

# Computational Phenotyping of Obstructive Airway Diseases: A Systematic Review

Muwada Bashir Awad Bashir<sup>1</sup>, Gregorio Paolo Milani<sup>2,3,\*</sup>, Valentina De Cosmi<sup>4,5,\*</sup>,  
Alessandra Mazzocchi<sup>3,\*</sup>, Guoqiang Zhang<sup>1</sup>, Rani Basna<sup>1</sup>, Linnea Hedman<sup>6</sup>, Anne Lindberg<sup>7</sup>,  
Linda Ekerljung<sup>1</sup>, Malin Axelsson<sup>8</sup>, Lowie EGW Vanfleteren<sup>9</sup>, Eva Rönmark<sup>6</sup>, Helena Backman<sup>6</sup>,  
Hannu Kankaanranta<sup>1,10,11</sup>, Bright I Nwaru<sup>1,12</sup>

<sup>1</sup>Krefting Research Centre, Department of Internal Medicine and Clinical Nutrition, Institute of Medicine, University of Gothenburg, Gothenburg, Sweden; <sup>2</sup>Pediatric Unit, Foundation IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy; <sup>3</sup>Department of Clinical Science and Community Health, University of Milan, Milan, Italy; <sup>4</sup>Department of Food Safety, Nutrition and Veterinary Public Health, Istituto Superiore Di Sanità - Italian National Institute of Health, Roma, Italy; <sup>5</sup>Department of Clinical Sciences and Community Health, University of Milan, Milano, Italy; <sup>6</sup>Department of Public Health and Clinical Medicine, Section of Sustainable Health/ the OLIN Unit, Umeå University, Umeå, Sweden; <sup>7</sup>Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden; <sup>8</sup>Department of Care Science, Faculty of Health and Society, Malmö University, Malmö, Sweden; <sup>9</sup>COPD Center, Sahlgrenska University Hospital, University of Gothenburg, Gothenburg, Sweden; <sup>10</sup>Department of Respiratory Medicine, Seinäjoki Central Hospital, Seinäjoki, Finland; <sup>11</sup>Tampere University Respiratory Research Group, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland; <sup>12</sup>Wallenberg Centre for Molecular and Translational Medicine, Institute of Medicine, University of Gothenburg, Gothenburg, Sweden

\*These authors contributed equally to this work

Correspondence: Muwada Bashir Awad Bashir, Krefting Research Centre, Institute of Medicine, University of Gothenburg, P.O. Box 424, Gothenburg, SE-405 30, Sweden, Tel +46707847667, Fax +46 317863101, Email muwada.bashir@gu.se; mawadah99@yahoo.com

**Introduction:** Computational sciences have significantly contributed to characterizing airway disease phenotypes, complementing medical expertise. However, comparing studies that derive phenotypes is challenging due to varying decisions made during phenotyping. We conducted a systematic review to describe studies that utilized unsupervised computational approaches for phenotyping obstructive airway diseases in children and adults.

**Methods:** We searched for relevant papers published between 2010 and 2020 in PubMed, EMBASE, Scopus, Web of Science, and Google Scholar. Additional sources included conference proceedings, reference lists, and expert recommendations. Two reviewers independently screened studies for eligibility, extracted data, and assessed study quality. Disagreements were resolved by a third reviewer. An in-house quality appraisal tool was used. Evidence was synthesized, focusing on populations, variables, and computational approaches used for deriving phenotypes.

**Results:** Of 120 studies included in the review, 60 focused on asthma, 19 on severe asthma, 28 on COPD, 4 on asthma-COPD overlap (ACO), and 9 on rhinitis. Among asthma studies, 31 focused on adults and 9 on children, with phenotypes related to atopy, age at onset, and disease severity. Severe asthma phenotypes were characterized by symptomatology, atopy, and age at onset. COPD phenotypes involved lung function, emphysematous changes, smoking, comorbidities, and daily life impairment. ACO and rhinitis phenotypes were mostly defined by symptoms, lung function, and sensitization, respectively. Most studies used hierarchical clustering, with some employing latent class modeling, mixture models, and factor analysis. The comprehensiveness of variable reporting was the best quality indicator, while reproducibility measures were often lacking.

**Conclusion:** Variations in phenotyping methods, study settings, participant profiles, and variables contribute to significant differences in characterizing asthma, severe asthma, COPD, ACO, and rhinitis phenotypes across studies. Lack of reproducibility measures limits the evaluation of computational phenotyping in airway diseases, underscoring the need for consistent approaches to defining outcomes and selecting variables to ensure reliable phenotyping.

**Keywords:** asthma, COPD, severe asthma, rhinitis, unsupervised, phenotyping

## Introduction

Chronic obstructive airway diseases, such as asthma and COPD, are heterogeneous conditions that exhibit diverse clinical presentations due to a variety of endogenous and exogenous factors.<sup>1,2</sup> Obstructive airway diseases have distinct mechanistic pathways and heterogeneous clinical presentations known as phenotype.<sup>3</sup> Identification of specific phenotypes of airway diseases is important as this will help better to target therapies, personalize clinical interventions, and improve diagnostic accuracy.<sup>4</sup>

Over the past two decades, there has been an increase in the use of data-driven approaches in identifying phenotypes of chronic obstructive airway diseases.<sup>5</sup> These approaches rely on unsupervised methods to extract latent patterns of the disease that are not known beforehand.<sup>6</sup> This allows for the identification of disease subgroups that are more reflective of natural disease phenomena and that can guide clinical decision-making<sup>1,6</sup>. However, studies employing these methods, and the resulting phenotypes have been challenging to compare, perhaps due to differences in participants' profiles, study settings, phenotyping methods employed, and number and types of variables used.<sup>2,6</sup> To gain clear appreciation of the landscape of computational phenotyping of chronic obstructive airway diseases, a systematic synthesis of the underlying evidence is valuable. Through this, the methodological underpinning of studies can be ascertained, and the quality of evidence appraised, thus helping to identify potential research gaps in moving the field forward.

This review aimed at identifying, critically appraising, and synthesizing data from studies that have utilized computational approaches to phenotype chronic obstructive airway diseases in both children and adults. The review set out to characterize and compare the populations included in studies, assess and compare the criteria used to select participants, evaluate and compare the variables used to derive phenotypes of chronic airway diseases across studies, and assess the choices informing inclusion of variables. Additionally, the review described and compared the computational approaches used across studies and described and assess the number and characteristics of phenotypes derived across studies in terms of their clinical interpretation.

## Methods

### Protocol and Registration

We developed a protocol that outlined the review processes and methods before undertaking this work, which was registered in PROSPERO (CRD42020164898) and published.<sup>7</sup>

### Eligibility Criteria

Table 1 shows the full information on inclusion and exclusion criteria of studies into the review based on aspects of study design, setting, outcome, method of phenotyping, participants' age, study year and language.

**Table 1** Inclusion and Exclusion Criteria

	Inclusion Criteria	Exclusion Criteria
<b>Study design</b>	Observational, cross-sectional, cohort,	Interventional, randomized, and nonrandomized clinical trials, case studies and case series.
<b>Study setting</b>	General population Clinical setting	
<b>Outcome</b>	Asthma, Chronic obstructive pulmonary diseases, emphysema, rhinitis	Others
<b>Method of phenotyping</b>	Computational, unsupervised, or data driven	Hypothesis based approaches, supervised method, expert-based methods.
<b>Participants age</b>	Adults and Children.	
<b>Study Year</b>	From 2010 to 2020	Before 2010 and after 2020.
<b>Language</b>	All	

## Information Source

To identify relevant studies for the review, we searched PubMed, Embase, Web of Science, Scopus, and Google Scholar. For unpublished materials, such as conference proceedings, we searched databases of proceedings of conferences and databases of the literature, such as Open Grey. We also contacted experts in the field to request for any paper that was missed from our database searches. Finally, we screened the reference lists of included studies to identify any additional papers.

## Search Strategy

We developed search strategies for all the databases to identify relevant studies for the review. The search strategies (Supplementary file 1) were first developed in PubMed and then adapted in searching the other databases.

## Study Records

### Data Management and Selection Process

The search results from the different databases were exported to EndNote for screening. The first stage of the literature review involved removal of duplicates from the database searches; then, we performed title and abstract screening. Two reviewers independently screened the studies on the basis of the review inclusion and exclusion criteria; any discrepancies were resolved by discussion, or a third reviewer arbitrated if a consensus was not reached. The final stage involved full-text screening of the studies potentially meeting the eligibility criteria on the basis of the titles and abstracts. We documented the screening process using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart.<sup>8</sup>

### Data Collection Process

Reviewers, in pairs, independently extracted relevant data from included studies onto a data extraction form that was developed for the review; any discrepancies were resolved by discussion, or a third reviewer arbitrated if a consensus was not reached. We developed a data extraction form specifically designed for this review. The form was initially piloted on three included studies; any amendment was undertaken prior to using the form on all included studies.

### Data Items

Information on the following data items were collected from included studies into the data extraction form: general information (author's name, publication year and study time, aim of the study, and data source); information describing populations characteristics (population size, recruitment characteristics, sample size, children/adults, inclusion and exclusion criteria); type of chronic obstructive airway disease and how was the outcomes defined; information about the variables selected for phenotyping (number and description of variables, rational of selection, variable measurement and definition); type and features of computational approach used; and information of the derived phenotypes (number of phenotypes, characteristics of each phenotype, and clinical interpretation).

## Outcome and Prioritization

We included studies focusing on computational phenotyping of the following chronic obstructive airway diseases:

- Asthma
- COPD and asthma and COPD overlap
- Rhinitis
- Emphysema

## Quality Assessment of Included Studies

We appraised the general quality of included studies using an in-house developed checklist. Since, to our knowledge, there are no standard tools for assessing the quality of studies on computational disease phenotyping, we developed a checklist that enabled us to assess the quality of reporting specific aspects of the studies as they relate to performing a computational phenotyping. The aspects assessed were subjects' selection and inclusion in the phenotyping sample; missing data; outcome definition; variables included for the phenotyping; clinical and scientific relevance of the derived phenotypes; and reproducibility of the phenotyping process. To evaluate reproducibility, we examined aspects such as the disclosure of detailed information on methods used for phenotyping, computational aspects of data processing, and the utilization of software and tools for reproducible research frameworks. Detailed information and form of quality assessment can be found in the supplementary material.

## Data Synthesis

Data was narratively synthesized. We used tables and figures to summarize the results and different aspects of the studies, including study characteristics, methods of phenotyping, variables considered in phenotyping, counts of number of phenotypes, and description, as well as the results of the quality assessment.

## Deviation from the Study Protocol

None of the identified studies addressed emphysema as an outcome to be phenotyped. Instead, emphysematous changes as features of phenotyping obstructive airway diseases were reported within studies on COPD. Further, studies that included subjects with asthma and COPD overlap (ACO) were reported as separate outcome.

## Results

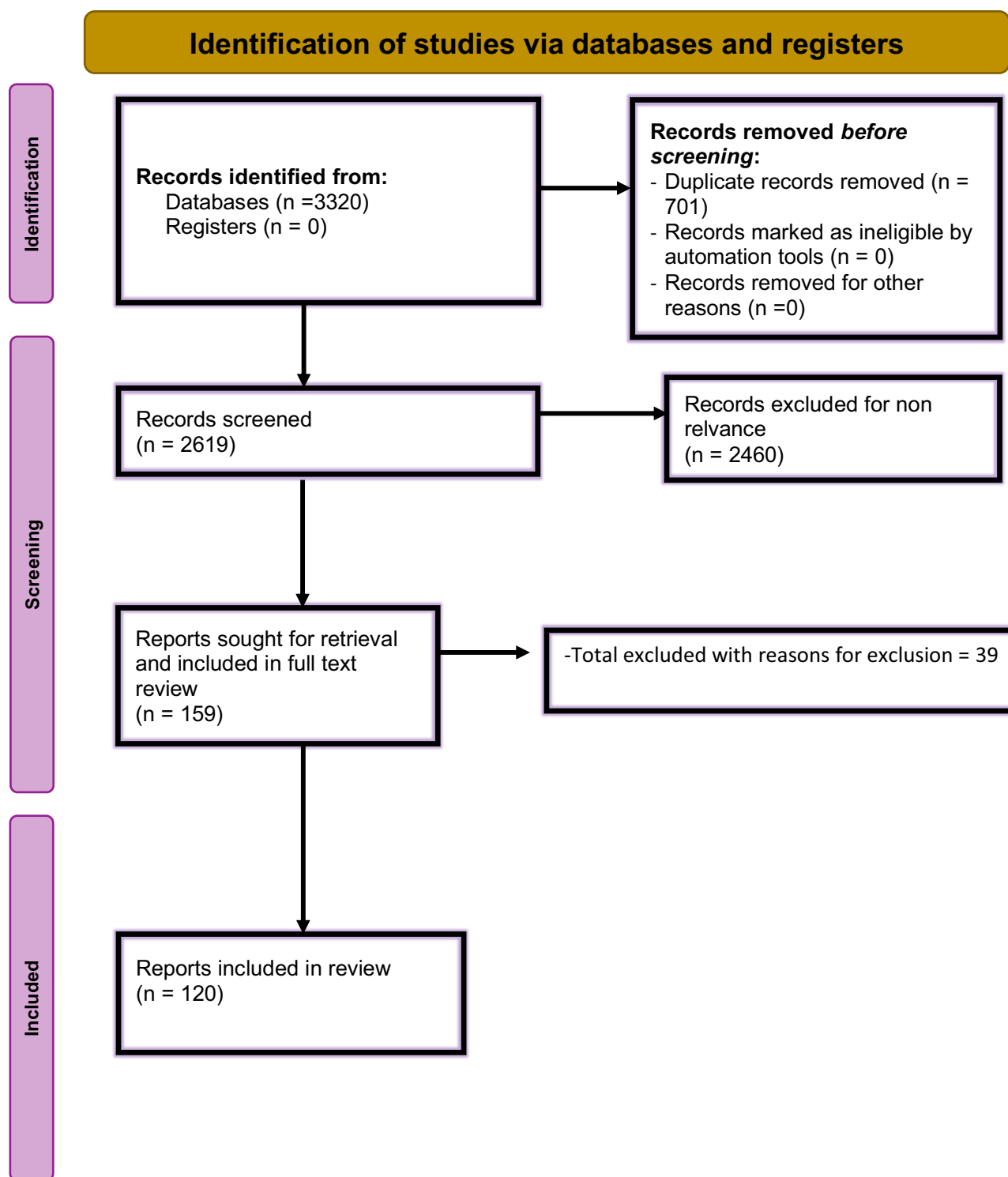
### Study Selection

A total of 3320 records were identified from the literature searches. After removal of duplicates, 2619 records were screened by title and/or abstract, of which 2460 records were excluded for not being eligible. A total of 159 records were considered for full-text review, of which 39 were excluded for different reasons, summarized in [Table S1](#) in the supplementary material. Finally, 120 studies were included in this review analysis. [Figure 1](#) shows the screening and selection of studies for this review.

## Study Characteristics

### Asthma

A total of 60 studies were on asthma.<sup>9–68</sup> The average number of subjects included in these studies was 1251, ranging from 50 to 9651 participants per study. The majority of studies were conducted among adults ( $n = 31$ )<sup>10,12–14,18–20,24,25,27,29–32,35–38,42–44,46–48,50,52,56,59–62</sup> and the remaining ( $n = 9$ )<sup>21–23,28,33,34,40,45,51</sup> in children, with remaining in mixed sample. Most studies were of cohort design ( $n = 30$ ),<sup>12,14,15,17,20,23,24,28,30–33,36–38,41,42,47,50,51,54–60,62,65,66</sup> while the rest were mostly cross-sectional ( $n = 17$ ).<sup>10,11,13,16,18,21,22,25,27,29,33,35,40,43,45,46,48</sup> Most studies were conducted in a clinical setting ( $n = 33$ )<sup>10–13,15,19–22,24,25,31,33–38,42,43,45,50,52–58,60–62,65</sup> with patients variously recruited from hospitals, pulmonary rehabilitation centers, and primary or tertiary care respiratory or general clinics. Studies with subjects selected from general population were 14,<sup>14,18,23,32,36,40,41,44,46,47,51,59,63,68</sup> while 8<sup>19,28,29,31,35,40,44,53</sup> studies did not report on the source of their participants. Full information on characteristics of studies in children and adults using unsupervised computational methods to phenotype asthma is presented in [Table 2](#).



**Figure 1** Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram illustrating the studies' selection process.

**Note:** This figure was adapted from Page, Matthew J., et al. 'The PRISMA 2020 statement: an updated guideline for reporting systematic reviews.' *bmj* 372 (2021). <https://doi.org/10.1136/bmj.n71>.

### Severe Asthma

A total of 19 studies<sup>69–87</sup> were on severe asthma. The average number of participants included in each study was 230, ranging from 40 to 1424 subjects per study. Most studies were conducted in a clinical setting (n = 17).<sup>69–72,74–77,79–87</sup> One study<sup>78</sup> in a general population setting and another study without a clear indication of setting.<sup>73</sup> Most were cohort studies

**Table 2** Characteristics of Studies in Children and Adults Using Unsupervised Computational Methods to Phenotype Asthma and Severe Asthma

	Reference and Country	Study design	Population and Participants	Population type	Study setting	Method used for phenotyping
	<b>ASTHMA</b>					
1.	Amaral et al 2019 US	Cohort study	Adults = 1059	Current asthma	General population	Latent class analysis
2.	Amelink et al 2013. Netherlands	Cross sectional	Adults n= 200	Patients with adult-onset asthma	Clinical setting	Nonhierarchical clustering
3.	Bhargava et al 2018. India	Cohort	Adults n = 113	Asthma based on GINA followed up for 6 months	Clinical setting	Hierarchical clustering
4.	Benton et al 2010. USA	Cross sectional	Children and adolescents n =154	Children with physician diagnosed asthma.	General population	Hierarchical and non- hierarchical clustering
5.	Bochenek et al 2014. Poland	Cross sectional	Adults n=201	Patients diagnosed with AERD by physicians	Clinical setting	Latent class analysis
6.	Boudier et al 2013. Multicounty	Cohort	Adults n=3320	Patients with asthma	General population	Latent class analysis
7.	Cabral et al 2017. Brazil	Cohort	Children and adolescents n =306	Children and nonsmoking adolescents (6–18 years of age) with a clinical diagnosis of asthma.	Clinical setting	Hierarchical clustering
8.	Celejewska-Wójcik et al 2020. Poland	Cross sectional	Adults n= 95	Patients with NERD based on provocation test and asthma based on GINA.	Not indicated	Latent class analysis
9..	Chanoine et al 2017. France	Cohort	Adults n= 4328	(20–75) years old women with reported ever having asthma.	General population	Hierarchical and non- hierarchical clustering
10.	Couto et al 2015; Portugal	Cross sectional	Adults n=150	Athletes with asthma diagnosis based on International Olympic Committee.	General population	Latent class analysis

11.	Cruz et al 2018. Brazil	Cohort	Adults n=966	Adults with asthma	Clinical setting	Nonhierarchical clustering
12.	Damiens K et al 2013. Canada	Cohort	Adults n=272	Work related asthma.	Clinical setting	Cluster analysis
13.	Deliu et al 2016. Turkey	Cross sectional	Children n = 201	Asthma patients, children aged 6–17 years.	Clinical setting	Hierarchical clustering
14.	Deliu et al 2018. Turkey	Cross sectional	Children n = 613	Children with asthma aged 6–18 years.	General population	Hierarchical clustering
15.	Depner et al 2013; Multicountry	Cohort	Children n = 953	Children from rural areas	Clinical setting	Latent class analysis
16.	Dudchenko et al 2018. Russia Russia	Not indicated	Adults n=300	Patients with asthma	Clinical setting	Cluster analysis
17.	Folz et al 2018. USA	Cohort	Adults n= 136	Asthma patients	General population	Hierarchical clustering
18.	Fontanella et al 2018. UK	Cross sectional	Children = 461	Children with available component resolved diagnostic data	General population	Hierarchical clustering
19.	Gonem S et al 2012. UK	Not indicated	Adults n=114	Asthma patients and healthy controls.	Not indicated	Hierarchical and non- hierarchical clustering
20.	Gower WA et al 2013. USA	Cohort	Children n = 942	Mild, moderate, or severe asthma.	Not indicated	Hierarchical clustering

(Continued)

Table 2 (Continued).

	Reference and Country	Study design	Population and Participants	Population type	Study setting	Method used for phenotyping
21.	Hilvering et al 2015; Netherland	Cross sectional	Adults n=115	Adult asthma patients.	Not indicated	Nonlinear Principal Component Analysis followed by discriminant analysis.
22.	Hsiao HP et al 2018; Taiwan	Cohort	Adults n=720	Age > 20 years, with stable physician diagnosed, mild-to-severe asthma.	Clinical setting	Orthogonal varimax factor analysis
23.	Ilmarinen P et al 2017. Finland	Cohort	Adults n=171	White patients with new onset adult-onset asthma diagnosed at age > 15 years.	Clinical setting	Hierarchical clustering
24.	Jeong A et al 2017. Switzerland	Cohort	Adults n=959	Self-reported asthma.	General population	Latent class analysis
25.	Just J et al 2012. France	Cohort	Children n = 315	Aged 6–12 years at the time of exploration; a history of asthma based on lung function; symptoms examined while stable	Clinical setting	Hierarchical clustering
26.	Just J et al 2014. France	Cross sectional	Children=125	Children aged between 6 to 12 years, having allergic asthma	Clinical setting	Hierarchical clustering
27.	Kaneko Y et al 2013. Japan	Cross sectional	Adults n=880	Physician diagnosed adult asthma based on either symptoms or spirometry measures or both.	Clinical setting	Hierarchical and non- hierarchical clustering
28.	Kim HJ et al 2018. Korea	Cohort	Adults n=1679	Asthma patients aged 18 to 79 years	General population	Hierarchical clustering
29.	Kim MA et al, 2017. Korea	Cohort	Adults n=259	Physician diagnosed asthma patients according to GINA.	Clinical setting	Hierarchical and non- hierarchical clustering
30.	Kim TB et al 2013. Korea	Cohort	Adults n=2567	Subjects with dyspnea, cough, sputum production or wheezing for >3 months, positive BHR, and stable conditions with regular medications	Clinical setting	Hierarchical and non- hierarchical clustering
31.	Koike et al 2018. (abstract) Not indicated.	Not indicated	Adults n = 458	Subjects with asthma presenting with cough and no other microbial infection.	Clinical setting	Hierarchical cluster analysis

32.	Kwon et al 2012. Korea	Cross sectional	Children n = 193	Children with physician diagnosed asthma	General population	Hierarchical and non- hierarchical clustering
33.	Lee E et al 2017. South Korea	Cohort	Children n=235	Children 6–8 years with parent reported physician diagnosed asthma in lifetime	Clinical setting	Latent class analysis
34.	Li et al 2016. USA	Cross sectional	Adults n= 2081	GOLD stage 0–4, smoking≥20 packs/year	Not indicated	Cluster analysis
35.	Loureiro et al 2015. Portugal	Cross sectional	Adults n=57	Asthma patients, aged >18 years; diagnosed according to GINA guidelines.	Clinical setting	Hierarchical cluster analysis
36.	Loza et al 2016. Multicounty	Cross sectional and cohort	Adults n=238	Not indicated	General Population	Non- hierarchical clustering
37.	Mahut et al 2011. France	Cross sectional	Children n = 169	Children who satisfy criteria of clinical and functional asthma.	Clinical setting	Nonhierarchical clustering
38.	Mäkikyrö et al 2017. Finland	Cross sectional	Adults n=1995	Asthma patients according to the Social Insurance Institution of Finland criteria who received reimbursement right for asthma medication.	General population	Latent class analysis
39.	Mason et al 2018. Italy	Cohort	Adults n=187	Subjects with diagnosis of occupation asthma based on (SIC) result to diisocyanate.	General population	Hierarchical and non- hierarchical clustering
40.	Mastalerz I et al 2015. Poland	Cross sectional	Adults n=137	Patients with asthma, aspirin tolerant asthma, severe asthma according to ATS/ERS recommendation of 2013, atopic asthma, and asthma with rhinosinusitis.	Not indicated	Latent class analysis
41.	Nadif et al 2018. France	Cohort	Adults n = 318	Adult with current asthma	Clinical setting	Mixture models
42.	Nagasaki et al 2014. Japan	Cohort	Adults n=224	Patients with stable asthma.	Clinical setting	Hierarchical clustering
43.	Nasreen et al 2019. Canada	Cohort	Children n = 403	Asthma patients based on parents' report.	General population	Latent class analysis

(Continued)

Table 2 (Continued).

	Reference and Country	Study design	Population and Participants	Population type	Study setting	Method used for phenotyping
44.	Qui et al 2018. China	Not indicated	Adults n=218	Asthma based on symptoms or physician diagnosis according to GINA. Severity defined based on need for SCS.	Clinical setting	Hierarchical and non- hierarchical clustering
45.	Sakagami et al 2011. (abstract) Japan	Not indicated	Adults = 591	Patients from primary, secondary, and tertiary care centers.	Clinical setting	Cluster analysis
46.	Schatz et al, 2013. USA	Cohort	Children and adolescents and adults n =4130	Asthma patients aged $\geq 15$ years according to ATS.	Clinical setting	Hierarchical clustering
47.	Schimdlin et al 2015. (Abstract) USA	Cohort	Children n = 72	Asthma patients based on parents' report and lung function test.	General population	Hierarchical clustering
48.	Seino et al 2018. Japan	Cohort	Adults n= 2273	Adults with asthma diagnosis according to GINA.	Clinical setting	Hierarchical clustering
49.	Sendín-Hernández et al 2018. Spain	Retrospective record-based study	Adults n= 225	Allergic asthma patients based on lung function test and symptoms ages 18–65 years (ATS) criteria.	Clinical setting	Hierarchical clustering
50.	Seys et al 2017. Belgium	Cohort	Adults, asthma =205, healthy =80	Asthma based on symptoms and reversibility test at age 18–65 years.	Clinical setting	Hierarchical clustering
51.	Siroux V et al 2011. Multi country	ECRHSII = Cohort	Adults n= 1895	Age > 14 years; asthma diagnosed following GEMA 2009,	General population	Latent class analysis
		EGEA2 = Case–control and family-based	Adults n= 641	Adults who had ever had asthma.	General population	Multivariate exploratory data clustering preceded by multiple correspondence analysis
52.	Tay et al 2019. Singapore	Cohort	Adults n=630	Asthma based on objective measure of airflow limitation or physician diagnosis in addition to asthmatics in stage 4 or 5 GINA treatment.	Clinical setting	Latent class analysis

53.	Tsukioka et al 2017. Japan	Cohort	Adults n=104	Elite athletes with asthma.	Clinical setting	Hierarchical clustering
54.	Wang LL et al 2017; (abstract) China	Cohort	Adults n=284	Clinically stable asthma patients.	Clinical setting	Hierarchical and non- hierarchical clustering
55.	Watanabe et al 2016 (abstract); Japan	Japan	Adults n=120	Non-smokers, aged $\geq 60$ with diagnosis of asthma.	Clinical setting	Hierarchical and non- hierarchical clustering
56.	Wisnivesky et al;2019 USA	Cohort	Adults = 330	Physician diagnosed asthmatic among population of exposed to 11th/9 disaster.	General population	Kamila Algorithm
57.	Wu et al 2017. China	Cohort	Adults n=110	Patients with nasal polyposis and comorbid asthma; 18–65 years of age	Clinical setting	Non- hierarchical clustering
58.	Zaihra T et al 2016; Canada	Cohort	Adults N =125, Severe asthmatic =77, moderate =48	Asthmatic patients according to GINA.	Clinical setting	Non- hierarchical clustering
59.	Zhang X et al 2019 (abstract); China	Retrospective case control	Adults n= 825	Patients with NPcA diagnosed by pathology after biopsy.	Clinical setting	Hierarchical clustering
60.	Zoratti E et al 2018. USA	Cohort	Children n = 717	Asthma and high dose or systematic steroids or symptoms, lung function, treatment dependency, low or moderate dose of ICS.	Clinical setting	Boruta feature selection algorithm
<b>Severe asthma</b>						
61.	Brinkman et al, 2011. multicountries	Cross sectional	Adults n = 77	Severe asthma based on the IMI criteria.	Clinical	Hierarchical cluster analysis
62.	Desai et al 2011. UK	Not indicated	Adults n = 164	Patients attending difficult asthma clinic	Clinical	Non-hierarchical clustering

(Continued)

Table 2 (Continued).

	Reference and Country	Study design	Population and Participants	Population type	Study setting	Method used for phenotyping
63.	Diver et al 2018. UK	Cohort	Adults n = 63	Severe asthma, moderate to severe COPD. maintenance oral corticosteroid therapy; subjects with sputum samples adequate for microbiome sequencing.	Clinical	Mixtur modelling
64.	Fitzpatrick et al 2018. USA	Cohort	Children n = 161	6 to 17 years of age, never smoked, physician diagnosed asthma based on spirometry at baseline or during exacerbation.	Clinical	Hierarchical cluster analysis
65.	Freitas PD et al 2018. (abstract); Brazil	Cross sectional	Adults = 119	Patients with moderate to severe asthma under optimal treatment according to GINA.	Not indicted	Hierarchical cluster analysis
66.	Gomez et al 2017. USA	Cohort	Adults n = 156	Subjects with severe asthma	General population	Hierarchical clustering
67.	Jang et al 2017. Korea	Cohort	Adults n = 86	Patients with refractory asthma	Clinical	Hierarchical and non-hierarchical clustering
68.	Konstantellou et al 2015. Greece	Cohort	Adults n = 170	Patients with asthma diagnosis followed up in 1st and 2 <sup>nd</sup> centers based on GINA.	Clinical	Hierarchical and non-hierarchical clustering
69.	Lau et al, 2017: Singapore	Cohort	Adults n = 55	Patients with asthma exacerbation of status asthmatics requiring ICU admission and intubation	Clinical	Cluster analysis
70.	Moore et al 2010; USA	Cohort	Adults n = 304	Refractory asthma.	General population	Hierarchical cluster analysis
71.	Newby et al 2018; UK	Cohort	Adultsn = 349	Severe asthma as OCS 50% of the year or high dose ICS plus add on medications either at baseline or follow-up.	Clinical	Mixture models
72.	Raherson et al, 2018; France	Cohort	Adults n = 1424	Not indicated	Clinical	Hierarchical and non-hierarchical clustering

73.	Sekiya et al 2015; Japan	Cohort	Adults n = 190	Stable, on maintenance asthma therapy 4 weeks prior to study, asthma diagnosis as per GINA and severity status as per (ERS) and (ATS).	Clinical	Non-hierarchical clustering
74.	Serrano Pariente et al 2015; Spain	Cohort	Adults n = 84	Asthma as per ATS criteria, Aged > 15 years	Clinical	Non-hierarchical clustering
75.	Simpson et al (abstract);2017; Multi-country	Cross sectional	Adults = 421	Severe asthma based on U-BIOPRED consensus criteria	Clinical	Not indicated
76.	Taniguchi et al; 2014; Japan	Cohort	Adults = 127	Severe refractory asthma based on ATS criteria without exclusion based on smoking history	Clinical setting	Unbiased cluster analysis
77.	Wu et al; 2014; USA	Cohort	Adults = 378	Severe asthma based on ATS criteria	General population	Non-hierarchical clustering
78.	Weng-Jing Ye et al 2017; China	Cohort	Adults n = 203	Severe asthma as per the IMI-criteria.	Clinical	Hierarchical and non-hierarchical clustering
79.	Youroukova et al 2017; Bulgaria	Cohort	Adults n = 40	Severe refractory asthma as use of OCS 50% of the year or high dose ICS plus add on medications at baseline or follow-up.	Clinical	Hierarchical cluster analysis

**Abbreviations:** AERD, aspirin exacerbated respiratory diseases; GINA, Global Initiative for Asthma; ATS/ERS, American Thoracic society/European respiratory society; SCS, systematic corticosteroids; NSAID, non-steroidal anti-inflammatory diseases; GEMA, Spanish asthma management guidelines; NPcA, nasal polyposis and comorbid asthma; ICS, inhaled corticosteroids; IMI, innovative medicine initiative; OCS, oral corticosteroids, SIC, specific inhalation challenge.

(n = 11),<sup>71,74–76,78–81,86,87</sup> while the remaining were cross-sectional studies. Characteristics of studies in children and adults using unsupervised computational methods to phenotype severe asthma are presented in [Table 2](#).

## COPD

A total of 28 studies<sup>4,88–114</sup> were on COPD. The average number of subjects per study was 5218, ranging from 46 to 104143 subjects per study. Most studies were conducted within a clinical setting (n = 17),<sup>4,88–93,95,97,98,100,107,109–113</sup> with cohort studies<sup>88,89,92–96,98,100,101,109–111,114</sup> being the most reported study design (n = 14), while the second common were of cross-sectional design (n = 7).<sup>4,90,91,97,107,112,113</sup> Full information on characteristics of studies using unsupervised computational methods to phenotype COPD and ACO is given in [Table 3](#).

## Asthma and COPD Overlap (ACO)

Four of the included studies were on asthma and COPD overlap. The average number of participants included in these studies was 255, ranging from 47 to 435 participants per study. All were cross-sectional studies. Three studies were conducted in a clinical setting,<sup>115–117</sup> while one was conducted in a general population setting.<sup>118</sup> Characteristics of studies in adults using unsupervised computational methods to phenotype COPD and ACO are presented in [Table 3](#).

## Rhinitis

A total of 9 studies were on rhinitis.<sup>119–127</sup> The average number of participants included in each study was 516, ranging from 115 to 1831 participants per study. Most studies were conducted in a clinical setting (n = 6),<sup>119–121,123,126,127</sup> while three studies<sup>122,124,125</sup> were conducted in the general population. Most were cohort studies (n = 7),<sup>119,121,123–127</sup> one was cross-sectional,<sup>120</sup> and one case–control study.<sup>122</sup> Characteristics of studies using unsupervised computational methods to phenotype rhinitis are presented in [Table 4](#).

## Phenotypes of Respiratory Diseases

### Asthma

In total, 251 phenotypes were reported in studies on asthma with considerable degree of overlap between them. In characterizing asthma phenotypes, atopy was the most common feature included in most studies,<sup>13–15,18,20,22,28,30,31,33,35,37,38,40,41,43,44,47,48,53,55,57,59–62,65,68</sup> resulting in differentiation of atopic and non-atopic asthma phenotypes (reported in 29 studies and featured in 100 of the reported asthma phenotypes). Atopic status was defined mostly based on skin prick test, serum IgE levels or subjects' report of familial atopy. Atopic asthma phenotype was reported in 28 studies,<sup>13–15,18,20,21,28,30,31,33,35,37,38,40,41,43,44,47,48,53,55,57,59–62,65,68</sup> while non-atopic asthma was reported in 22 studies.<sup>14,15,18,20–22,30,35–38,40,41,43,44,53,55,57,59,61,65,68</sup> The second feature was lung function measures, which featured in 85 of the reported phenotypes and was considered in 30 studies.<sup>10,14–16,18–22,24,27,28,30,33,36–38,40–44,50,52,55,60–62,65,68</sup> Time at asthma onset featured in 74 phenotypes and reported in 27 studies.<sup>11,13,15,18,21–23,28,30,31,35–38,43,44,48,50,52,53,55,59–61,65</sup> The definition of early and late onset asthma varied among different studies. When studying both children and adolescents, asthma that developed during childhood and adolescence was referred to as early onset while adulthood developed asthma as late-onset asthma. However, when examining only adults or children, researchers measured the average age at which asthma onset and the standard deviation across different phenotypes. In these cases, the terms early and late onset asthma were defined differently, with the groups having younger individuals labeled as early onset and those with older individuals classified as late-onset asthma. Early onset asthma phenotype was reported in 19 studies<sup>15,18,21,22,28,35,37,38,43,44,48,50,52,53,55,59–61,65</sup> while late – onset asthma was reported in also 18 studies.<sup>11,13,18,21–23,30,31,35,37,38,43,50,53,55,59,60,65</sup> Level of asthma control was also a commonly reported feature, occurring in 45 of phenotypes and reported in 17 studies.<sup>11,13,18,21,24,27,33,42–46,48,50,53,56,60</sup> Well-controlled asthma phenotype was reported in 11 studies,<sup>21,24,42–46,48,53,56,60</sup> while uncontrolled asthma was reported in 16 studies.<sup>11,13,18,21,24,27,33,42–44,46,48,50,53,56,60</sup>

Sex featured in 65 of the reported phenotypes. Female asthma phenotype was reported in 20 studies,<sup>10,11,13,16,18,20,21,30,31,35,42,43,45,48,52,53,55,60,61,63</sup> while male asthma phenotype featured in 19 studies.<sup>10,11,13,18,20,22,28,30,40,42,43,45,47,52,53,55,60,61,63</sup> Eleven studies reported on obesity-related asthma phenotypes.<sup>10,11,13,19,30,31,33,43,48,56,60</sup>

**Table 3** Characteristics of Studies Using Unsupervised Computational Methods to Phenotype COPD

	Reference and Country	Study Design	Population and Participants	Population Type	Study Setting	Method Used for Phenotyping
	<b>COPD</b>					
1.	Augustin et al; 2018 Netherland	Cohort	n =518	Clinically stable COPD	Clinical population	Self-Organizing Maps
2.	Bafadhel et al (abstract). 2011. UK	Cohort	n =86	Physician diagnosis COPD based on spirometry, age> 40, GOLD stage I–IV and current sever exacerbation.	Clinical population	Cluster analysis
3.	Bertini et al; 2013. Italy	Cross sectional	n =62	Clinically confirmed exacerbation within 3 months prior to study.	Clinical population	Non- hierarchical clustering
4.	Burgel et al; 2017. France	Cohort	n =2409	Subjects with COPD diagnosed based on spirometry whether stable or on exacerbation.	Clinical population	Classification and regression tree
5.	Burgel et al; 2010. France	Cross sectional	n =322	Stable COPD diagnosed based on spirometry.	Clinical population	Hierarchical clustering
6.	Burgel et al; 2012 Belgium	Cross sectional	n =527	Smoking history ≥15 pack-years, age > 50 years, and 154 patients diagnosed of COPD based on spirometry.	General and clinical population	Hierarchical clustering
7.	Chen et al; 2014 Taiwan and China	Cohort	n =332	Men diagnosed with COPD, ≥40 years old, based on symptoms and spirometry.	Clinical population	Cluster analysis
8.	Chubachi et al (abstract); 2016 Japan	Cohort	n =311	COPD and completed data.	Not Indicated	Hierarchical clustering
9.	De Torres et al; 2017 Spain	Cohort	n =521	Active and former smokers with COPD.	Clinical population	Non- hierarchical clustering
10.	Divo et al (abstract); 2016 Not indicated	Cohort	n =120	Patients with COPD.	Not Indicated	Hierarchical clustering
11.	Fens et al; 2013. Netherland	Cross sectional	n =157	Smoking history of at least 15 pack-years, COPD and chronic bronchitis based on GOLD.	General population	Hierarchical clustering
12.	Guillamet RV et al; 2018 USA	Cohort	n =3144	>40 years, COPD, and ever hospital admission	Clinical population	Sphere exclusion method

(Continued)

Table 3 (Continued).

	Reference and Country	Study Design	Population and Participants	Population Type	Study Setting	Method Used for Phenotyping
13.	Harrison SL et al; 2014 UK	Cohort	n =92	Asthmatics based on spirometry and COPD as per signs and symptoms, age > 40 years, and smoking history.	Clinical population	Hierarchical and non-hierarchical clustering
14.	Haghighi et al, 2019. USA	Cross sectional	n= 406	Former smokers	General population	Non- hierarchical clustering
15.	Kim WJ et al; 2017 Asia	Cohort	n =1676	Asian, age 40 years and with COPD based on post bronchodilator spirometry.	General population	Hierarchical clustering
16.	Kim S et al 2017; Korea	Cohort	n = 272	Age > 40 years and (FEV1/FVC) <0.7	General population	Hierarchical clustering
17.	Kukol et al 2019; (abstract). Russia	Not indicated	Not indicated	Elderly patients with COPD	Not indicated	Cluster analysis
18.	Lee et al, 2019. Korea	Cohort	N = 1195	Patients with available follow up data on the first acute exacerbation of COPD.	Clinical	Non- hierarchical clustering
19.	li et al abstract;2016. USA	Not Indicated	n =2081	GOLD stage 0–4, smoking ≥20 packs/year	Not Indicated	Hierarchical clustering
20.	Liang et al, 2019. Korea	Prospective cross sectional	n = 102 patients n = 18 controls	Age > 40, COPD diagnosis based on GOLD, post BD EFV1 <0.7.	Clinical population	Hierarchical clustering
21.	Lopes et al, 2019. Brazil	Cross sectional	n =150	Stable COPD diagnosis within the last 30 days, based on GOLD	Clinical population	Hierarchical and non-hierarchical clustering
22.	Ning et al; 2016, (abstract). China	Not indicated	Not indicated	Subjects with wheeze for more than 12 months.	General population	Hierarchical clustering
23.	Peters et al; 2016. Netherlands	Cohort	n =160	COPD based on spirometry and GOLD grades 2–3, receiving treatment as usual.	Clinical population	Hierarchical clustering
24.	Pikoula et al; 2019 UK	Cohort	n =30961	Age >35 year, with at least one diagnostic code of COPD based on GOLD and complete data.	Clinical population	Hierarchical and non- hierarchical clustering
25.	Rodrigues et al; 2018. Brazil	Cohort	n =141	Stable COPD, with no sever cardiac or musculoskeletal comorbidity	Clinical population	Non- hierarchical clustering

**Abbreviations:** GOLD: global initiative for chronic obstructive lung disease, GINA: global initiative of asthma.

**Table 4** Characteristics of Studies Using Unsupervised Computational Methods to Phenotype Rhinitis

	Rhinitis					
	Reference and Country	Study Design	Population and Participants	Population Type	Study Setting	Method Used for Phenotyping
1.	Adnane et al 2017. Morocco	Cohort	Adults n = 131	Patients with medical refractory chronic rhinosinusitis (CRS).	Clinical setting	Discriminant analysis
2.	Agache et al 2010. Romania	Cohort	Adults and children n = 115	Adults and children (aged 5–11 years), physician diagnosed SAR, asthma as per symptoms and spirometry test. SAR as per symptoms SPT and at least 1 of seasonal allergens.	General population	Non-hierarchical clustering.
3.	Bousquet et al 2015. France	Case control and family study	Adults n = 825	Age >18 years, diagnosis of AR and symptoms at time of examination.	General population	Hierarchical clustering
4.	Burte et al 2015. France	Cohort	Adults n = 983	Age ≥16 years, complete data, allergic sensitization, and asthma	Clinical setting	Mixture models
5.	Herr M et al 2012. France	Cohort	Infants n = 1831	Singleton full-term newborns, normal birth weight and an uncomplicated birth and neonatal period with parents who can commit to and participate in study.	Clinical setting	Hierarchical clustering
6.	Kurukulaaratchy et al 2015. UK	Cohort	Adolescents n = 468	Self-report of current rhinitis at age 18.	Clinical setting	Non-hierarchical clustering.
7.	Lee E.L. et al 2016. Korea	Cohort	Children n = 512	Children with parental-reported, physician- diagnosed rhinitis and symptoms of rhinitis in the previous 12 months.	General population	Latent class analysis
8.	T.A Nakayama et al 2012. Japan	Cohort	Adults n = 425	CRSnP as per the European position paper and non-response to medical treatment and undergone surgery.	Clinical setting	Non-hierarchical clustering.
9.	ZM Soler et al 2015. USA and Canada	Cohort	Adults n = 382	Consensus criteria: 3 months of at least 2 cardinals' symptoms and evidence of inflammation on sino-nasal endoscopy and CT scan, and active symptoms after initial treatment.	Clinical setting	Hierarchical clustering

**Abbreviations:** SAR, seasonal allergic rhinitis; SPT, skin prick test; AR, allergic rhinitis; CT, computed tomography; RSnP, rhinosinusitis with nasal polyposis.

Disease activity was characterized variously across asthma studies, based on either symptoms' activity or disease severity. Frequency of symptoms and rate of exacerbation featured phenotypes of high or low symptoms' activity, while disease severity defined using standard criteria of asthma severity characterized phenotypes of mild, moderate or severe asthma. Severe asthma phenotypes as indicated by investigators were reported in 12 studies,<sup>10,16,18,20–22,28,34,35,48,56,67</sup> while 20 studies<sup>10,14,15,18,21–24,33,41–43,50,55,57,59,60,62,66,68</sup> reported on asthma phenotypes with high symptoms or exacerbation rates. Across identified studies, labeling a phenotype as severe was not entirely based on standard GINA criteria for defining severe asthma or physician decision, although some studies applied such approach.<sup>21,33,34,56</sup> Otherwise, most investigators identified severity of phenotypes based on symptom frequency, need for high dosage of treatment and disease impairment of daily life, with no clear reporting on how severity was defined.<sup>10,16,20,28,35,48,67</sup>

Inflammation was considered in deriving asthma phenotypes using different indicators like inflammatory cell counts in peripheral or sputum induced samples, fractional exhaled nitric oxide (FeNO),<sup>10,18,44</sup> or measure of inflammatory cytokines.<sup>58</sup> A total of 36 phenotypes were described based on high or low levels of eosinophilic inflammatory cells in sputum or peripheral blood. Some of those were reported in 17 studies<sup>10,11,15,16,20,21,30,33,40,41,43,44,48,50,52,58,65</sup> as asthma with high eosinophilia. Variants of neutrophilic asthma phenotypes, in turn, were less commonly reported in 10 studies.<sup>11,13,30,33,42–44,50,52,58</sup> See full results on number of derived phenotypes and their descriptions for studies on asthma and severe asthma in Table 5.

**Table 5** Number of Derived Phenotypes and Their Descriptions for Studies on Asthma

Reference, Country	Population Age	Number of Phenotypes	Phenotype Description
Amaral et al 2019. USA	Adults	3	High symptoms with poor lung function and more emergency visits among those > 40 years. Low symptoms with better lung function and less emergency visits among those < 40 years Low symptoms with better lung function and less emergency visits among those > 40 years
Amelink et al 2013. Netherlands	Adults	3	Severe eosinophilic inflammation-predominant asthma and persistent airflow limitation despite high-intensity anti-inflammatory treatment, with relatively low symptom scores Obese female, frequent symptoms, high HCU and low eosinophilia. Males with mild to moderate, well controlled asthma and low HCU, low symptoms and normal function.
Bhargava et al 2018. India	Adults	4	Treatment responsive, normal weight, childhood asthma. Poor treatment response, obese, males with late onset disease. Poor treatment response, old, male, with late-onset, obese. Good treatment response, obese females with late onset disease.
Benton et al 2010. USA	Children and adolescents	4	Male neutrophilic asthma Female obese, late onset asthma Atopic, uncontrolled, obese eosinophilic asthma Mild asthma
Bochenek et al 2014. Poland	Adults	4	Asthma with a moderate course, intensive upper airway symptoms, and blood eosinophilia. Asthma with a mild course, relatively well controlled, and with low health care use. Asthma with a severe course, poorly controlled, and with severe exacerbations and airway obstruction. Poorly controlled asthma with frequent and severe exacerbations in female subjects.
Boudier et al 2013. Multicountry	Adults	7	Mild atopic with no treatment needed. Non atopic mild with no treatment need Non atopic sever with treatment needed. Atopic, severe with high reactivity and treatment need. Atopica moderate asthma with high reactivity Atopic moderate asthma with normal function. Non atopic moderate with no need to treatment.

(Continued)

Table 5 (Continued).

Reference, Country	Population Age	Number of Phenotypes	Phenotype Description
Cabral et al 2016. Brazil	Children and adolescents	3	Normal lung function, mild eosinophile, late onset, mild atopy and few exacerbations.
			Normal lung function, moderate atopy, more severe eosinophilia and early onset disease.
			Poor lung function, frequent exacerbations, severe eosinophilia, and severe atopy.
Celejewska-Wójcik et al 2020. Poland	Adults	3	Younger females with mild paucigranulocytic, low eicosanoids and low pro inflammatory markers.
			Older Females with severe, eosinophilic asthma and high proinflammatory markers.
			Old women, obese, mild eosinophilic asthma with high pro and anti-inflammatory markers
Chanoine et al 2017. France	Adults	5	Never regular treatment
			Persistence high ratio of maintenance therapy
			Increasing ration of maintenance therapy from low to high.
			Initiating maintenance therapy at high level
			Cessation of maintenance therapy during study period.
Couto et al 2015. Portugal	Adults	2	Atopic asthma with rhinitis, allergic comorbidities and high FeNO.
			Sport asthma: non-allergic exercise induced asthma with high BHR.
Cruz et al 2018 (abstract). Brazil	Adults	4	Normal baseline function with less treatment needed.
			Low baseline function with obesity.
			Low baseline function and high medication and reversibility.
			Moderate baseline impairment and low revisability.
Damiens K et al 2013. Canada	Adults	6	Old sever atopic WRA with rhinitis. <b>(Clusters were identified in subjects with WRA)</b>
			Young mild atopic, eosinophilic OA with rhinitis with exposure to HMW agents.
			Non atopic, men, with WEA with exposure to LMW.
			Atopic, LMW agents' exposure and delayed asthma reaction and low occupational rhinitis. <b>(Clusters were identified in the OA subgroup)</b>
			HMW exposure, immediate long duration reaction.
			Nonsmoking women with normal function and immediate eosinophilic reaction.
Deliu et al 2016. Turkey	Children	6	Early onset mild asthma with impaired function and high medication use.
			Non atopic children with normal function and controlled asthma.
			Late onset eosinophilic asthma with impaired function, high medication uses and sensitization.
			Elderly with late onset severe, poor controlled asthma.
			Elderly with mild asthma and impaired function
			Mild atopic asthma with normal function

(Continued)

**Table 5** (Continued).

Reference, Country	Population Age	Number of Phenotypes	Phenotype Description
Deliu et al 2018. Turkey	Children	5	High symptom and medication use. <b>(HC after dimensionality reduction)</b>
			Males with late onset severe disease with normal function.
			Late onset, mild asthma with impaired function and multiple sensitizations.
			Female with early onset, mild, atopic asthma.
			Mild atopic asthma
			Female, early onset severe asthma. (HC using all available variables).
			late onset mild atopic asthma
			Moderate, highly atopic asthma,
			Male mild non atopic asthma.
			Late onset, atopic, severe asthma.
			Difficult asthma. <b>(HC using the informative set of features)</b>
			Early-onset mild atopic asthma.
			Early-onset mild non-atopic asthma.
			Late-onset asthma.
			Exacerbation-prone asthma.
Depner et al 2013. Multicountry	Children	5	Persistent wheeze.
			Late-onset wheeze.
			Intermediate wheeze.
			Transient wheeze.
			No/infrequent wheeze.
Dudchenko et al 2018. Russia	Adults	7	Mild symptoms, infrequent attacks relieved by SABA, moderate S of B. and limited physical activity and sensitivity to weather changes.
			Uncontrolled asthma, severe inflammation, daily attacks and severely impaired physical activity and night sleep
			Uncontrolled disease, high catarrhal manifestation, daily frequent attacks, moderate symptoms with high sleep and activity impairment.
			Mild, well controlled, minimal inflammation, low symptoms severity and frequency, low need for SABA and minimal impairment of PA and night sleep.
			Controlled to partially controlled, moderate catarrhal manifestation, within normal lung function, with excretion related limitation of PA and moderate symptoms.
			Uncontrolled asthma, moderate catarrhal symptoms, occasional symptoms, and usage of low dose BD and mild to moderate lung function impairment.
			Variable control, moderate inflammatory features, frequent breathlessness and diverse pattern of nocturnal disturbance, high dependence on weather changes and severely impaired lung function.
Folz et al 2018. USA	Adults	3	Three clusters that shows significant difference in BMI, FEV1%, FVC, FEV1, exercises induced asthma, asthma control, age at asthma onset and Feno.

(Continued)

Table 5 (Continued).

Reference, Country	Population Age	Number of Phenotypes	Phenotype Description
Fontanella et al 2018. UK	Children	4	Multiple sensitizations with sIgE to multiple components
			Predominantly dust mite sensitization (sIgE responses mainly to components from C. sIgE-1)
			Predominantly grass and tree sensitization (sIgE to multiple components across C. sIgE-4–7)
			Lower grade sensitization.
Gonem S et al 2012 (abstract). UK	Adults	2	High air entrapment, abnormal lung mechanics, limited function, poor control and QOL.
			Normal lung physiology compared to I.
Gower WA et al 2013 (abstract). USA	Children	5	Non-Hispanic white, early onset and normal baseline function.
			Hispanic, early onset and low baseline function.
			Obese, late onset, low baseline function.
			Males with low baseline function.
			African American, obese, atopic, high disease duration and hospitalization.
Hilvering et al 2015 (abstract). Netherlands	Adults	Non indicated	Eosinophilic and non-eosinophilic asthma phenotypes that are differentiated by peripheral blood eosinophil count, FeNO (Fraction of Exhaled Nitric Oxide), ACQ (Asthma Control Questionnaire), medication use, nasal polyposis, aspirin sensitivity and neutrophil/eosinophil responsiveness upon stimulation.
Hsiao et al 2018. Taiwan	Adults	6	Females, late onset, non-atopic asthma with normal function and low inflammation.
			Females, young, atopic, eosinophilic, low neutrophilic asthma.
			Females, late onset, non-atopic, obese, low eosinophilic, high neutrophilic asthma.
			Males, late onset, non-atopic, low eosinophilic and normal function.
			Males, young, atopic, smoking, neutrophilic asthma.
			Males who are former smokers with late onset high eosinophilic disease.
Ilmarinen et al 2017. Finland	Adults	5	Non-rhinitis asthma.
			Smoking asthma.
			Female asthma.
			Obesity-related asthma.
			Early-onset atopic adult asthma.
Jeong, A. et al 2016. Switzerland	Adults	4	Women, persistent multiple symptoms, late onset asthma with high percent body fat.
			Symptom presenting asthma with obesity.
			Symptom-free atopic asthma, atopic, high rhinitis,
			Symptom-free non-atopic asthma.
Just, J. et al 2012. France	Children	3	Atopic, severe, uncontrolled, eosinophilic, and basophilic, with high treatment asthma.
			Obesity, non-atopic, low function, neutrophilic asthma.
			Mild asthma

(Continued)

**Table 5** (Continued).

Reference, Country	Population Age	Number of Phenotypes	Phenotype Description
Just, J. et al 2014. France	Children	4	Multiple Allergies and Severe Asthma.
			Pollen Sensitization with Severe Exacerbations.
			Multiple Allergic Sensitizations and mild Asthma.
			House dust mite (HDM) Sensitization and Mild Asthma.
Kaneko, Y. et al 2013; Japan	Adults	6	Late-Onset, mild, less-Atopic.
			Early-Onset, mild, atopic.
			Early-Onset, moderate-to-Severe Atopic.
			Late-Onset, severe.
			Middle-Age onset, female-Dominant.
			Late-Onset, moderate Less Atopic
Kim JH. et al 2018; Korea	Adults	5	Persistent normal lung functions, female, middle/old age, obese, low treatment, and low symptoms.
			Persistent normal lung functions, young, female, atopic, low dose treatment, and low symptoms.
			Mild baseline impairment with slight improvement, low function, early onset, long duration, high dosage treatment and higher symptoms.
			Baseline marked impairment with fast improvement, atopic, high symptoms, low reversibility, and high dosage treatment.
			Marked baseline impairment with slow improvement, non-atopic, low reversibility, long duration, high symptoms, and high dosage treatment usage.
Kim MA. et al 2017; Korea	Adults	4	Young, early onset, atopic asthma with normal function.
			Elderly, late onset, non-atopic asthma with low function.
			High atopic with severely impaired function.
			Elderly, late onset, slight atopy, and normal function asthma.
Kim TB. et al 2013; Korea	Adults	4	Male, late onset, smokers, with preserved lung function.
			Atopic, late onset, high HCU, low function, low reversibility, and high severity.
			Young, early onset, and atopic asthma.
			Older, late onset, mild asthma with high function.
Koike et al 2018 Not indicated.	Adults	4	High FeNo asthma
			Low value of exhaled R5-R20
			Abnormal exhaled reactance.
			Abnormal resistance and reactance.
Kwon et al 2012 (abstract). Korea	Children	3	Atopic asthma
			Male, eosinophilic asthma.
			Non-atopic asthma

(Continued)

Table 5 (Continued).

Reference, Country	Population Age	Number of Phenotypes	Phenotype Description
Lee E et al 2017. South Korea	Children	4	Early onset, atopic asthma with mild symptoms and life impairment.
			Mild asthma with infrequent symptoms attacks and high SES, normal function and low BHR.
			Atopic, frequent symptoms attacks, daily impairment, eosinophilic, low SES and high BHR.
			Non atopic, infrequent symptoms attacks with minimal daily impairment, low eosinophilia and BHR and normal lung function.
Liang et al 2016. China	Adults	3	Males, high basophilic, uncontrolled, high anti (IL- 10, TGF-b, and sRAGE) and pro inflammatory markers (IFN-g, IL- 4, IL-5, IL-6, IL-9, IL-17, IL-23, EGF, GM-CSF, and TNF-a) and high baseline function.
			Low pro (INF-g, IL-4, IL-5, IL-6, IL-8, IL-9, IL-13, IL-17, IL- 23, EGF, GM-CSF, TNF-a, and VEGF) and anti-inflammatory markers (IL-10 and sRAGE) uncontrolled, neutrophilic, and basophilic with high baseline function.
			Females, controlled, neutrophilic, low basophilic, low baseline function and high leptin and VEGF levels but low sRAGE.
Loureiro et al 2015. Portugal	Adults	5	Young males, normal weight, well controlled, early onset, mild asthma, high function and low HCU.
			Old females, obese, long duration, good control and less severity, neutrophilic asthma.
			Young females, early onset atopic asthma, severe and frequent symptoms, no eosinophilia, normal function
			Females, obese, late onset, non-atopic, mixed low inflammation, uncontrolled, high impairment on daily life, high medication and HCU and normal lung function.
			Male, smoking, atopic, obese, late onset, aspirin sensitive, comorbid, long duration, eosinophilic, severe asthma with low function and high medication and HCU.
Loza et al 2016. Multicountry	Adults	4	Early onset, mild disease, low BDR, low AHR, and low inflammation and predominant T2 high inflammation.
			Atopic, moderate to severe, mild reversable, eosinophilic, high BHR asthma with high T2 inflammation.
			Non atopic, controlled, low function, neutrophilic, with mixed symptoms severity, with T2 low inflammation
			Non atopic, severe, uncontrolled, high BDR and AHR with mixed inflammation with moderate T2 inflammation.
Mahut et al 2011. France	Children	4	Nonsmoking exposed male with well controlled asthma.
			Nonsmoking exposed females with well controlled asthma.
			Nonsmoking exposed, well controlled disease with high airways tone.
			Parental smoking, well controlled asthma with small airways and lung size ratio.
Mäkikyrö et al 2017. Finland	Adults	5	Well controlled mild asthma
			Partially controlled mild asthma
			Partially controlled moderate severity asthma.
			Uncontrolled with mixed severity.
			Uncontrolled severe asthma.

(Continued)

**Table 5** (Continued).

Reference, Country	Population Age	Number of Phenotypes	Phenotype Description
Mason et al 2018. Italy	Adults	3	Male, nonsmoking, overweight, toluene diisocyanate (TDI) sensitized, late specific inhalation challenge (SIC) response with long exposure duration.
			Sensitization to TDI and late SIC response.
			Atopic, sensitized to methylene diisocyanate (MDI) with early SIC response.
Mastalerz et al 2015. Poland	Adults	4	High CRS, high atopy, mixed inflammation, mild to moderate intermittent, mostly controlled, low dose of ICS and higher levels of PGE2.
			Mostly AERD, females, High CRS, severe uncontrolled, eosinophilic, Aspirin sensitive, highest PGD2, LTE4, LTD 4 in ISS.
			Atopic, controlled, paucigranulocytic inflammation, no CRS, and low levels of proinflammatory mediators.
			Atopic, female, obese, early onset, severe, uncontrolled asthma, with low proinflammatory mediators, cyclLTs and PGD2, and high levels of PGE2.
Nadif et al 2018. France	Adults	3	Adult-onset asthma, poor lung function, treatments, cough and phlegm, exacerbations, high neutrophil count, and high fluorescence's oxidative products (FIOPs) level.
			Paucigranulocytic asthma, normal lung function, rhinitis and low IgE level.
			Predominantly men with childhood-onset asthma, eosinophilic asthma, allergic comorbidities, and high IgE level.
Nagasaki et al 2014. Japan	Adults	4	Late onset, paucigranulocytic asthma with high function.
			Early onset, mild eosinophilic asthma.
			Late onset, severe, eosinophilic with serum Periostin.
			Uncontrolled, severe, serum neutrophilic, high IL6 and high comorbidities.
Nasreen et al 2019. Canada	Children	3	Low initial attack rates that increase to high.
			Medium initial attack rates that decrease to none.
			High initial attack rates that decrease to medium.
Qui et al 2018. China	Adults	4	Young, early onset, sputum neutrophilia and low eosinophilia, with moderate function impairment.
			Female, severe disease, eosinophilic, hypoxemic asthma with impaired function.
			Females, elderly, neutrophilic asthma with moderate to severe function impairment.
			Male, smoking, mixed inflammation asthma with moderate to severe impairment.
Sakagami et al 2011. (abstract) Japan	Adults	5	Female, early onset, controlled long term asthma.
			Female, uncontrolled, and high depression.
			Female, atopic, late onset well controlled asthma.
			Elderly, female, non-atopic late onset asthma.
			Chinese, elderly, female, late onset well controlled asthma.

(Continued)

Table 5 (Continued).

Reference, Country	Population Age	Number of Phenotypes	Phenotype Description
Schatz et al 2013. USA	Adults and children	5	<b>Children</b>
			Atopic white male patients with no smoking exposure, normal function
			Mostly females, normal function
			Non atopic asthma
			Passive smoking exposure
			White race children, with higher BMI.
			<b>Adolescents and adults</b>
			White female, adult onset, non-aspirin sensitive with lower total IgE levels
			High atopy and atopic dermatitis.
			Mostly male patients
			Nonwhite race patients
			Aspirin sensitive asthma
Schimdlin et al 2015. (Abstract); USA	Children	4	Atopic, males with late onset wheezing.
			Non-atopic, males, early onset, and low function.
			Atopic asthma with persistent wheeze.
			Female, infrequent wheeze and normal function.
Sendín-Hernández et al 2018. Spain	Adults	3	Non atopic, familial atopy, mild, persistent asthma.
			Mild or intermittent asthma.
			Atopic, severe asthma that needs high treatment
Seino et al 2018. Japan	Adults	3	Elderly, severe, uncontrolled, high treatment and high adherence barrier.
			Elderly, normal weight, severe, uncontrolled asthma, with no adherence barrier.
			Young, obese, controlled, persistent asthma with no adherence barrier.
Seys et al 2017. Belgium	Adults	5	Sustained low function, eosinophilic neutrophilic sputum, and high IL-5-, IL-17/A/F and IL-25.
			High IL-5 and/or IL-10 and normal IL-17 F levels.
			High IL-6 profile.
			High IL-6 and IL-1 $\beta$ profile.
			Normal levels of all above cytokines.

(Continued)

**Table 5** (Continued).

Reference, Country	Population Age	Number of Phenotypes	Phenotype Description
Siroux V et al 2011. Multicountry	Adults	4	<b>EGEA</b>
			Inactive allergic childhood onset asthma
			Active allergic childhood onset asthma
			Inactive adult-onset asthma
			Active adult-onset asthma
			<b>ECRHS</b>
			Inactive asthma
			Active allergic asthma
			Severe asthma
			Adult-onset non-allergic asthma
Tay et al 2019. Singapore	Adults	3	Chinese, female, old, late onset obese, with controlled low symptom asthma.
			Non-Chinese, females, obese, uncontrolled, high symptoms, high GINA step and highest comorbidity.
			Male, multiethnic, atopic, early onset, smokers with moderate control level.
Tsukioka et al 2017. Japan	Adults	3	Athletes with moderate levels of FeNO and IgE.
			Female athletes, lowest FeNO and IgE levels, worst function despite low symptoms.
			Male athletes' childhood onset asthma, atopic, higher FeNO and IgE and higher response to methacholine.
Wang LL et al 2017. (abstract) China	Adults	5	Men, former or current smoker with high depression and anxiety, well controlled asthma with high lung function.
			Young, non-smoking women, atopic, with high psychological morbidity.
			Smoking, high HCU and sustained low function.
			Smoking men, high psychological morbidity and poor control, low SES with high HCU
			Women, non-smoking, non-allergic, slight obstruction, worst control and high anxiety and depression.
Watanabe et al 2016 (abstract). Japan	Adults	3	Older men with high morbidity of COPD and hypertension
			Middle age men with higher treatment adherence behavior.
			Women, middle age, low treatment levels.
Wisnivesky et al;2019 USA	Adults	3	Decreased lung function, poor control, high hospitalization, and high sinusitis and depression rates.
			Normal lung function, good control, moderate hospitalization rate, high depression, and post-traumatic stress disorders.
			Men, near normal lung function, good control, low rate of emergency hospitalization and medication use, low mental health diseases and GERD.
Wu et al 2017. China	Adults	3	Smoking, nonatopic, late onset, eosinophilic, short duration, NSAID sensitive, low function and impaired CT scan asthma with prior sinus surgery.
			Elderly, non-atopic, late onset, non-eosinophilic, long duration asthma, with prior sinus surgery.
			Male atopic asthma.

(Continued)

**Table 5** (Continued).

Reference, Country	Population Age	Number of Phenotypes	Phenotype Description
Zaihra T et al 2016. Canada	Adults	4	Severe asthma, late onset, high dose ICS, least cluster stability
			Females, obese, severe asthma with highest cluster stability.
			Mild air entrapment with proximal airway remodeling, early onset, reduced lung functions.
			Moderate asthma with good lung functions.
Zhang X et al 2019 (abstract). China	Adults	3	Moderate air entrapment with/without proximal airway remodeling
			Sever air entrapment with proximal airway remodeling
			Non atopic, infrequent symptoms and rhinitis asthma with normal function.
Zoratti E et al 2018. USA	Adults	4	Non-atopic, low inflammation, frequent symptoms and rhinitis and high treatment.
			Moderate atopy, mild symptoms and rhinitis, low medication, and minimal function impairment.
			Atopic, high inflammation, infrequent symptoms, and intermediate treatment.
			Atopic, frequent symptoms, high treatment, and highly impaired function.

## Severe Asthma

The total number of reported severe asthma phenotypes was 61 with considerable degree of overlap between them. The most reported features that differentiated severe asthma phenotypes were atopy, featuring in 28 phenotypes; age at disease onset, featuring in 25 phenotypes; treatment defined as medication dosage or treatment step; inflammation measures, featuring in 14 phenotypes; disease activity as frequency of symptoms and exacerbations, featuring in 14 phenotypes; and age and sex that featured 13 phenotypes.

Regarding allergic status and time at disease onset, 10 studies<sup>72,74–76,78–81,86,87</sup> reported phenotypes of atopic severe asthma, while non-atopic severe asthma phenotypes were reported by 8 studies.<sup>74–76,78–80,86,87</sup> Early onset severe asthma phenotypes were reported in 8 studies,<sup>70,72,74,75,78,79,86,87</sup> while late-onset severe asthma phenotype variants were reported in 7 studies.<sup>70,72,74,78,79,86,87</sup> Defining age of disease onset in most studies was based on measuring the mean and standard deviation of age at disease onset and comparing phenotypes.<sup>72,74,79,81,87</sup> Only one study defined more than 12 years as cutoff for late onset.<sup>78</sup>

Disease activity in terms of symptoms differentiated phenotypes of severe asthma with high symptoms presentation in 6 studies,<sup>70,74,78,79,81,82</sup> as well as in another 4 studies<sup>70,74,79,81</sup> with low symptoms. Based on medication usage, phenotypes of severe asthma that require extra higher treatment were described in 6 studies.<sup>69,74,76,78,79,81</sup> Those were in form of extra higher doses of ICS, oral corticosteroids (OCS), additional controller, regular use of systematic CS, or more frequent need for OCS and short controllers. In turn, lower to more moderate medication usage or requirement that was reported in 5 studies.<sup>69,74,76,79,81</sup> Although spirometry measures were not as commonly reported as other indicators of disease activity, highly obstructed variants of severe asthma phenotypes were reported in 7 reports,<sup>72,74,75,78,79,86,87</sup> while moderate to mild obstructed severe asthma in 5 reports.<sup>69,72,74,78,79</sup>

For demographic characteristics, variants of female severe asthma phenotypes were described in 4 studies,<sup>74,75,78,81</sup> and male severe asthma phenotypes in similar count of records.<sup>74,75,78,80</sup> Elderly related variants of severe asthma phenotypes were described in 4 studies,<sup>74,75,78,81</sup> and young age severe asthma phenotypes in 3 records.<sup>75,78,82</sup> Obesity-related variants of severe asthma phenotypes were reported in 5 studies.<sup>70,73,74,78,79</sup> See full results on number of derived phenotypes and their descriptions for studies on severe asthma in Table 6.

**Table 6** Number of Derived Phenotypes and Their Descriptions for Studies on Severe Asthma

Reference ID and Country	Number of Phenotypes	Phenotypes Description
Brinkman et al, 2011. (abstract). Multi-countries	3	Moderate disease with moderate use of OCS and low eosinophilia.
		Low eosinophilia and mild disease with low OCS usage.
		Obstructed eosinophilic with high OCS usage.
Desai et al 2011. (abstract). UK	4	Obesity with discordant high symptoms/ low eosinophilia.
		Late onset disease with concordant high symptom/eosinophilia.
		Early onset, discordant low symptoms/eosinophilia, low FEV1.
		Early onset concordant symptom and eosinophilia.
Diver et al 2018. UK	2	Microbial predominance with Hemophilus and Moraxella, high Gammaproteobacterial (G) to Firmicutes (F) ratio.
		Low microbial prevalence and low G:F ratio.
Fitzpatrick et al, 2018. USA	4	Late-onset, non-atopic, non- Hispanic white, normal function, and symptomatic asthma.
		Early-onset atopic asthma, lower lung function and high symptoms and medication usage.
		Early-onset atopic asthma with high comorbidities, high bronchial responsiveness, and low lung function.
		Early-onset atopic asthma with severe obstruction, highest symptoms and medication usage.
Freitas PD et al 2018 (abstract). Brazil	3	Physically inactive, obese, with depression and high treatment.
		Physically active, young, less comorbidities and good control and QOL.
		Intermediate physical activity.
Gomez et al, 2017. USA	4	Young females' mild atopic asthma with low YKL levels.
		Males with non-atopic asthma of good treatment response and low YKL level.
		Severe asthma, elderly with high treatment and high YKL levels.
		Obese elderly with severe short term disease and high YKL levels.
Jang et al, 2017; Korea	4	Mild, atopic, low rhinitis, eosinophilic asthma among young
		Non atopic neutrophilic severe asthma.
		Female highly reactive asthma with mixed inflammation
		Male smoking severely obstructed with high rhinitis.
Konstantellou et al 2015. Greece	3	Non atopic mild obstruction with low treatment and non-SRA criteria.
		Atopic, severe obstructed, with treatment and positive SRA criteria.
		Atopic, mild disease with low treatment and non-SRA criteria.
Lau et al, 2017: Singapore	2	High Absolute eosinophilic asthma, young, males and requires more hospital admission
		Low eosinophilic asthma

(Continued)

Table 6 (Continued).

Reference ID and Country	Number of Phenotypes	Phenotypes Description
Moore et al 2010; USA	4	Young females, childhood onset mild atopic asthma with low treatment and health care utilization.
		Old females' childhood onset mild disease with high treatment use.
		Elderly, late onset, obesity, non-atopic severe disease with high treatment usage and HCU.
		Early onset, atopic severe asthma, with treatment requirement and HCU.
Newby et al, 2018; UK	4	Atopic, early onset, low function, high HCU, high treatment, and high BDR.
		Late onset, obese, frequent symptoms, mild function decline, high treatment and high depression.
		Non atopic, normal function, infrequent symptoms and low treatment.
		Marked function decline, high treatment and infrequent exacerbations.
Raherson et al, 2018 (abstract); France	3	Atopic asthma.
		Male dominant eosinophilic asthma.
		Non atopic asthma.
Sekiya et al 2015; Japan	5	Early onset, severe symptoms, activity limitation, high treatment and high HCU.
		Elderly, female, long disease duration and high CRS and nasal polyps.
		Atopic low treatment and better hospitalization prognosis.
		Elderly, male, with COPD.
		Mild disease with previous hospitalization.
Simpson et al; 2017. Multi-country	6	Classic asthma: high airway reversibility and high eosinophils.
		Pulmonary treatable traits asthma (high prevalence of multiple pulmonary traits.
		Steroid insensitive: high eosinophils despite good medication adherence.
		Reflux and cough: cluster highly identified by traits of reflux and cough.
		High treatable traits: high prevalence of treatable traits across subjects.
		Low prevalence of treatable traits across subjects.
Taniguchi et al; 2014; (abstract) Japan	5	Young onset atopic severe asthma.
		Older onset, female, obesity severe asthma
		High smoking pack years and high eosinophilia.
		High smoking pack years, low DLCO and low eosinophils
		Non indicated
Serrano Pariente et al 2015. Spain	3	Elderly, with severe asthma.
		Respiratory arrest, impaired consciousness and need for mechanical ventilation
		Young, no sufficient treatment and sensitization to alternata and soybean.

(Continued)

**Table 6** (Continued).

Reference ID and Country	Number of Phenotypes	Phenotypes Description
Wu et al; 2014. USA	6	Healthy controls with normal lung function and no symptoms.
		Mild asthma, less symptoms, better quality of life, early onset, less atopic, more allergen skin test reaction, better clinical outcome, and high BAL eosinophile and neutrophile than cluster 1.
		Hispanic women, frequent symptoms, low QOL, high allergic sensitization, low inflammation, and near normal FEV1 value with low hospitalization.
		Female, non- Caucasian, high BMI, high symptoms, and early onset asthma; high familial asthma, low lung function high BHR and high inflammation.
		Elderly, late onset asthma, non-allergic, high nasal polyposis and sinusitis and high CS dosage; low function, high eosinophils and neutrophils.
		Early onset, high symptoms, lowest lung function, high HCU, high sinusitis, high OCS, high FeNO, eosinophils and neutrophils and high osteoporosis.
Weng-jing Ye et al 2017. China	4	Early onset atopic asthma.
		Small airway obstruction and atopic asthma.
		Late-onset and non-atopic asthma.
		Severe airflow obstruction and obvious airway remodeling.
Youroukova et al 2017. Bulgaria	4	Late-onset, non-atopic asthma with impaired lung function.
		Late-onset, atopic asthma.
		Late-onset, aspirin sensitivity, eosinophilic asthma.
		Early-onset, atopic asthma.

## COPD

The total number of reported COPD phenotypes was 57. The most reported feature for defining COPD phenotypes was lung function measured by spirometry that differentiated 44 phenotypes. Other commonly reported features were age, featuring in 26 phenotypes; symptoms and frequency of exacerbations, featuring in 24 phenotypes; sex, featuring in 17 phenotypes; and cardiovascular, metabolic, and psychiatric comorbidities, featuring in 14–17 phenotypes.

Based on spirometry lung function measures, COPD phenotypes were classified as mild, moderate, or severely obstructed disease. Severe to moderately obstructed phenotypes of COPD were reported in 10 studies,<sup>4,88,91,93,95–97,102,105,109,111,112</sup> while 5 studies<sup>91,93,96,102,112</sup> reported mild obstructed COPD phenotypes. Other measures of lung function used for deriving COPD phenotypes were measures of accompanying emphysematous changes like lung diffusion capacity for carbon monoxide (DLCO),<sup>88</sup> computed tomography (CT) measure of lung density and airway wall thickness.<sup>4,90,97,105</sup> The latter identified COPD phenotypes with high, moderate to low emphysematous changes in 4 studies.<sup>4,90,97,105</sup>

Demographic and social characteristics like age, sex, body mass index and smoking were also used to define COPD phenotypes. Elderly related COPD phenotype was reported in 9 COPD studies,<sup>88,91–93,95,96,98,102,113</sup> while 7 studies<sup>88,91,93,95,96,98,113</sup> described COPD phenotypes that were characterized by young age. Variants of female-related COPD phenotypes were reported in 3 records,<sup>88,91,97</sup> while male sex-related COPD was reported in 4 studies.<sup>88,91,96,102</sup> Both over- and underweight were associated with COPD phenotypes when considering BMI. Obesity-related COPD was reported in two studies by Burgel et al<sup>4,91</sup> while under or low weight-related COPD phenotypes were reported in 5 studies.<sup>4,93,95,111,113</sup> Heavy, persistent, high rate or long duration smoking-related COPD phenotypes were reported in 3 studies,<sup>91,97,105</sup> while 2 studies reported low smoking-related COPD phenotypes.<sup>91,102</sup>

Disease activity/severity was characterized in studies of COPD phenotyping variously using frequency of symptoms and exacerbations and level of treatment. COPD phenotypes with high frequency of symptoms and exacerbations were reported in 11 studies of COPD,<sup>4,91,93,95,97,100,102,107,109,111,112</sup> while COPD phenotypes with low symptoms in 8 studies.<sup>4,91,93,97,100,102,107,112</sup> Four studies<sup>91,93,97,105</sup> reported on COPD phenotypes with utilization of high treatment doses and 2 others with low dosage treatment.<sup>91,97</sup>

**Table 7** Number of Derived Phenotypes and Their Descriptions for Studies on COPD and Asthma COPD Overlap (ACO)

Study ID; Year	Number of Phenotypes	Phenotypes Description
Augustin et al; 2018. Netherland	7	Male, mild obstruction with mildly impaired diffusion.
		Elderly, males, moderate obstruction with moderately impaired diffusion.
		Sex non differential, moderate to severe obstruction and diffusion impairment.
		Sex non differential, moderate to severe obstruction with mild diffusion impairment.
		Elderly, males, severe obstruction, severe hyperinflation, moderate diffusion impairment and respiratory muscle weakness.
		Elderly, female, severe obstruction, sever hyperinflation and moderate diffusion impairment.
		Elderly, males, severe obstruction, severe hyperinflation, sever impaired diffusion, respiratory muscle weakness and alveolar hypoventilation.
Bafadhel et al; 2011 UK	4	Bacteria-predominant.
		Eosinophil-predominant.
		Virus- predominant.
		Pauci-inflammatory reaction. Clusters further varied by Sputum IL-1b, serum CXCL10 as biomarker.
Bertini et al; 2013. Italy	3	High Formate, Serine, Valine, Lysine, Acetate, Alanine, Isoleucine and Leucine.
		High radiologic emphysema.
		Low Proline.
Burgel et al; 2017 France	5	Elderly, high mortality, high CVD and diabetes and less severe disease.
		Intermediate mortality, low comorbidities and severe to moderate disease.
		Elderly, obese, intermediate mortality and high comorbidities.
		High mortality, severe disease with low comorbidities.
		Low mortality, mild disease, and low comorbidities.
Burgel et al; 2010 France	4	Young, severe airflow limitation (GOLD stage 3 and 4), low BMI, frequent exacerbations, high levels of anxiety and depression; cardiovascular comorbidities were infrequent.
		Older individuals, mild airflow limitation, low dyspnea, mild overweight, low anxiety and depression levels, almost no exacerbations, and mild impairment in HRQoL; higher prevalence of cardiovascular morbidity.
		Moderate to severe obstruction, young with low prevalence of cardiovascular and depression comorbidity.
		Moderate to severe obstruction, older with high prevalence of cardiovascular and depression comorbidity.

(Continued)

**Table 7** (Continued).

Study ID; Year	Number of Phenotypes	Phenotypes Description
Burgel et al; 2012 Belgium	3	Mild to moderate obstruction, mild emphysema, mild dyspnea, normal nutrition status and low comorbidities.
		Younge, males, underweight, Severe obstruction, marked emphysema, sever dyspnea, impaired QOL, high musculoskeletal disease
		Olde, obese, moderate to severe obstruction with bronchial thickening, mild emphysema than C2, high rates of CVD and diabetes.
Chen et al; 2014. Taiwan and China	5	Young, mild obstruction, mild symptoms, and infrequent exacerbations.
		Elderly, mild obstruction, mild and infrequent symptoms, and high CIS usage.
		Elderly, underweight, moderate obstruction with sever exacerbation and dyspnea symptoms.
		Severe obstruction, high symptoms, low BOS and mild exacerbations.
		Severe obstruction, low BOS, severe frequent exacerbations, high ICS and high mortality.
Chubachi et al (abstract); 2016 Japan	5	Low comorbidity
		Lung and other cancers.
		Metabolic and CVD.
		Psychological and GERD.
		Cachectic, anemia and Osteoporosis.
De Torres et al; 2017. Spain	3	Young, mild obstruction, low BMI, and low CVD.
		Characteristics between C1 and C3.
		Elderly, high BMI, marked obstruction, high symptoms and high HTN, diabetes, OSA and CVD.
Divo et al abstract; 2016. Not indicated	4	Young, male, obese, moderate impairment on QOL, physical ability and health, moderate diffusion capacity
		Elder, obese, sex undifferentiated, severe obstruction, moderate diffusion capacity with marked health, physical impairment, and high mortality.
		Elderly, female, obese, mild obstruction, high diffusion capacity, least health impairment and low mortality.
		Reference: Young, female, no obstruction, high diffusion capacity, normal to overweight, with the least overall impairment and mortality.
Fens et al; 2013. Netherlands	4	Females, mild obstruction, low symptoms, good QOL with high lung density and little emphysema.
		Combined CB and emphysema, moderate to severe obstruction, GOLD stage 3, impaired diffusion capacity and emphysema.
		Mild obstruction, GOLD stage I, Hypercholesterolemia and low lung density and high emphysema.
		Smoking, high symptoms, preserved function, low QOL, with moderate emphysema and lung density.

(Continued)

**Table 7** (Continued).

Study ID; Year	Number of Phenotypes	Phenotypes Description
Guillamet RV et al; 2018. USA	9	Elderly, high depression and mild comorbidities.
		Low comorbidities and low remission rates.
		Elderly, with high CHD and CHF.
		Young, low comorbidities, high medication, and readmission rates.
		Advanced disease and frequent readmissions.
		Young, with high CVA.
		Young, atopic asthma with high readmission rates.
		Young, high CKD and diabetes with few readmissions.
		Advanced disease with frequent readmissions.
Harrison SL et al; 2014 UK	3	Controlled, infrequent symptoms, low emotional sensitivity, and short illness duration.
		Uncontrolled, frequent symptoms, high emotional sensitivity, and short illness duration.
		Coherent illness, frequent symptoms, emotional sensitivity, and cyclical illness timeline.
Haghighi et al, 2019. USA	4	Asymptomatic, normal airway structure, normal lung function and airway wall thickening and moderate emphysema.
		Obese, female, increase tissue fraction at inspiration and minimal emphysema and lowest emphysema progression rate.
		Elderly, male, small airway narrowing, decreased tissue fraction at expiration and high air entrapment.
		Lean, male, severe COPD and high emphysema progression rate.
Kim WJ et al; 2017. Asia	3	Elderly, male, low function, and few symptoms,
		Very Low function, high symptoms of dyspnea and health impairment.
		Nonsmoking, obese, mild disease and normal function.
Kim S et al 2017. Korea	3	Young, mild, low symptoms, low disease impairment on life and low inflammation.
		Male, old, heavy smokers, severe obstruction, underweight, high inflammation, high symptoms and disease impairment of daily life, high emphysematous changes.
		Females, non-smokers, moderate dyspnea and disease impairment on life.
Kukol et al 2019; (abstract). Russia	5	Normal body weight, with disease duration more than 5 years; with a frequency of exacerbations less than 2 times a year among women
		Elderly, younger than cluster 1, overweight, disease duration <5years and exacerbation less than twice a time among women.
		Overweight, less than 6 years of disease duration and exacerbation less than twice a year among men.
		Under weight, more 7 years of disease duration, and exacerbation less than twice yearly among men.
		Overweight more than 8 years disease duration and less than 2 exacerbations yearly.

(Continued)

**Table 7** (Continued).

Study ID; Year	Number of Phenotypes	Phenotypes Description
Lee et al, 2019, Korea	4	ACO, second best lung function to mild COPD, second highest age, highest BMI, least smoking, longest walking distance and lowest CAT score and highest rate of asthma.
		Mild COPD
		Moderate COPD
		Severe COPD
Li et al, 2016, USA	5	Resistant smokers with normal function and early emphysema.
		Resistant smokers with mild function decline and no emphysema.
		Heavy smokers, mild COPD minimal obstruction and emphysema.
		Less smokers, moderate COPD high HCU and ICS usage.
		Sever COPD, sever obstruction, sever emphysema and high ICS and HCU usage.
Liang et al, 2019: Korea	4	Seven trajectories of association between clinical and autoantigens parameters, retrospective exacerbation (AE) increased with CAT score.
		Five trajectories of association between clinical and immunological parameters, retrospective exacerbation associated negatively with age, lung function and sputum antibodies (P0, Scl70, Sm, UI-SnRNP, PR3 and Ro/SSA) and serum globulin (Glb) and positively related to white blood cells, sputum anti-PR3, sputum anti-Ro/SSA, and sputum anti-UI-SnRNP were significantly negatively correlated with AE
		Five trajectories of association between clinical and immunological parameters, AE was positively associated with the CAT score and sputum autoantibodies (UI-SnRNP, PR3, MPO and Ro/SSA).
		Six trajectories of association between clinical and immunological parameters, AE was negatively associated with serum uric acid and blood neutrophil count.
Lopes et al; 2019. Brazil	2	Frequent symptoms, high consequences, and cyclical timeline of illness, less coherence and high emotional sensitivity. (destressed)
		Infrequent symptoms, less cyclical timeline of illness, low emotional sensitivity, and high illness coherence.(coping)
Ning et al; 2016, (abstract). China	4	Chronic bronchitis in smokers with normal pulmonary function.
		Chronic bronchitis or mild chronic obstructive pulmonary disease (COPD) patients with mild airflow limitation.
		Heavy smoking, poor quality of life and severe airflow limitation.
		Atopic patients with mild airflow limitation, elevated serum IgE and clinical features of asthma
Peters et al; 2016 Netherlands	3	Moderate COPD, normal weight, high exercise performance and mild impact on quality of life, functionality, and symptoms.
		Moderate COPD, overweight, high exercise performance and high impact on symptoms, functionality, and quality of life.
		Severe COPD, overweight, moderate exercise performance and mild impact on symptoms, quality of life and functionality.

(Continued)

**Table 7** (Continued).

Study ID; Year	Number of Phenotypes	Phenotypes Description
Pikoula et al; 2019. UK	5	Young, female, smoking, high psychological illness, highly deprived group with high IMD.
		Male, late diagnosis, severe airflow obstruction, frailty and the lowest comorbidities.
		Elderly, male, former smokers, high IHD, CVD and diabetes.
		Under weight, females, eosinophilic, severe disease with low CVD high obesity and atopy.
		Females, smoking, obese, atopic mild asthma high CRS and GERD.
Rodrigues et al; 2018. Brazil	2	Under to normal weight, severe disease, low function, impaired daily activity, and low muscle strength, compared to cluster 2.
Scarlata et al; 2018 Italy	3	Mild central and peripheral obstruction and high KCO.
		Intermediate peripheral obstruction with no comorbidities
		Sever obstruction, low KCO and short estimated survival.
Xavier F et al; 2019. Brazil	3	Young, physically active, low obstruction, low dyspnea, and good body composition.
		Elderly, physical inactive with low mortality.
		Physical inactive, poor QOL, and low mortality
Yoon et al; 2019. Korea	4	Asthma and COPD overlap
		Mild COPD
		Moderate COPD
		Severe COPD
<b>Asthma COPD overlap</b>		
Rootmensen et al 2016. Netherlands	4	Cluster 1: excessive smoking history COPD without sign of emphysema
		Cluster 2: emphysematous COPD
		Cluster3: patients with allergic asthma characteristics
		Cluster4 features suggesting asthma and COPD overlap
De Vries et al 2018. Netherlands	5	Cluster 1: female, obese, high symptoms, combined asthma and COPD, and low inflammation.
		Cluster 2: males, combined asthma, and COPD, eosinophilic, high FENO and low treatment step.
		Cluster3: non-Caucasian, combined asthma, and COPD, highly obstructed, eosinophilia, low exacerbations, and low treatment steps.
		Cluster4: atopic, combined asthma and COPD, neutrophilia, and high exacerbations,
		Cluster5: mostly asthmatics, high lung function, and low exacerbations.

(Continued)

**Table 7** (Continued).

Study ID; Year	Number of Phenotypes	Phenotypes Description
Fingleton et al 2017. New eland and China	5	Cluster1: smokers, atopic, severe, late onset, asthma/ chronic bronchitis/ emphysema overlaps with systemic inflammation.
		Cluster2: smokers, moderately severe, early onset asthma combined asthma and COPD, type two dominant inflammation.
		Cluster3: minimal smoking, atopic, eosinophilia and type 2 inflammation with minimal airflow obstruction.
		Cluster4: late onset non atopic with minimal airflow obstruction.
		Cluster 5: atopic, early onset and mild/ intermittent phenotype.
Gorska K t al.2017; Poland	12	Cluster1: atopic asthmatics.
		Cluster2: non atopic asthmatics,
		Cluster 3: smokers with highly obstructed COPD
		Cluster 3:1: smokers, highly obstructed COPD with high eosinophilia.
		Cluster 3:2: smokers, highly obstructed COPD with low eosinophilia.
		Cluster 1: asthmatics
		Cluster 2: combined asthma and COPD that has further 5 subgroups:
		2:1: mainly COPD with high MM6.
		2:2: smokers, mainly COPD, highly obstructed,
		2:3: mainly combined asthma and COPD
		2:4: atopic, asthma with high eosinophilia.
		2:5: low smoking, mainly COPD, with minimal obstruction.

Concerning comorbidities, the mostly reported ones to differentiate COPD phenotypes were cardiovascular diseases and diabetes and metabolic diseases,<sup>4,92,94,95,98</sup> together with depression and anxiety.<sup>91,94,98,100,107</sup> Additionally, features considered in characterizing COPD phenotypes were disease impairment on physical and daily activity, respiratory health, quality of life and mortality. COPD phenotypes with impaired quality of life were reported in 3 studies,<sup>97,109,113</sup> while high mortality-related COPD phenotypes were reported in 5 studies.<sup>92,93,96,112,113</sup> See full results on number of derived phenotypes and their descriptions for studies on COPD in [Table 7](#).

### Asthma and COPD Overlap (ACO)

A total of 21 phenotypes of ACO were identified. The most reported features considered for differentiating ACO phenotypes were smoking status, which identified 7 phenotypes; inflammation status which identified 9 phenotypes; atopy that identified 7 phenotypes; spirometry measures identifying 5 phenotypes and disease activity/severity as per symptoms identifying 5 phenotypes.

Regarding socio-demographic aspect, smoking-related ACO phenotypes were reported in two studies,<sup>116,118</sup> along with female ACO phenotype and obesity-related ACO,<sup>115</sup> each one record. Lung function measures featured a highly obstructed ACO phenotype that was reported in two studies<sup>115,116</sup> and a high symptom phenotype of ACO was reported in 1 study.<sup>115</sup> With respect to inflammation status, eosinophilic variants of ACO were reported in one study,<sup>115</sup> as well as neutrophilic ACO phenotype.<sup>115</sup>

For other disease characteristics, early onset ACO phenotypes were reported in one study,<sup>118</sup> while 3 records reported a variant of atopic ACO phenotype.<sup>115,116,118</sup> See full results on number of derived phenotypes and their descriptions for studies on COPD and ACO in Table 7.

## Rhinitis

The total number of reported rhinitis phenotypes was 45. The most considered features for differentiating phenotypes of rhinitis were sex, which featured in 19 phenotypes; disease severity, which featured in 18 phenotypes; impairment on quality of life, which featured in 14 phenotypes and disease activity per symptoms that featured 10 phenotypes.

Considering socio-demographic characteristics, sex, age, and socio-economic status (SES) identified several rhinitis phenotypes. Variants of female-related rhinitis as well as male-related phenotypes of rhinitis were reported in near half of the reports (n = 5).<sup>119–121,124,126</sup> Phenotypes of old age-related rhinitis were reported in 2 studies,<sup>121,126</sup> as well as young age-related ones.<sup>126</sup> SES featured phenotypes of high and low SES-related rhinitis which was reported by Lee et al.<sup>125</sup> Alcohol intake further identified high intake-related phenotypes of rhinitis that was reported by Soler et al.<sup>126</sup>

Disease activity in terms of frequency of symptoms, classification of disease based on severity status as well as medication intake were commonly used to differentiate rhinitis phenotypes. High symptom phenotypes of rhinitis were reported in 3 studies.<sup>121,123,125</sup> Severe rhinitis phenotypes, in turn, were reported in 3 studies,<sup>120,121,123</sup> while Lee et al.<sup>125</sup> reported rhinitis phenotypes which require high treatment doses.

Measures of airways or lung function that were used to feature rhinitis included CT scanning diagnostics of rhinitis, endoscopy score, and FeNO, in addition to spirometry, bronchodilator reversibility, bronchial hyperresponsiveness in subjects with accompanying asthma.<sup>124,125</sup> Two records described rhinitis phenotypes with highly obstructed airways.<sup>124,125</sup> Rhinitis with high endoscopic and CT score were reported by two studies.<sup>119,126</sup> One study reported on rhinitis phenotypes among asthmatics that is characterized by high inflammation indicated by FeNO.<sup>124</sup> Among asthmatic with rhinitis, phenotypes of rhinitis with low to moderate BDR and BHR were also reported.<sup>124,125</sup>

**Table 8** Number of Derived Phenotypes and Their Descriptions for Studies on Rhinitis

Refence ID and Country	Number of Phenotypes	Phenotypes Description
Adnane et al 2017. Morocco	3	Female, eosinophilic, high CRS with nasal polyps, low endoscopy score, low CT and sino-nasal outcome test
		Male, non-eosinophilic rhinitis.
		Female, high CRSsNP, high endoscopy score, CT and sino-nasal outcome score.
Agache et al 2010. France	5	Short breast-feeding duration and severe rhinitis- children.
		Male, polysensitization and severe rhinitis - children.
		Severe rhinitis with polysensitization – adults.
		Male, high pets' exposure and severe rhinitis- adults.
		High atopy and polysensitization- adults.
Bousquet et al 2015. France	4	Moderate to severe rhinitis, low QOL, high symptoms, high disease burden than C3.
		Elderly, Female, early onset, mild intermittent rhinitis, high QOL, low symptoms and low comorbidities.
		Males, moderate to severe rhinitis, less symptoms and better QOL than C1.
		Young, female, atopic, later onset, severe to moderate persistent rhinitis, high symptoms, low QOL and high disease burden.

(Continued)

**Table 8** (Continued).

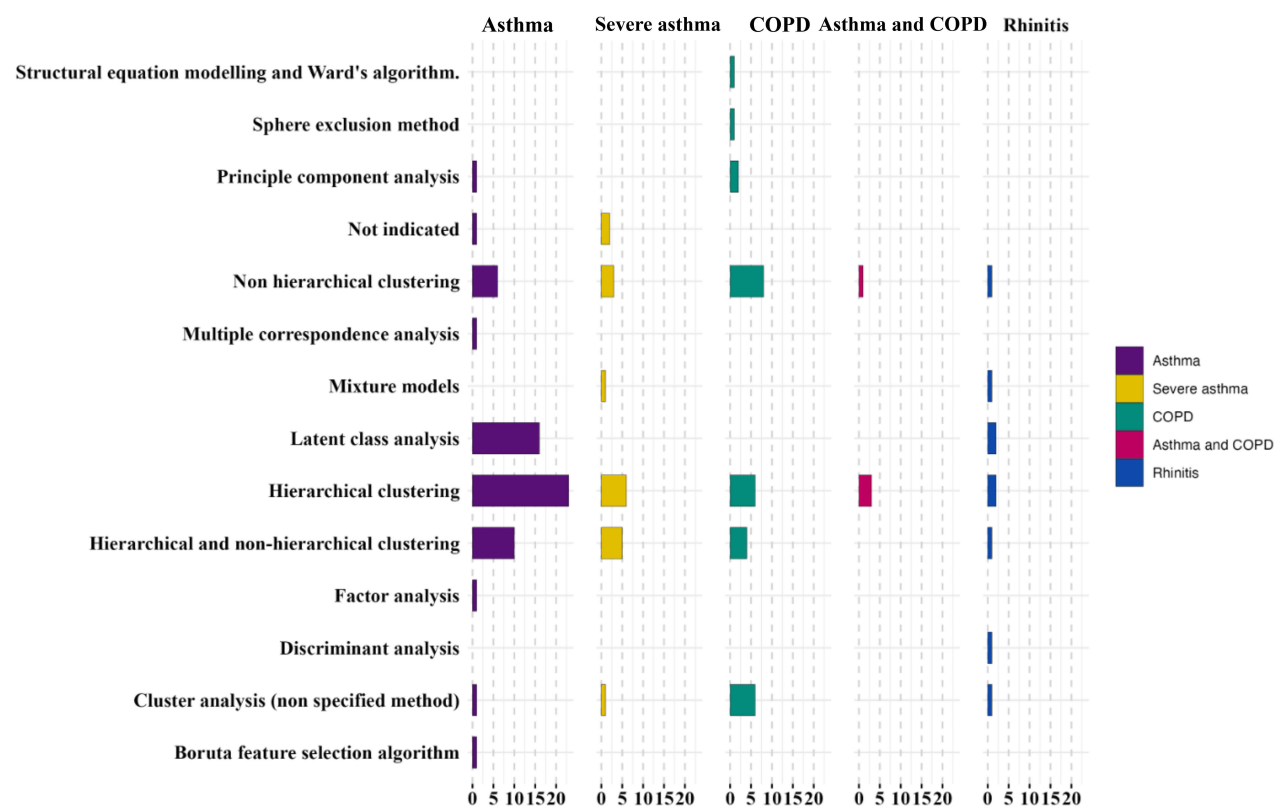
Refence ID and Country	Number of Phenotypes	Phenotypes Description
Burte et al 2015. France	3	No nasal symptoms,
		High nasal symptoms through year; high sinusitis, and low sensitization.
		Nasal symptoms at spring, high sensitization, high hay fever, high allergic rhinitis and conjunctivitis and high polysensitization.
Herr M et al 2012. France	3	Highly atopic, severe occasional wheeze, low HCU, high sensitivity to cow milk, egg white, nuts, cat and house mites' dust.
		Mild atopic, severe wheeze, impaired activity with high nocturnal cough and respiratory infections.
		Reference: mild occasional wheeze, with low respiratory and allergic outcomes.
Kurukulaaratchy et al 2015. UK	4	Atopic, early onset, moderate severity, normal function, low BHR, BDR and prevalence of asthma.
		Female, non-atopic, late onset, mild severity, normal function with low BHR and BDR, low asthma and low inflammation
		Severe earliest-onset rhinitis with asthma, the youngest rhinitis onset, the highest comorbid asthma (of simultaneous onset) and atopy.
		Male, atopic, early onset, seasonal disease of mild severity, low function with moderate BHR and BDR, moderate asthma and high inflammation.
Lee E.L. et al 2016. Korea	4	Non atopic rhinitis with low SES.
		Atopic rhinitis with normal lung function.
		Atopic, low function, high food allergy, high atopic dermatitis, high symptoms with high treatment need,
		Non atopic rhinitis with high SES.
T.A Nakayama et al 2012. Japan	4	Young, low peripheral eosinophil, basophil and mucosal eosinophil count, low CT score and polyps score and low symptoms.
		Low peripheral eosinophil, basophil and mucosal eosinophil count, higher CT and polyp score than cluster 1, and high symptoms score.
		Highest mucosal eosinophil count and peripheral basophil and eosinophil counts and low polyp and symptoms score.
		High CT and polyps score, the highest incidence of ATA as comorbidity, high peripheral eosinophil, basophil and mucosal eosinophil count and high symptoms.
Soler et al 2015. USA and Canada	5	Elderly, males, high alcohol intake, high diabetes, low depression, moderate severity with low QOL.
		Elderly, male, high depression, moderate QOL and low endoscopy and control score.
		Female, high depression, low endoscopy score and intermediate control score.
		Young, males, high depression, mild disease and low QOL.
		Females, high depression and fibromyalgia, low control, and endoscopy score and low QOL.

Based on disease characteristics like time of onset and seasonality, atopy status and accompanying nasal polyposis, both early and late onset variants of rhinitis phenotypes were reported in two studies,<sup>121,124</sup> while seasonal rhinitis and accompanying nasal polyposis phenotypes were reported by in the same count of studies.<sup>122,124</sup> Variants of atopic rhinitis were reported in 3 studies,<sup>123–125</sup> while polysensitization rhinitis phenotypes were also reported in 3 studies.<sup>120,122,123</sup>

The aspect of disease impairment on QOL and comorbidities was also frequently considered in featuring phenotypes of rhinitis. Phenotypes of rhinitis with impaired QOL were reported in 3 studies.<sup>121,126,127</sup> Rhinitis phenotypes with related comorbidities of depression, fibromyalgia, diabetes, and dermatitis were reported in one study by Soler et al.<sup>126</sup> See full results on number of derived phenotypes and their descriptions for studies on rhinitis in Table 8.

## Methods of Phenotyping

Various methods of unsupervised computational phenotyping of respiratory diseases were used across the reported studies (Figure 2). The most frequently implemented and reported unsupervised approaches for phenotyping of chronic airway diseases were hierarchical and non-hierarchical clustering<sup>10,32,38,40,46,54,70,81,82,90,95,101,110–112,114,117</sup> with some records (n = 19) reported the implementation of the two approaches in the same study.<sup>17,19,21,24,35,38,45,47,62,75,76,79,80,86,93,100,107,113</sup> In addition, latent class modelling<sup>10,13–16,20,23,31,41,42,44,48,50,58–60,123,125</sup> was also frequently used. Other non-model-based methods of dimensionality reduction, such as factor analysis, principal component analysis, discriminant analysis and multiple correspondence analysis were also reported as methods for deriving phenotypes, albeit less frequently. Over years, hierarchical and non-hierarchical clustering were common particularly between 2010 and 2018. However, between 2015 and later, there was an increase in the use of other methods such as mixture-based model,<sup>69,126</sup> structural equation modelling,<sup>88</sup> and factor analysis with latent class modeling.<sup>128</sup> Figure 2 shows the count of studies reporting the methods applied for phenotyping.



**Figure 2** Number of studies using each unsupervised phenotyping method for each respiratory outcome.

# Quality Assessment of the Included Studies

Overall, the comprehensiveness of variables included in deriving the phenotypes was the best quality aspect reported in majority of studies on asthma,<sup>10,11,14–16,18,20,21,24,28,30,31,33,34,41,43–45,47,48,50,56,59–61,65,66</sup> COPD,<sup>4,88,91–93,95,96,101,102,107,109–111,114</sup> severe asthma,<sup>72,78,82,86,87</sup> ACO<sup>116–118</sup> and rhinitis.<sup>120,122,124–126</sup> Random sampling of study subjects, however, was less frequently performed among studies on asthma (n = 8)<sup>23,32,41,46,51,59–61</sup>, severe asthma (n = 2),<sup>78,81</sup> COPD (n = 1),<sup>105</sup> ACO (n = 2)<sup>117,118</sup> and rhinitis (n = 1).<sup>125</sup> Majority of studies excluded subjects based on either clinical, social, or demographic characteristics. With respect to method of outcome definition, the most reported approach was usage of physician diagnosis assisted by clinical and biomarkers, which was reported in 33 of studies on asthma,<sup>10,12,13,15,16,18,21,24,27,30,31,33–38,40,42,44,46,48,52,54–63</sup> 16 of studies on COPD<sup>4,88–91,93–95,97,101,102,109–112,114</sup>, 12 of studies on severe asthma,<sup>69,71–73,75,76,78,80–82,86,87</sup> four of studies on rhinitis,<sup>119,120,126,127</sup> two of studies on ACO.<sup>116,117</sup> Overall reporting on reproducibility practices was uncommon, including how investigators handled noise and variation in data, rationale for selecting statistical methods for phenotyping, visualization techniques, and utilization of available tools for implementing reproducibility. With respect to clinical, biological or scientific relevance of the derived phenotypes, most studies reported on this aspect: 26 studies on asthma,<sup>10–24,27,28,30–37</sup> 18 studies on COPD,<sup>4,88,91–98,100–102,107,109,110,112,113</sup> 11 studies on severe asthma,<sup>69–73,75,76,78,81,87</sup> three studies<sup>115,116,118</sup> on ACO and nine studies on rhinitis.<sup>119–127</sup> Full information on the quality assessment results can be found in Figure 3 and Table S2 in Supplementary Material.

# Endotypes and Phenotypes of Airways Diseases

Efforts to define phenotypic subtypes of airway diseases involved utilizing various approaches, including the assessment of serum and sputum-induced inflammatory cells as well as other biomarkers associated with inflammation and related

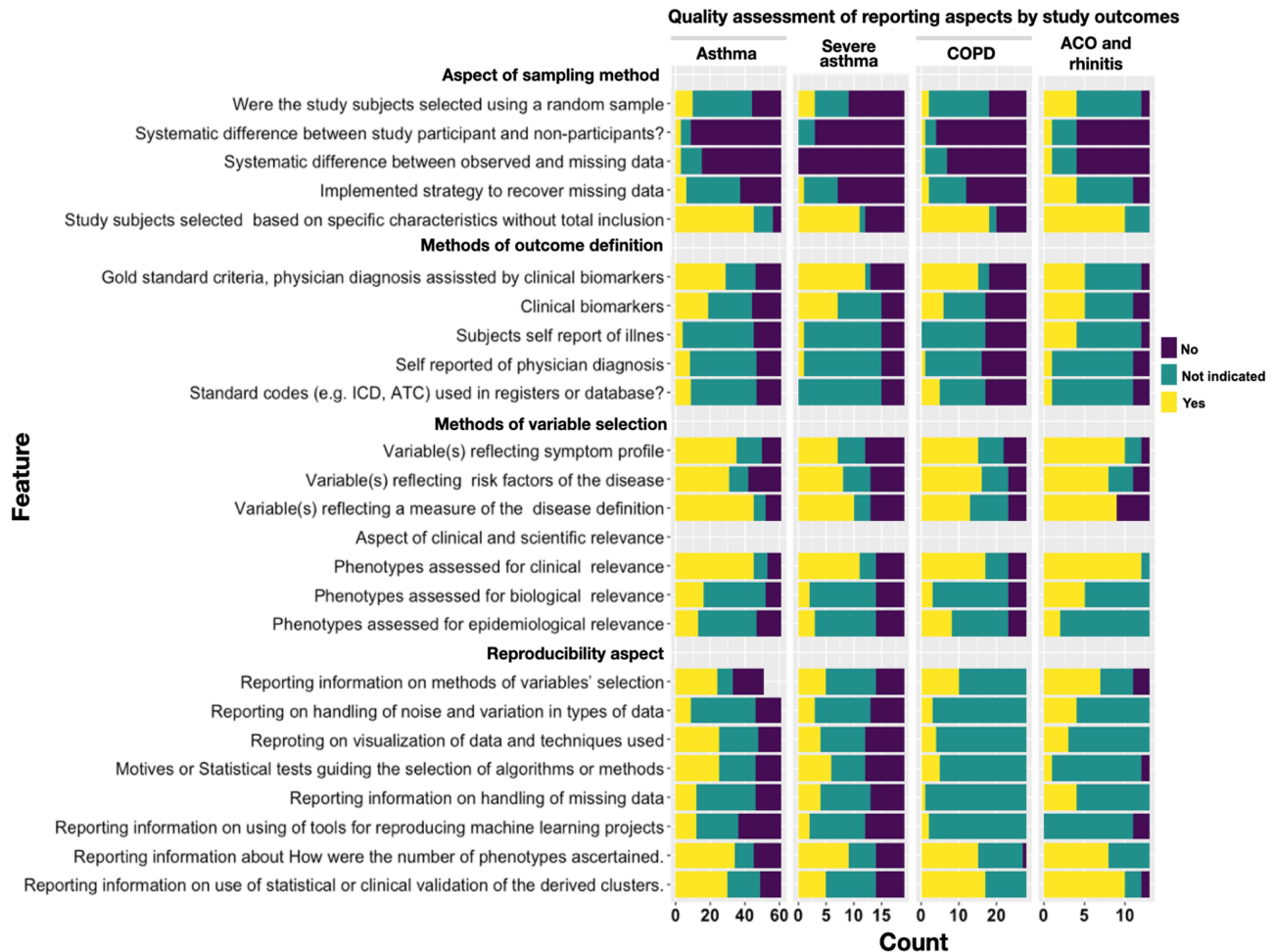


Figure 3 Quality assessment items reporting for studies on asthma, severe asthma, COPD, asthma and COPD and rhinitis.

processes. These biomarkers include cytokines, airway-inducible inflammatory mediators, and the composition of the airway microbiome.

Mastalerz et al<sup>48</sup> and Liang et al<sup>42</sup> conducted studies investigating the role of airway-induced pro- and anti-inflammatory lipid eicosanoid mediators in asthma. In their research, Mastalerz et al<sup>48</sup> identified three distinct asthma phenotypes based on high levels of anti-inflammatory mediators. One phenotype exhibited chronic rhinosinusitis (CRS), good control, and mixed inflammation. The second phenotype showed atopy, no CRS, good control, and mixed inflammation. The third phenotype had poor control, aspirin sensitivity, and eosinophilia. Additionally, Mastalerz et al<sup>48</sup> identified a phenotype characterized by high levels of anti-inflammatory mediators among obese women with early-onset, atopic, and severe asthma.

Liang et al<sup>42</sup> derived a phenotype that is characterized by a mixture of pro- and anti-inflammatory mediators. This phenotype exhibited low basophils, high functional activity, and poor control. Liang et al<sup>42</sup> also identified a phenotype with low basophils, high anti-inflammatory mediators, and low control. Similarly, Caljwaska et al<sup>16</sup> utilized induced sputum supernatants for phenotyping of non-steroidal anti-inflammatory drug-exacerbated respiratory disease (NERD). Seys et al<sup>58</sup> focused on the expression of inflammatory cytokines in airway sections to further subtype asthma patients. Their work reported an unexpected pattern of cytokine predominance among Th2-high asthmatics. It, further, identified subclusters with high levels of IL-5, IL-10, IL-25, and IL-17, associated with low lung function, high eosinophils, neutrophils, and fractional exhaled nitric oxide (FeNO). Another cluster consisted of IL-5 and/or IL-10 high asthmatics, while a separate cluster showed high levels of IL-6. A fifth cluster exhibited a normal pattern of high Th2 cytokines but with high eosinophils and low neutrophils. These findings indicate further heterogeneity among individuals with Th2-high asthma. However, this phenotyping approach did not include other clinical parameters such as symptoms, lung function, and outcomes. Instead, the derived clusters were modeled against clinical outcomes for further evaluation.

Nagasaki et al<sup>50</sup> identified phenotypes characterized by high serum periostin. One phenotype exhibited high eosinophilia, early-onset disease with good control, while the other phenotype showed high periostin, mixed inflammation, severe disease with poor control, high IL-6, mixed inflammation, and multimorbidity.

In the context of COPD, Bafadhel et al<sup>89</sup> identified five distinct phenotypes of COPD exacerbations based on biological biomarkers. These phenotypes were characterized by different predominant factors, including bacteria, viruses, eosinophils, paucigranulocytemia, and elevated levels of Sputum IL-1b and serum CXCL10. Britani et al<sup>90</sup> employed exhaled breath condensate analysis to assess inflammatory biomarkers in COPD, revealing unique metabolite profiles such as the low proline phenotype and high serine, valine, lysine, acetate, alanine, isoleucine, and other metabolite phenotypes.

In severe asthma, Diver et al<sup>71</sup> focused on airway microbiology and identified clusters of severe asthma with varying levels of *Haemophilus* and *Moraxella* sputum communities, as well as different ratios of Gammaproteobacteria (G) to Firmicutes (F). Gomez and colleagues<sup>74</sup> explored severe asthma phenotypes characterized by varying levels of the chitinase-like protein YKL-40 as an inflammatory mediator which was inversely associated with disease severity, control, and treatment response.

An endotype of high eosinophilic chronic rhinosinusitis with high nasal polyposis and low CT and endoscopy score was identified by Adnane et al,<sup>119</sup> while Nakayama et al<sup>127</sup> reported a high eosinophil and basophil rhinitis that is characterized by comorbid asthma, high symptoms, high CT and endoscopy score. Low eosinophilic rhinitis reported by Adnane et al<sup>119</sup> was variants among male subjects, and others among females but with high CRSnP and CT and endoscopy score. Nakayama and colleagues<sup>127</sup> variants of low eosinophilic rhinitis were non-differential in symptoms, endoscopy and CT score.

## Discussion

### Summary of Key Findings

Our review reveals a wide variation in the phenotypes derived across all obstructive airway diseases investigated, both in children and adults, as well as variations in the methods used for phenotyping, study participants and population settings from where they have been recruited, and variables included in deriving the phenotypes. For asthma, the most reported phenotypes related to atopic status, time at disease onset, sex differences, disease symptomatology, and severity. For severe asthma, lung

function measures, atopic status and age at disease onset were the most characterizing features for defining phenotypes. COPD phenotypes were mostly characterized by lung function measures, as well as accompanying comorbidities, disease impairment on daily activity, and mortality and smoking status. Phenotypes of asthma and COPD overlap were mostly defined by smoking status, lung function measures, inflammation, and disease activity. Phenotypes of rhinitis were mostly defined by sex, disease severity, disease impairment on life and seasonality. The most reported unsupervised methods used for phenotyping were hierarchical and non-hierarchical clustering, particularly between 2010 and 2018. However, between 2015 and later, there was an increase in the use of other methods such as mixture-based model,<sup>69,126</sup> structural equation modelling,<sup>88</sup> and factor analysis with latent class modeling.<sup>10,13–16,20,23,31,42,44,48,50,58–60,123,125</sup>

## Results Interpretation

### Study Setting and Reporting of Airway Diseases Phenotypes

The majority of studies across outcomes enrolled participants from clinical settings, including from primary care centers,<sup>15</sup> tertiary hospitals,<sup>38</sup> pulmonary rehabilitation centers,<sup>24</sup> outpatient clinics,<sup>58</sup> and emergency departments.<sup>11</sup> Asthma studies that recruited participants from the general population reported commonly observed asthma phenotypes, such as mild early-onset asthma and mild atopic asthma<sup>14,18,41</sup>, which were comparable to those observed in clinical settings.<sup>12,15,21,22,31,34,35,37,38,50</sup> The phenotyping of asthma within clinical settings enabled the derivation of phenotypes mostly defined by measures easily obtained in clinics, compared to other epidemiological risk factors. For instance, phenotypes that were characterized by high response to treatment; low/high treatment adherence,<sup>12,56</sup> as well as courses of disease progression measured by symptoms or lung function were reported.<sup>19</sup> However, two studies conducted in general population settings reported asthma phenotypes of persistent and gradually improved wheeze or lung function,<sup>23,37</sup> as well as varying trajectories of attack rates progression and remission.<sup>51</sup> The work by Moore et al<sup>78</sup> is the only reported attempt of phenotyping severe asthma using a sample from general population. The characterized phenotypes of early-onset, atopic asthma, and late onset, non-atopic, severe obesity-related asthma with high utilization of health care were frequently reported in other studies conducted in clinical setting.<sup>79,81,82,86,87</sup>

Similarly, Kim et al<sup>101</sup> was the only reported work on phenotyping COPD based on a general population sample, it reported phenotypes of highly obstructed older males with mild COPD; highly obstructed with high symptoms and mild COPD among non-smoking obese with near normal functions. However, similar characteristics of such were reported in COPD variants reported from clinical setting.<sup>4,88,91–93,100</sup> Studies on COPD phenotypes derived from clinical setting, however, distinctly reported phenotypes of COPD with high to low emphysematous and air-entrapment changes,<sup>4,88,93,96,97,100,105</sup> as well as comorbidities.<sup>4,91,92,94,98,110,114</sup> Three studies phenotyped rhinitis based on samples from the general population.<sup>122,124,125</sup>

Burte et al<sup>122</sup> reported rhinitis phenotypes that are characterized by high nasal symptoms overall, high nasal symptoms through the year with low sensitization and seasonal spring nasal symptoms with high sensitization, while Kurukulaarachy et al<sup>124</sup> and Lee et al<sup>125</sup> reported variants of non-atopic rhinitis and atopic rhinitis with normal lung function. Rhinitis phenotypes derived from clinical setting reported similar variants of atopy-related rhinitis,<sup>120,123</sup> in addition to others related to high eosinophilia and nasal polyposis,<sup>119,127</sup> as well as quality of life and comorbidities.<sup>121,126</sup> Regarding asthma and COPD overlap, the reported study from general population sample was Fingleton et al<sup>118</sup> and revealed variants of smoking related, onset related, as well as atopic related ACO, which were comparable to similar reporting from clinical setting report.<sup>116</sup>

Overall, the characteristics of phenotypes derived from the general population settings were not always consistent with those derived from clinical settings. The observed variations could be due to variation in the types of variables readily available for measurement in general versus clinical settings. Further, presentation of patients of specific degree of disease severity, control, and treatment regimens varied across settings, and this could contribute to the variations between studies.

### Aspects Considered for Phenotyping Airway Diseases

In delineating the various aspects that contribute to the characterization of airway disease phenotypes, researchers have usually included variables that encompass etiological or risk factors, indicators of disease manifestation, and treatment behaviors, and prognostic indicators. Notably, among studies undertaken within the clinical setting, a significant focus has been placed on the physiological aspects of the disease, particularly lung function measures as vital parameters in phenotyping asthma and COPD. However, in the phenotyping of asthma, bronchodilatation (reversibility) is often not considered, which can potentially help in distinguishing different phenotypes with obstructive characteristic. Similarly, in

the context of COPD, the consideration of reversibility is also limited, despite its importance in differentiating COPD from other phenotypes, such as asthma-COPD overlap (ACO).<sup>129</sup>

Furthermore, the inclusion of bronchoprovocation tests in asthma phenotyping was seldom reported. Boudier et al<sup>14</sup> identified two phenotypes characterized by bronchial hyperresponsiveness, with overlapping allergic status but differing in symptom severity. Similarly, only two studies have considered airway remodeling in asthma phenotyping. Considering the importance of bronchial hyperresponsiveness and airway remodeling in asthma severity and monitoring of effectiveness of treatments, their inclusion in phenotyping will provide valuable insights. Number and list of variables used for phenotyping airway diseases are listed in [Tables S3–S7](#) in the supplementary material.

## Comparison with Previous Work

Our finding that the most reported asthma phenotypes were the ones differentiated by atopy, age at disease onset, and disease severity is consistent with the results from the review by Cunha et al and colleagues.<sup>130</sup> Our review similarly noted hierarchical cluster analysis as the most commonly reported method of phenotyping. Our review also found that majority of studies recruited participants from specialized healthcare centers. Pinto et al<sup>131</sup> reported similar results, indicating that hierarchical and non-hierarchical clustering were the most commonly reported methods to derive COPD phenotypes.

## Strength and Limitations

The current review followed recommended rigorous systematic review processes, including a priori protocol development, registration, and publication, and a comprehensive search of the literature across five leading healthcare databases, supplemented by grey literature and expert consultations. This approach minimized the risk of missing important studies. However, the variations in the included studies – in study design, outcome disease, methodological approaches employed, variables used for phenotyping, and derived phenotypes – inhibited comparison between studies. In addition, the review highlighted consistent issues with poor reporting practices, particularly regarding reproducibility, emphasizing the need for methodological improvements to enhance research quality and comparability in computational phenotyping of obstructive airway diseases.

## Future Research Implications

Harmonizing methodological approaches in computational phenotyping of obstructive airway diseases is essential. Developing a consensus on key variables for phenotyping and standardizing participant selection will enhance the comparability and interpretation of findings. The quality assessment tool created for this study addresses a significant gap, and broader application could lead to further improvements and consensus on evaluating study quality in this field. Many studies lack essential details that ensure validity, such as transparent reporting of data processing, handling of missing data, and the rationale for computational choices. Future research should prioritize full transparency and incorporate reproducibility tools, including code and data sharing, version control, and environment management systems, to improve documentation and sharing, ultimately advancing the quality and consistency of research in this area.

## Conclusion

The use of computational data-driven methods to derive phenotypes of airway diseases such as asthma, COPD, severe asthma, and ACO has resulted in significant variation in derived phenotypes across studies. This variability may be attributed to differences in sample selection, outcome measures, definitions, and variable selection used for phenotyping. The infrequent use of reproducibility measures in computational phenotyping research hinders the possibility of investigating the causes behind such variation. To achieve a better understanding and validity of the derived phenotypes and their clinical and scientific utility, a consistent approach to outcome definition and variable selection, as well as reproducible methods for phenotyping airway diseases, is needed.

## Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation. All authors took significant part in drafting, revising or

critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

## Funding

This study was performed under the Nordic Epilung project and the West Sweden Asthma Study which are supported by Nordforsk, the VBG Group Herman Krefting Foundation for Asthma and Allergy Research, the Swedish Heart-Lung Foundation, the Swedish Research Council, the Research Foundation of the Swedish Asthma and Allergy Association, and the Swedish government under the ALF agreement between the Swedish government and the county councils. None of the sponsors had any involvement in the planning, execution, drafting or write-up of this study.

## Disclosure

H. Kankaanranta reports personal fees for lectures and consulting from AstraZeneca, Boehringer-Ingelheim, Chiesi Pharma, Covis Pharma, GSK, MSD, Novartis, Orion Pharma and Sanofi Genzyme, outside the submitted work. H. Backman reports personal fees for lectures from AstraZeneca, Boehringer-Ingelheim, and GSK, outside the submitted work. A Lindberg reports personal fees from AstraZeneca, Boehringer Ingelheim, GSK and Novartis, outside the submitted work. All authors of this work declare no conflict of interests.

The protocol of this systematic review has been published in Systematic Reviews [[systematicreviewsjournal.biomedcentral.com](https://systematicreviewsjournal.biomedcentral.com)] as a Journal article: [<https://systematicreviewsjournal.biomedcentral.com/articles/10.1186/s13643-022-02078-0>]

## References

1. Weatherall M, Shirlcliffe P, Travers J, Beasley R. Use of cluster analysis to define COPD phenotypes. *Eur Respir J*. 2010;3(3):472–474. doi:10.1183/09031936.00035210
2. Castaldi PJ, Benet M, Petersen H, et al. Do COPD subtypes really exist? COPD heterogeneity and clustering in 10 independent cohorts. *Thorax*. 2017;72(11):998–1006. doi:10.1136/thoraxjnl-2016-209846
3. Weatherall M, Travers J, Shirlcliffe P, et al. Distinct clinical phenotypes of airways disease defined by cluster analysis. *Eur Respir J*. 2009;34(4):812–818. doi:10.1183/09031936.00174408
4. Burgel PR, Paillasseur JL, Peene B, et al. Two distinct chronic obstructive pulmonary disease (COPD) phenotypes are associated with high risk of mortality. *PLoS One*. 2012;7(12):e51048. doi:10.1371/journal.pone.0051048
5. Howard R, Rattray M, Prosperi M, Custovic A. Distinguishing Asthma Phenotypes Using Machine Learning Approaches. *Curr Allergy Asthma Rep*. 2015;15(7):38. doi:10.1007/s11882-015-0542-0
6. Prosperi MC, Sahiner UM, Belgrave D, et al. Challenges in identifying asthma subgroups using unsupervised statistical learning techniques. *Am J Respir Crit Care Med*. 2013;188(11):1303–1312. doi:10.1164/rccm.201304-0694OC
7. Bashir MBA, Basna R, Zhang G-Q, et al. Computational phenotyping of obstructive airway diseases: protocol for a systematic review. *Syst Rev*. 2022;11(1):216. doi:10.1186/s13643-022-02078-0
8. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev*. 2015;4(1):1. doi:10.1186/2046-4053-4-1
9. Amaral R, Pereira AM, Jacinto T, et al. Identification of asthma phenotypes in the US general population: a latent class analysis approach. *Allergy*. 2019;74:295–296.
10. Amelink M, de Nijs SB, De Groot JC, et al. Phenotypes of adult-onset asthma by cluster analysis. *Eur Respir J*. 2012;40:2.
11. Benton AS, Wang ZY, Lerner J, Foerster M, Teach SJ, Freishtat RJ. Overcoming Heterogeneity in Pediatric Asthma: tobacco Smoke and Asthma Characteristics Within Phenotypic Clusters in an African American Cohort. *J Asthma*. 2010;47(7):728–734. doi:10.3109/02770903.2010.491142
12. Bhargava S, Mahesh PA, Holla AD, et al. Cluster analysis identifies distinct clinical phenotypes with poor treatment responsiveness in asthma. *Eur Respir J*. 2018;52:2.
13. Bochenek G, Kuschill-Dziurda J, Szafraniec K, Plutecka H, Szczeklik A, Nizankowska-Mogilnicka E. Certain subphenotypes of aspirin-exacerbated respiratory disease distinguished by latent class analysis. *J Allergy Clin Immunol*. 2014;133(1):98–103. doi:10.1016/j.jaci.2013.07.004
14. Boudier A, Curjuric I, Basagana X, et al. Ten-Year Follow-up of Cluster-based Asthma Phenotypes in Adults A Pooled Analysis of Three Cohorts. *Am J Respir Crit Care Med*. 2013;188(5):550–560. doi:10.1164/rccm.201301-0156OC
15. Cabral ALB, Sousa AW, Mendes FAR, De Carvalho CRF. Phenotypes of asthma in low income children and adolescents: cluster analysis. *J Bras Pneumol*. 2017;43(1):44–50. doi:10.1590/s1806-37562016000000039
16. Celejewski-Wojcik N, Wojcik K, Ignacak-Popiel M, et al. Subphenotypes of NSAID-exacerbated respiratory disease identified by latent class analysis. *Allergy*. 2019;75(4):831–840.
17. Chanoine S, Pin I, Sanchez M, et al. Asthma Medication Ratio Phenotypes in Elderly Women. *J Allergy Clinical Immunol*. 2018;6(3):897–906. doi:10.1016/j.jaip.2017.07.014
18. Couto M, Stang J, Horta L, et al. Two distinct phenotypes of asthma in elite athletes identified by latent class analysis. *J Asthma*. 2015;52(9):897–904. doi:10.3109/02770903.2015.1067321

19. Cruz AA, Bansal AT, Ponte EV, et al. Unbiased clinical cluster analysis of patients of ProAR Asthma Cohort. *Eur Respir J*. 2018;52:2.
20. Damiens K, Paquet J, Lemiere C. Identification Of Work-Related Asthma Phenotypes By Cluster Analysis. *Am J Respir Crit Care Med*. 2013;187(1):1. doi:10.1164/rccm.201211-1966ED
21. Deliu M, Yavuz ST, Sperrin M, et al. Hierarchical clustering identifies novel subgroups of childhood asthma. *Allergy*. 2016;71:98.
22. Deliu M, Yavuz TS, Sperrin M, et al. Features of asthma which provide meaningful insights for understanding the disease heterogeneity. *Clin Exp Allergy*. 2018;48(1):39–47. doi:10.1111/cea.13014
23. Depner M, Fuchs O, Genuneit J, et al. Clinical and Epidemiologic Phenotypes of Childhood Asthma. *Am J Respir Crit Care Med*. 2014;189(2):129–138. doi:10.1164/rccm.201307-1198OC
24. Dudchenko LS, Savchenko VM. Cluster analysis classification of asthmatic pathologic manifestations during stay at the resort. *Tuberculosis and Lung Diseases*. 2018;96(2):16–21. doi:10.21292/2075-1230-2018-96-2-16-21
25. Folz RJ, Myers J, Jorayeva A, Folz RJ. Phenotyping Older Adults with Asthma Via Clinical Cluster Analysis. *Am J Respir Crit Care Med*. 2018;197(1):2. doi:10.1164/rccm.201708-1549ED
26. Fontanella S, Frainay C, Murray CS, Simpson A, Custovic A. Machine learning to identify pairwise interactions between specific IgE antibodies and their association with asthma: a cross-sectional analysis within a population-based birth cohort. *PLoS Med*. 2018;15(11):22. doi:10.1371/journal.pmed.1002691
27. Gonem S, Natarajan S, Hartley R, et al. Cluster analysis reveals a distinct small airway-predominant phenotype of asthma. *Thorax*. 2012;67:A7.
28. Gower WA, Li X, Li H, et al. Cluster Analysis Of Asthma Phenotypes In Children Ages 6 To 18 Years From The Eve Consortium. *Am J Respir Crit Care Med*. 2013;187(1):1.
29. Hilvering B, Jansen J, Lammers JWJ, Koenderman L. Phenotyping Asthma Using An Unsupervised Prediction Model Based On Blood Granulocyte Responsiveness. *Am J Respir Crit Care Med*. 2015;191:1.
30. Hsiao HP, Lin MC, Wu CC, Wang CC, Wang TN. Sex-Specific Asthma Phenotypes, Inflammatory Patterns, and Asthma Control in a Cluster Analysis. *J Allergy Clinical Immunol*. 2019;7(2):556–567. doi:10.1016/j.jaip.2018.08.008
31. Ilmarinen P, Tuomisto LE, Niemela O, Tammola M, Haanpää J, Kankaanranta H. Cluster Analysis on Longitudinal Data of Patients with Adult-Onset Asthma. *J Allergy Clinical Immunol*. 2017;5(4):967–978. doi:10.1016/j.jaip.2017.01.027
32. Jeong A, Imboden M, Hansen S, et al. Heterogeneity of obesity-asthma association disentangled by latent class analysis, the SAPALDIA cohort. *Respir Med*. 2017;125:25–32. doi:10.1016/j.rmed.2017.02.014
33. Just J, Gouvis-Echraghi R, Rouve S, Wanin S, Moreau D, Annesi-Maesano I. Two novel, severe asthma phenotypes identified during childhood using a clustering approach. *Eur Respir J*. 2012;40(1):55–60. doi:10.1183/09031936.00123411
34. Just J, Saint-Pierre P, Gouvis-Echraghi R, et al. Childhood Allergic Asthma Is Not a Single Phenotype. *J Paediatr*. 2014;164(4):815–820. doi:10.1016/j.jpeds.2013.11.037
35. Kaneko Y, Masuko H, Sakamoto T, et al. Asthma phenotypes in Japanese adults - their associations with the CCL5 and ADRB2 genotypes. *Allergol Int*. 2013;62(1):113–121. doi:10.2332/allergolint.12-OA-0467
36. Kim JH, Chang HS, Shin SW, et al. Lung function trajectory types in never-smoking adults with asthma: clinical features and inflammatory patterns. *Allergy Asthma Immunol Res*. 2018;10(6):614–627. doi:10.4168/air.2018.10.6.614
37. Kim MA, Shin SW, Park JS, et al. Clinical characteristics of exacerbation-prone adult asthmatics identified by cluster analysis. *Allergy Asthma Immunol Res*. 2017;9(6):483–490. doi:10.4168/air.2017.9.6.483
38. Kim TB, Jang AS, Kwon HS, et al. Identification of asthma clusters in two independent Korean adult asthma cohorts. *Eur Respir J*. 2013;41(6):1308–1314. doi:10.1183/09031936.00100811
39. Koike F, Otani Y, Oyama S, et al. Cluster analysis of cough variant asthma using exhaled value of forced oscillation technique. *Eur Respir J*. 2018;52:3.
40. Kwon J, Seo J, Kim H, et al. Asthma Phenotypes in School-aged Children from the Population Study: cluster Analysis. *J Allergy Clin Immunol*. 2012;129(2):AB7–AB7. doi:10.1016/j.jaci.2011.12.885
41. Lee E, Lee SH, Kwon JW, et al. Persistent asthma phenotype related with late-onset, high atopy, and low socioeconomic status in school-aged Korean children. *BMC Pulm Med*. 2017;17(1):11. doi:10.1186/s12890-017-0387-5
42. Liang ZY, Liu LY, Zhao HJ, et al. A Systemic Inflammatory Endotype of Asthma With More Severe Disease Identified by Unbiased Clustering of the Serum Cytokine Profile. *Medicine*. 2016;95(25):7. doi:10.1097/MD.00000000000003774
43. Loureiro CC, Sa-Couto P, Todo-Bom A, Bousquet J. Cluster analysis in phenotyping a Portuguese population. *Revista Portuguesa de Pneumologia*. 2015;21(6):299–306. doi:10.1016/j.rppnen.2015.07.006
44. Loza MJ, Djukanovic R, Chung KF, et al. Validated and longitudinally stable asthma phenotypes based on cluster analysis of the ADEPT study. *Respir Res*. 2016;17(1):21. doi:10.1186/s12931-016-0482-9
45. Mahut B, Peyrard S, Delclaux C. Exhaled nitric oxide and clinical phenotypes of childhood asthma. *Respir Res*. 2011;12(1):65. doi:10.1186/1465-9921-12-65
46. Makikyro EMS, Jaakkola MS, Jaakkola JJK. Subtypes of asthma based on asthma control and severity: a latent class analysis. *Respir Res*. 2017;18(1):11. doi:10.1186/s12931-017-0508-y
47. Mason P, Frigo AC, Scarpa MC, Maestrelli P, Guarnieri G. Cluster analysis of occupational asthma caused by isocyanates. *J Allergy Clin Immunol*. 2018;142(6):2011. doi:10.1016/j.jaci.2018.08.018
48. Mastalerz L, Celejewski-Wojcik N, Wojcik K, et al. Induced sputum supernatant bioactive lipid mediators can identify subtypes of asthma. *Clin Exp Immunol*. 2015;45(12):1779–1789. doi:10.1111/cea.12654
49. Nadif R, Febrissey M, Andrianjafimasy M, et al. Adult asthma phenotypes identified by a cluster analysis on clinical and biological characteristics. *Eur Respir J*. 2018;52:2.
50. Nagasaki T, Matsumoto H, Kanemitsu Y, et al. Integrating longitudinal information on pulmonary function and inflammation using asthma phenotypes. *J Allergy Clin Immunol*. 2014;133(5):1474–U406. doi:10.1016/j.jaci.2013.12.1084
51. Nasreen S, Wilk P, Mallowney T, Karp I. Asthma exacerbation trajectories and their predictors in children with incident asthma. *Ann Allergy Asthma Immunol*. 2019;123(3):293. doi:10.1016/j.anai.2019.05.013
52. Qiu RH, Xie JX, Chung KF, et al. Asthma Phenotypes Defined From Parameters Obtained During Recovery From a Hospital-Treated Exacerbation. *J Allergy Clinical Immunol*. 2018;6(6):1960–1967. doi:10.1016/j.jaip.2018.02.012

53. Sakagami T, Hasegawa T, Koya T, et al. Identification Of Clinical Asthma Phenotypes By Using Cluster Analysis With Simple Measurable Variables In Japanese Population. *Am J Respir Crit Care Med*. 2011;183(1):1. doi:10.1164/rccm.201009-1482ED
54. Schatz M, Hsu JWY, Zeiger RS, et al. Phenotypes determined by cluster analysis in severe or difficult-to-treat asthma. *J Allergy Clin Immunol*. 2014;133(6):1549–1556. doi:10.1016/j.jaci.2013.10.006
55. Schimdlin KA, Brokamp C, LeMasters GK, et al. Distinct Phenotypes of Childhood Asthma: cluster Analysis in a Longitudinal Birth Cohort. *J Allergy Clin Immunol*. 2015;135(2):AB85–AB85. doi:10.1016/j.jaci.2014.12.1210
56. Seino Y, Hasegawa T, Koya T, et al. A Cluster Analysis of Bronchial Asthma Patients with Depressive Symptoms. *Internal Medicine*. 2018;57(14):1967–1975. doi:10.2169/internalmedicine.9073-17
57. Sendin-Hernandez MP, Avila-Zarza C, Sanz C, et al. Cluster Analysis Identifies 3 Phenotypes within Allergic Asthma. *J Allergy Clinical Immunol*. 2018;6(3):955–961. doi:10.1016/j.jaip.2017.10.006
58. Seys SF, Scheers H, Van den Brande P, et al. Cluster analysis of sputum cytokine-high profiles reveals diversity in T(h) 2-high asthma patients. *Respir Res*. 2017;18(1):10. doi:10.1186/s12931-017-0524-y
59. Siroux V, Basagaña X, Boudier A, et al. Identifying adult asthma phenotypes using a clustering approach. *Pneumologia*. 2011;60(3):166–173.
60. Tay TR, Choo XN, Yii A, et al. Asthma phenotypes in a multi-ethnic Asian cohort. *Respir Med*. 2019;157:42–48. doi:10.1016/j.rmed.2019.08.016
61. Tsukioka K, Koya T, Ueno H, et al. Phenotypic analysis of asthma in Japanese athletes. *Allergol Int*. 2017;66(4):550–556. doi:10.1016/j.alit.2017.02.009
62. Wang L, Liang R, Zhou T, et al. Identification and validation of asthma phenotypes in Chinese population using cluster analysis. *Ann Allergy Asthma Immunol*. 2017;119(4):324–332. doi:10.1016/j.anai.2017.07.016
63. Watanabe S, Koya T, Hasegawa T, et al. Cluster Analysis Of Uncontrolled Asthma In Japanese Population. *Am J Respir Crit Care Med*. 2016;193(1):1. doi:10.1164/rccm.201509-1801ED
64. Wisnivesky JP, Rojano B, Chen S, et al. Asthma Phenotypes in World Trade Center Workers: a Cluster Analysis. *Am J Respir Crit Care Med*. 2019;199(1):2. doi:10.1164/rccm.201807-1299ED
65. Wu DW, Bleier BS, Li L, et al. Clinical Phenotypes of Nasal Polyps and Comorbid Asthma Based on Cluster Analysis of Disease History. *J Allergy Clinical Immunol*. 2018;6(4):1297–1305. doi:10.1016/j.jaip.2017.09.020
66. Zaihra T, Walsh CJ, Ahmed S, et al. Phenotyping of difficult asthma using longitudinal physiological and biomarker measurements reveals significant differences in stability between clusters. *BMC Pulm Med*. 2016;16(1):8. doi:10.1186/s12890-016-0232-2
67. Zhang XX, Xia TT, Lai ZD, et al. Uncontrolled asthma phenotypes defined from parameters using quantitative CT analysis. *Eur Radiol*. 2019;29(6):2848–2858. doi:10.1007/s00330-018-5913-1
68. Zoratti EM, Krouse RZ, Babineau DC, et al. Asthma phenotypes in inner-city children. *J Allergy Clin Immunol*. 2016;138(4):1016–1029. doi:10.1016/j.jaci.2016.06.061
69. Brinkman P, Wagener AH, Bansal AT, et al. Electronic Noses Capture Severe Asthma Phenotypes By Unbiased Cluster Analysis. *Am J Respir Crit Care Med*. 2014;189:2.
70. Desai D, May RD, Haldar P, et al. Cytokine Profiling In Severe Asthma Subphenotypes Using Factor And Cluster Analysis. *Am J Respir Crit Care Med*. 2011;183(1):1.
71. Diver S, Richardson M, Desai D, et al. Hierarchical cluster analysis of subjects with severe asthma according to microbiological profile. *Thorax*. 2018;73:A20.
72. Fitzpatrick AM, Teague WG, Meyers DA, et al. Heterogeneity of severe asthma in childhood: confirmation by cluster analysis of children in the National Institutes of Health/National Heart, Lung, and Blood Institute Severe Asthma Research Program. *J Allergy Clin Immunol*. 2011;127(2):382–389. doi:10.1016/j.jaci.2010.11.015
73. Freitas PD, Xavier RF, Da Silva STC, et al. Identification of asthma phenotypes using cluster analysis. *Eur Respir J*. 2018;52:2.
74. Gomez JL, Yan XT, Holm CT, et al. Characterisation of asthma subgroups associated with circulating YKL-40 levels. *Eur Respir J*. 2017;50(4):12. doi:10.1183/13993003.00800-2017
75. Jang AS, Kwon HS, Cho YS, et al. Identification of subtypes of refractory asthma in Korean patients by cluster analysis. *Lung*. 2013;191(1):87–93. doi:10.1007/s00408-012-9430-8
76. Konstantellou E, Papaioannou AI, Loukides S, et al. Persistent airflow obstruction in patients with asthma: characteristics of a distinct clinical phenotype. *Respir Med*. 2015;109(11):1404–1409. doi:10.1016/j.rmed.2015.09.009
77. Lau PPK, Yii ACA, Chan AKW, Ganguly R, Devanand A, Koh MS. Cluster analysis of severe life-threatening asthma: identification of a distinct phenotype characterized by high serum eosinophils. *Eur Respir J*. 2017;50:2.
78. Moore WC, Meyers DA, Wenzel SE, et al. Identification of Asthma Phenotypes Using Cluster Analysis in the Severe Asthma Research Program. *Am J Respir Crit Care Med*. 2010;181(4):315–323. doi:10.1164/rccm.200906-0896OC
79. Newby C, Heaney LG, Menzies-Gow A, et al. Statistical cluster analysis of the British thoracic society severe refractory asthma registry: clinical outcomes and phenotype stability. *PLoS One*. 2014;9(7):e102987. doi:10.1371/journal.pone.0102987
80. Raherison-Semjen C, Prudhomme A, Nocent-Eijnani C, et al. FASE-CPHG Study: identification of asthma phenotypes in the French Severe Asthma Study using cluster analysis. *Eur Respir J*. 2018;52:2.
81. Sekiya K, Nakatani E, Fukutomi Y, et al. Severe or life-threatening asthma exacerbation: patient heterogeneity identified by cluster analysis. *Clin Exp Immunol*. 2016;46(8):1043–1055. doi:10.1111/cea.12738
82. Serrano-Pariente J, Rodrigo G, Fiz JA, Crespo A, Plaza V. Identification and characterization of near-fatal asthma phenotypes by cluster analysis. *Allergy*. 2015;70(9):1139–1147. doi:10.1111/all.12654
83. Simpson A, Hekking PP, Shaw D, et al. Late Breaking Abstract - Cluster analysis of treatable traits in the U-BIOPRED adult severe asthma cohort. *Eur Respir J*. 2017;50:1.
84. Taniguchi N, Konno S, Makita H, et al. Cluster Analysis Of Severe Adult Asthma Patients, Including Smokers, In A Japanese Population. *Am J Respir Crit Care Med*. 2014;189(1):1. doi:10.1164/rccm.201310-1751ED
85. Wu W, Bleecker E, Moore W, et al. Unsupervised phenotyping of Severe Asthma Research Program participants using expanded lung data. *J Allergy Clin Immunol*. 2014;133(5):1280–1288. doi:10.1016/j.jaci.2013.11.042

86. Ye WJ, Xu WG, Guo XJ, et al. Differences in airway remodeling and airway inflammation among moderate-severe asthma clinical phenotypes. *J Thoracic Dis.* 2017;9(9):2904–2914. doi:10.21037/jtd.2017.08.01
87. Youroukova VM, Dimitrova DG, Valerieva AD, et al. Phenotypes Determined by Cluster Analysis in Moderate to Severe Bronchial Asthma. *Folia Med.* 2017;59(2):165–173. doi:10.1515/folmed-2017-0031
88. Augustin IML, Spruit MA, Houben-Wilke S, et al. The respiratory physiome: clustering based on a comprehensive lung function assessment in patients with COPD. *PLoS One.* 2018;13(9):14. doi:10.1371/journal.pone.0201593
89. Bafadhel M, McKenna S, Terry S, et al. Acute Exacerbations of Chronic Obstructive Pulmonary Disease Identification of Biologic Clusters and Their Biomarkers. *Am J Respir Crit Care Med.* 2011;184(6):662–671. doi:10.1164/rccm.201104-0597OC
90. Bertini I, Luchinat C, Miniati M, Monti S, Tenori L. Phenotyping COPD by 1H NMR metabolomics of exhaled breath condensate. *Metabolomics.* 2014;10(2):302–311. doi:10.1007/s11306-013-0572-3
91. Burgel PR, Paillasseur JL, Caillaud D, et al. Clinical COPD phenotypes: a novel approach using principal component and cluster analyses. *Eur Respir J.* 2010;36(3):531–539. doi:10.1183/09031936.00175109
92. Burgel PR, Paillasseur JL, Janssens W, et al. A simple algorithm for the identification of clinical COPD phenotypes. *Eur Respir J.* 2017;50(5):11. doi:10.1183/13993003.01034-2017
93. Chen CZ, Wang LY, Ou CY, Lee CH, Lin CC, Hsiue TR. Using Cluster Analysis to Identify Phenotypes and Validation of Mortality in Men with COPD. *Lung.* 2014;192(6):889–896. doi:10.1007/s00408-014-9646-x
94. Chubachi S, Tsutsumi A, Kameyama N, et al. Cluster Analysis Based On Comorbidities For Japanese COPD Patients. *Am J Respir Crit Care Med.* 2016;193(1):1.
95. De Torres JP, Marin JM, Martinez-Gonzalez C, De lucas-ramos P, Cosio B, Casanova C. The importance of symptoms in the longitudinal variability of clusters in COPD patients: a validation study. *Respirology.* 2018;23(5):485–491. doi:10.1111/resp.13194
96. Divo M, Casanova C, Marin JM, et al. Identification of clinical phenotypes in patients with and without COPD using cluster analysis. *Eur Respir J.* 2016;48(1):3. doi:10.1183/13993003.01046-2016
97. Fens N, van Rossum AGJ, Zanen P, et al. Subphenotypes of Mild-to-Moderate COPD by Factor and Cluster Analysis of Pulmonary Function, CT Imaging and Breathomics in a Population-Based Survey. *COPD.* 2013;10(3):277–285. doi:10.3109/15412555.2012.744388
98. Guillaumet RV, Ursu O, Iwamoto G, Moseley PL, Oprea T. Chronic obstructive pulmonary disease phenotypes using cluster analysis of electronic medical records. *Health Informatics Journal.* 2018;24(4):394–409. doi:10.1177/1460458216675661
99. Haghighi B, Choi S, Choi J, et al. Imaging-based clusters in former smokers of the COPD cohort associate with clinical characteristics: the SubPopulations and intermediate outcome measures in COPD study (SPIROMICS). *Respir Res.* 2019;20(1):14. doi:10.1186/s12931-019-1121-z
100. Harrison SL, Robertson N, Graham CD, et al. Can we identify patients with different illness schema following an acute exacerbation of COPD: a cluster analysis. *Respir Med.* 2014;108(2):319–328. doi:10.1016/j.rmed.2013.10.016
101. Kim S, Lim MN, Hong Y, Han SS, Lee SJ, Kim WJ. A cluster analysis of chronic obstructive pulmonary disease in dusty areas cohort identified three subgroups. *BMC Pulm Med.* 2017;17(1):9. doi:10.1186/s12890-017-0553-9
102. Kim WJ, Gupta V, Nishimura M, et al. Identification of chronic obstructive pulmonary disease subgroups in 13 Asian cities. *Int J Tuberc Lung Dis.* 2018;22(7):820. doi:10.5588/ijtld.17.0524
103. Kukol LVP, A S. Determination of phenotypic characteristics of chronic obstructive lung disease in elderly patients. *Adv Gerontol.* 2019;32(3):445–450.
104. Lee J, Yoon H, Lee S, et al. Identification of Subtypes Predicting Exacerbations in COPD Using Cluster Analysis. *Am J Respir Crit Care Med.* 2019;199(1):2.
105. Li X, Han MK, Ortega VE, et al. Cluster Analysis Of Chronic Obstructive Pulmonary Disease (COPD) Related Phenotypes In The Subpopulations And Intermediate Outcome Measures In COPD Study (spiromics). *Am J Respir Crit Care Med.* 2016;193(1):1.
106. Liang ZY, Long F, Wang FY, et al. Identification of clinically relevant subgroups of COPD based on airway and circulating autoantibody profiles. *Mol Med Rep.* 2019;20(3):2882–2892. doi:10.3892/mmr.2019.10498
107. Lopes AC, Xavier RF, Pereira A, et al. Identifying COPD patients at risk for worse symptoms, HRQoL, and self-efficacy: a cluster analysis. *Chronic Illness.* 2019;15(2):138–148. doi:10.1177/1742395317753883
108. Ning PG, F. Y, Sun TY, Zhang HS, Chai D, Li XM. Study of the clinical phenotype of symptomatic chronic airways disease by hierarchical cluster analysis and two-step cluster analyses. *Zhonghua Nei Ke Za Zhi.* 2016;55(9):679–683. doi:10.3760/cma.j.issn.0578-1426.2016.09.005
109. Peters JB, Boer LM, Molema J, Heijdra YF, Prins JB, Vercoulen JH. Integral Health Status-Based Cluster Analysis in Moderate-Severe COPD Patients Identifies Three Clinical Phenotypes: relevant for Treatment As Usual and Pulmonary Rehabilitation. *Int J Behav Med.* 2017;24(4):571–583. doi:10.1007/s12529-016-9622-3
110. 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. Inf. Decis. Making.* 2019;19(1):14. doi:10.1186/s12911-019-0805-0
111. Rodrigues A, Camillo CA, Furlanetto KC, et al. Cluster analysis identifying patients with COPD at high risk of 2-year all-cause mortality. *Chron Respir Dis.* 2018;16:8.
112. Scarlata S, Finamore P, Santangelo S, et al. Cluster analysis on breath print of newly diagnosed COPD patients: effects of therapy. *J Breath Res.* 2018;12(3):10. doi:10.1088/1752-7163/aac273
113. Xavier RFP, Lopes AC, Cavalheri V, et al. Identification of Phenotypes in People with COPD: influence of Physical Activity, Sedentary Behaviour, Body Composition and Skeletal Muscle Strength. *Lung.* 2019;197(1):37–45. doi:10.1007/s00408-018-0177-8
114. Yoon HY, Park SY, Lee CH, et al. Prediction of first acute exacerbation using COPD subtypes identified by cluster analysis. *Int J Chronic Obstr.* 2019;14:1389–1397. doi:10.2147/COPD.S205517
115. De Vries R, Dagelet YWF, Spoor P, et al. Clinical and inflammatory phenotyping by breathomics in chronic airway diseases irrespective of the diagnostic label. *Eur Respir J.* 2018;51(1):10. doi:10.1183/13993003.01817-2017
116. Gorska K, Paplinska-Goryca M, Nejman-Gryz P, Goryca K, Krenke R. Eosinophilic and Neutrophilic Airway Inflammation in the Phenotyping of Mild-to-Moderate Asthma and Chronic Obstructive Pulmonary Disease. *COPD.* 2017;14(2):181–189. doi:10.1080/15412555.2016.1260539
117. Rootmensen G, van Keimpema A, Zwinderman A, Sterk P. Clinical phenotypes of obstructive airway diseases in an outpatient population. *J Asthma.* 2016;53(10):1026–1032. doi:10.3109/02770903.2016.1174258

118. Fingleton J, Huang KW, Weatherall M, et al. Phenotypes of symptomatic airways disease in China and New Zealand. *Eur Respir J*. 2017;50(6):10. doi:10.1183/13993003.00957-2017
119. Adnane C, Adouly T, Khallouk A, et al. Using preoperative unsupervised cluster analysis of chronic rhinosinusitis to inform patient decision and endoscopic sinus surgery outcome. *Europ Archiv Oto-Rhino-Laryngol*. 2017;274(2):879–885. doi:10.1007/s00405-016-4315-8
120. Agache I, Ciobanu C. Risk Factors and Asthma Phenotypes in Children and Adults with Seasonal Allergic Rhinitis. *Physician and Sportsmedicine*. 2010;38(4):81–86. doi:10.3810/psm.2010.12.1829
121. Bousquet PJ, Devillier P, Tadmouri A, Mesbah K, Demoly P, Bousquet J. Clinical Relevance of Cluster Analysis in Phenotyping Allergic Rhinitis in a Real-Life Study. *Int Arch Allergy Immunol*. 2015;166(3):231–240. doi:10.1159/000381339
122. Burte E, Bousquet J, Varraso R, et al. Characterization of Rhinitis According to the Asthma Status in Adults Using an Unsupervised Approach in the EGEA Study. *PLoS One*. 2015;10(8):18. doi:10.1371/journal.pone.0136191
123. Herr M, Just J, Nikasinovic L, et al. Risk factors and characteristics of respiratory and allergic phenotypes in early childhood. *Journal of Allergy and Clinical Immunology*. 2012;130(2):389–396.
124. Kurukulaaratchy RJ, Zhang HM, Patil V, et al. Identifying the heterogeneity of young adult rhinitis through cluster analysis in the Isle of Wight birth cohort. *J Allergy Clin Immunol*. 2015;135(1):143–U225. doi:10.1016/j.jaci.2014.06.017
125. Lee E, Lee SH, Kwon JW, et al. A rhinitis phenotype associated with increased development of bronchial hyperresponsiveness and asthma in children. *Ann Allergy Asthma Immunol*. 2016;117(1):21–28.e1. doi:10.1016/j.anai.2016.04.016
126. Soler ZM, Hyer JM, Ramakrishnan V, et al. Identification of chronic rhinosinusitis phenotypes using cluster analysis. *Int Forum Allergy Rhinol*. 2015;5(5):399–407. doi:10.1002/alr.21496
127. Nakayama T, Asaka D, Yoshikawa M, et al. Identification of chronic rhinosinusitis phenotypes using cluster analysis. *Am J Rhinol Allergy*. 2012;26(3):172–176. doi:10.2500/ajra.2012.26.3749
128. Hsiao HPL, C. M, Wu CC, Wang CC, Wang TN. Sex-Specific Asthma Phenotypes, Inflammatory Patterns, and Asthma Control in a Cluster Analysis. *J Allergy Clinical Immunol*. 2019;7(2):556.
129. Leung JM, Sin DD. Asthma-COPD overlap syndrome: pathogenesis, clinical features, and therapeutic targets. *BMJ*. 2017;358:j3772. doi:10.1136/bmj.j3772
130. Cunha F, Amaral R, Jacinto T, Sousa-Pinto B, Fonseca JA. A Systematic Review of Asthma Phenotypes Derived by Data-Driven Methods. *Diagnostics*. 2021;11(4):644. doi:10.3390/diagnostics11040644
131. Pinto LM, Alghamdi M, Benedetti A, Zaihra T, Landry T, Bourbeau J. Derivation and validation of clinical phenotypes for COPD: a systematic review. *Respir Res*. 2015;16(1):50. doi:10.1186/s12931-015-0208-4

## Journal of Asthma and Allergy

### Publish your work in this journal

The Journal of Asthma and Allergy is an international, peer-reviewed open-access journal publishing original research, reports, editorials and commentaries on the following topics: Asthma; Pulmonary physiology; Asthma related clinical health; Clinical immunology and the immunological basis of disease; Pharmacological interventions and new therapies. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/journal-of-asthma-and-allergy-journal>

**Dovepress**  
Taylor & Francis Group