measure verified that the initial 27 selected variables were adequate for exploratory factor analysis (Kaiser-Meyer-Olkin measure = 0.713; Bartlett sphericity test P ≤ .001). Second, based on the pattern of loading, eight factors were identified with an eigenvalue greater than 1, which accounted for 79.66% of variance in the factor analysis. Third, we selected 10 parameters representative of each factor based on a loading value of greater than 0.5 and correlation coefficient greater than 0.6. Because sex was an important clinical characteristic of asthma, we selected it as a cluster parameter. Finally, 11 variables were selected for cluster analysis, including age, sex, body mass index, asthma duration, smoking status, Asthma Control Test score, comorbid anxiety/depressive symptoms, prebronchodilator FEV1%, peripheral blood eosinophil count, peripheral blood neutrophil count, and lnICS (Table E3).

Participants' characteristics

In part I, 83 patients were female (55.7%), mean age 51.52 ± 14.07 years, with a body mass index of 22.98 ± 3.44 kg/m2. Among them, 32.2% were smokers (n = 48), 51.0% had atopy (n = 76), 12.1% had comorbid anxiety/depressive symptoms (n = 18), and 61.6% had rhinoconjunctivitis and other allergic diseases (n = 91). Mean age at asthma onset was 38 (24.50-54.00) years, and median duration of asthma was 7.00 years (interquartile range, 3.00-22.00 years). Among them, more than half of patients (55.1%) were controlled or partly controlled, with a median FEV1% predicted value of 66.46 ± 22.43. Median Asthma Control Test score was 20.00 (16.00-22.00).

Future outcomes of three identified clusters

The medication possession ratio for ICS adherence was 0.871 ± 0.255 during the 12-month follow-up. There was no significant difference in ICS adherence among clusters 1, 2, and 3 (0.917 ± 0.193 vs 0.857 ± 0.316 vs 0.782 ± 0.330, respectively, for the medication possession ratio; P = .265). The ICS dose of beclomethasone dipropionate equivalent was significantly reduced in cluster 1 (555.30 ± 482.03 vs 439.47 ± 498.94 μg/d; P = .002) in part II after the 1-year follow-up, but not in clusters 2 (712.50 ± 556.03 vs 600.00 ± 692.82 μg/d; P = .261) and 3 (735.29 ± 618.35 vs 690.00 ± 619.76 μg/d; P = .513).

Switching of inflammatory phenotype and neutrophilic asthma stability

Among the 91 patients who were included, the proportion of successful sputum induction was not differentiated among clusters 1 (n = 63; 63.4%), 2 (n = 8; 50.0%) and 3 (n = 20; 58.8%) in part IV (P = .575). There was no difference in controller medications such as ICS dose among clusters 1, 2, and 3 during the 1-month follow-up (477.05 ± 234.09 vs 725.00 ± 565.05 vs 590.00 ± 500.42 μg/d of beclomethasone dipropionate equivalent; P = .257). Further analysis indicated that the ICS dose was not associated with switching changes in airway inflammatory phenotypes or percentages of sputum neutrophils and eosinophils (all P > .05).