Published in final edited form as:
Am J Respir Crit Care Med. 2008 August 1; 178(3): 218–224. doi:10.1164/rccm.200711-1754OC.

Cluster Analysis and Clinical Asthma Phenotypes

Pranab Haldar[#1], **Ian D. Pavord**[#1], **Dominic E. Shaw**[1], **Michael A. Berry**[1], **Michael Thomas**[2], **Christopher E. Brightling**[1], **Andrew J. Wardlaw**[1], and **Ruth H. Green**[#1]

[1]Institute for Lung Health, Glenfield Hospital, Leicester, United Kingdom

[2]Department of General Practice, University of Aberdeen, Aberdeen, United Kingdom

[#] These authors contributed equally to this work.

Abstract

Rationale—Heterogeneity in asthma expression is multidimensional, including variability in clinical, physiologic, and pathologic parameters. Classification requires consideration of these disparate domains in a unified model.

Objectives—To explore the application of a multivariate mathematical technique, k-means cluster analysis, for identifying distinct phenotypic groups.

Methods—We performed k-means cluster analysis in three independent asthma populations. Clusters of a population managed in primary care (n = 184) with predominantly mild to moderate disease, were compared with a refractory asthma population managed in secondary care (n = 187). We then compared differences in asthma outcomes (exacerbation frequency and change in corticosteroid dose at 12 mo) between clusters in a third population of 68 subjects with predominantly refractory asthma, clustered at entry into a randomized trial comparing a strategy of minimizing eosinophilic inflammation (inflammation-guided strategy) with standard care.

Measurements and Main Results—Two clusters (early-onset atopic and obese, noneosinophilic) were common to both asthma populations. Two clusters characterized by marked discordance between symptom expression and eosinophilic airway inflammation (early-onset symptom predominant and late-onset inflammation predominant) were specific to refractory

Correspondence and requests for reprints should be addressed to Dr. P. Haldar, M.R.C.P., Institute for Lung Health, Glenfield Hospital, Leicester, LE3 9QP, UK. ph62@le.ac.uk.

Conflict of Interest Statement: P.H. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. I.D.P. received $2,000 for speaking at conferences organized by GlaxoSmithKline and $5,000 for speaking at conferences organized by AstraZeneca; he is in receipt of a $500,000 grant for a study of severe asthma from GlaxoSmithKline. D.E.S. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. M.A.B. has received lecture fees and conference support from AstraZeneca and GlaxoSmithKline. M.T. has received speaker's honoraria in the last 3 years for speaking at meetings sponsored by the following companies marketing respiratory products: AstraZeneca, Boehringer Ingelheim, GlaxoSmithKline, MSD, Schering-Plough, Teva; he has received honoraria for attending advisory panels with Altana, AstraZeneca, Boehringer Ingelheim, GlaxoSmithKline, MSD, Merck Respiratory, Schering-Plough, Teva; he has received sponsorships to attend international scientific meetings from GlaxoSmithKline, MSD, AstraZeneca; he has received funding for research projects from GlaxoSmithKline, MSD, AstraZeneca; he holds a research fellowship from Asthma UK. C.E.B. has received a total of $2.2 million in research funding over the last 3 years (or is pending) from AstraZeneca, Cambridge Antibody Technology, GlaxoSmithKline; he has received less than $10,000 per annum from consultancy fees from Cambridge Antibody Technology, AstraZeneca, GlaxoSmithKline, Roche, and Pfizer; he has participated as a speaker in scientific meetings or courses organized and financed by AstraZeneca, GlaxoSmithKline, Boehringer Ingelheim, MSD, and Pfizer. A.J.W. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. R.H.G. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript.

Europe PMC Funders Author Manuscripts

asthma. Inflammation-guided management was superior for both discordant subgroups leading to a reduction in exacerbation frequency in the inflammation-predominant cluster (3.53 [SD, 1.18] vs. 0.38 [SD, 0.13] exacerbation/patient/yr, $P = 0.002$) and a dose reduction of inhaled corticosteroid in the symptom-predominant cluster (mean difference, 1,829 μg beclomethasone equivalent/d [95% confidence interval, 307–3,349 μg]; $P = 0.02$).

Conclusions—Cluster analysis offers a novel multidimensional approach for identifying asthma phenotypes that exhibit differences in clinical response to treatment algorithms.

Keywords

taxonomy; corticosteroid response; multivariate classification

Asthma impacts significantly on the rising burden of chronic disease in developed countries. Approximately 5 to 10% of sufferers have refractory asthma that remains poorly controlled despite maximal inhaled therapy (1). Effective clinical care is complicated by heterogeneity in the physiologic, pathologic, and molecular abnormalities associated with refractory asthma (2). Current descriptions of asthma phenotypes are limited by subjectivity and poor coherence. A robust system of classification that incorporates the multidimensionality of asthma is needed to identify subgroups with consistent patterns of disease (3, 4). This may provide a framework for identifying distinct phenotypes, with specific pathophysiologic abnormalities that predict response to particular therapies (5) and help to focus current genetic and molecular studies.

The taxonomy of organisms remains the paradigm for biological models of classification. It is based empirically on the principle that similarity measured across a number of different characteristics predicts relationships of biological significance with greater probability. Cluster analysis refers to a group of multivariate mathematical algorithms that broadly perform two distinct functions: (*1*) quantification of similarity between individuals within a population on the basis of the (multiple) specified variables; (*2*) grouping of individuals into clusters such that similarity between members of the same clusters is strong and using between different clusters is weak (6, 7). The principal advantage of performing classification numerically is objectivity and using methodology for including multiple variables that assume equal weighting helps minimize *a priori* bias. Numerical taxonomy or taximetrics is the branch of taxonomy that has developed to use mathematical algorithms such as cluster analysis for this purpose (8), and the principle has been extended for use in other areas of biomedical science, notably bioinformatics and psychiatry (9). In the latter, cluster analysis techniques have been used to identify patterns of symptom expression that have been used to define diagnostic categories (9).

We postulated that cluster analysis could be applied for classifying clinical phenotypes of asthma. We examined this hypothesis using the k-means clustering algorithm to classify two distinct asthma populations: a group recruited from primary care with asthma of predominantly mild to moderate severity and a group from secondary care who met prespecified criteria for refractory asthma (10). The clinical relevance of these clusters was evaluated further by investigating differences in asthma outcomes between clusters identified in a separate cohort of patients with predominantly refractory asthma, who

Europe PMC Funders Author Manuscripts

participated in a recently completed randomized study at our center comparing a management strategy aimed at titrating steroid therapy to maintain a normal sputum eosinophil count, with a conventional clinical protocol (11). Some of the results of this study have been previously reported in the form of an abstract (12).

METHODS

Subjects

We studied three discrete populations with asthma. All patients had a physician diagnosis of asthma and sufficient symptoms to warrant at least one prescription for asthma therapy in the previous 12 months. All patients were current nonsmokers and ex-smokers had a less than 10 pack-year smoking history. The two larger datasets comprised cross-sectional data for performing cluster analysis to identify the major disease patterns existing, respectively, within primary-care and refractory asthma populations. Our first dataset comprised baseline data from patients with asthma (n = 184) recruited from primary-care practices for two prospective clinical studies at our center: the GLAD (GPIAG [General Practitioners in Asthma Group] and Leicester Asthma and Dysfunctional Breathing) study (n = 70) (trial number ISRCTN 47153522) and the recently completed Intensive Asthma Study (n = 114) (13). The studies shared common subject selection criteria and recruitment techniques.

Our second dataset (n = 187) comprised data from patients with a diagnosis of refractory asthma, made in accordance with American Thoracic Society (ATS) criteria (10) by a respiratory physician with a specialist interest in this field. All the patients attended our specialist Glenfield Hospital refractory asthma clinic (Leicester, UK) for assessment and management of their asthma. The analysis was performed on consecutive patients attending the clinic between 2004 and 2006, with a full complement of data collected as part of their routine baseline assessment during their first visit to our center. The systematic recording and validation of data for some etiologic factors such as nasal polyps, aspirin sensitivity, and ethnicity are not routinely performed at our center. These data were therefore not available as part of the analysis. However, to be representative of the secondary-care asthma population, we chose to include all patients meeting ATS criteria for refractory asthma. Thus, patients in whom nonadherence with therapy was likely to have been a major determinant were not excluded. This is in contrast to our third population (described *below*) who were recruited to a clinical trial in which suspected or documented therapy nonadherence was an exclusion criterion of the study.

The third dataset comprised baseline and longitudinal data collected from a prospective clinical study (11). The study compared severe exacerbation frequency over 12 months in 74 patients with predominantly refractory asthma managed according to regular monitoring of airway inflammation using induced sputum (sputum arm), with the aim of titrating steroid therapy to maintain normal eosinophil counts, or standard clinical care (clinical arm). Sufficient baseline data were available in 68 of the 74 study participants to perform cluster analysis. Fifty-nine of the 68 patients (86.7%) met ATS criteria for refractory asthma.

Cluster Analysis Methodology

Uniform cluster analysis methodology was applied to each population using a two-step approach. In the first step, hierarchical cluster analysis using Ward's method generated a dendrogram for estimation of the number of likely clusters within the studied population. This estimate was prespecified in a k-means cluster analysis that was used as the principal clustering technique (14). Variables chosen for cluster modeling were selected on the basis of their considered contribution to characterizing the asthma phenotype. Variable selection and cluster analysis methodology are discussed further in the online supplement.

All measurements were standardized using *z* scores for continuous variables and 0 or 1 for categorical variables. Continuous variables were log transformed to approximate a normal distribution where this was indicated. Discriminant function analysis was performed using both forward and backward stepwise algorithms on each cluster model to evaluate the input variables that were significant determinants of model structure. This is discussed in greater detail in the online supplement.

Statistical Methods

The between-cluster comparison of baseline parameters that were not input parameters was performed using one-way analysis of variance (ANOVA) for parametric variables, the χ^2 test for proportions, and Kruskal-Wallis for nonparametric variables. For the analysis of outcome data in the prospective study, our clustering algorithm was applied to the baseline study data, and outcomes were compared between study arms for each cluster using the independent *t* test. Univariate ANOVA with the cluster model as a covariate was performed to verify the significance of this as an independent factor for any observed differences in outcome (*see* the online supplement). The measured outcomes were prespecified and included the frequency of severe exacerbations, measured as the number of rescue courses of oral corticosteroid and the change in corticosteroid dose at 12 months. All statistical analyses were performed using SPSS version 14 (SPSS, Inc., Chicago, IL). In addition, STATA (Version 7.0; Stata Corp., College Station, TX) was used to perform repetitions of cluster models with the k-means algorithm for demonstrating repeatability.

Approval from the local research ethics committee was obtained for data analysis and publication following informed consent for the respective clinical studies and as part of a clinical database for patients attending the Glenfield Hospital Difficult Asthma Clinic.

RESULTS

Compared with our secondary-care, refractory asthma population, the primary-care population had milder disease with significantly fewer symptoms, less airway dysfunction, and lower levels of eosinophilic airway inflammation, while taking a significantly lower mean dose of inhaled corticosteroids (Table 1).

The cluster structure described for each population was reproducible when repeating the algorithm using STATA and within randomly selected subsets of each population (data not shown). Statistical validity for the results was supported by identifying similar clusters of refractory asthma within the independent study cohort of Green and colleagues (11).

A three-cluster model best fit the primary-care population dataset (Table 2; Figure 1). Cluster 1 described a subgroup with early-onset atopic asthma. This cluster had evidence of airway dysfunction, symptoms, and eosinophilic airway inflammation. Clinically, this cohort was associated with a significantly greater number of previous hospital attendances and asthma exacerbations requiring oral corticosteroids when compared with the other primary-care subgroups. Cluster 2 described an obese subgroup with a female preponderance, evidence of asthma symptoms, and an absence of eosinophilic airway inflammation. The third cluster was labeled benign asthma because cases within this subgroup had little evidence of active disease. Asthma symptoms, airway inflammation, and measures of airway dysfunction were frequently within normal limits, and 58% of this cohort did not have evidence of significant airway hyperresponsiveness at the time of assessment. Consistent with a milder disease profile, patients from this cluster had very low rates of hospital attendance for asthma and severe exacerbation frequency in the previous 12 months (Table 2).

We identified four clusters in the secondary-care, refractory asthma population (Table 3; Figure 1). Clusters 1 and 2 had a profile that closely resembled the respective clusters in primary care. Thus, early-onset atopic asthma and obese, noneosinophilic asthma were common to asthma populations across the spectrum of severity. The principal distinction between the clusters in each population was the difference in absolute values of different objective measures of disease severity. In comparison with primary care, early-onset atopic asthma in secondary care exhibited greater airway dysfunction, symptoms, and eosinophilic airway inflammation on a higher dose of corticosteroid therapy. However, the pattern of expression of these variables, demographic data, and measures of asthma control were consistent between clusters of the two populations. The sub-population of this phenotype with refractory asthma had a significantly higher rate of failed attendance of appointments in the 12 months after referral to the clinic compared with the other phenotypes of refractory asthma (Table 3).

Clusters 3 and 4 were specific to the refractory asthma population and both exhibited marked dissociation between eosinophilic inflammation and asthma symptoms. Cluster 3 described an early-onset, symptom-predominant group with minimal eosinophilic disease. Cluster 4 described an eosinophilic inflammation–predominant group with few symptoms, late-onset disease, and a greater proportion of males.

Discriminant function modeling identified the majority of input parameters used in the cluster analysis of both populations to be significant determinants of cluster membership (Table E1 of the online supplement). The discriminant function model of primary-care and refractory asthma clusters required seven of eight input parameters (excluding atopic status) and five of seven parameters (excluding atopic status and sex), respectively. The accuracy of the discriminant function models for predicting cluster membership was 94.6% (primary care) and 96.8% (refractory asthma).

Cluster analysis was performed from baseline data in 68 patients of the prospective study dataset. Three clusters were identified (Table E2); all were comparable with clusters observed in the larger refractory asthma population. The original study demonstrated a

significant reduction in severe exacerbation frequency in the sputum arm, with no significant difference in corticosteroid usage between the groups. The present cluster-specific analysis revealed that all of the benefit for preventing exacerbations occurred in the inflammation-predominant cohort (3.53 [SD, 1.18] vs. 0.38 [SD, 0.13] exacerbation/patient/yr, $P = 0.002$) (Table 4). In addition, sputum-guided therapy allowed successful downtitration of corticosteroid therapy in early symptom-predominant asthma (Table 4; mean difference, 1,829 μg beclomethasone equivalent/d [95% confidence interval, 307–349 μg]; $P = 0.02$), without compromising asthma control. A univariate ANOVA with the cluster model as a covariate identified both treatment grouping and the cluster model as significant determinants for observed differences in exacerbation frequency ($P = 0.002$, study groups; $P = 0.03$, cluster model), but only the cluster model was a significant determinant for differences in inhaled corticosteroid dose ($P = 0.07$ for treatment groups and $P = 0.005$ for cluster model).

DISCUSSION

The need for classifying asthma heterogeneity has gained urgency with the parallel development of better tools for measuring disease characteristics that highlight disparity in clinical, physiologic, and pathologic markers, together with novel and specific molecular therapies that are only likely to be efficacious in particular subgroups of asthma. This study is the first to apply principles of cluster analysis for the identification of clinical asthma phenotypes. We have further shown that phenotypes constructed in this way exhibit clinically relevant differences in outcome, with management strategies that use a measure of eosinophilic inflammation for titrating corticosteroid therapy.

Asthma classification is complicated by the multidimensional nature of the disease. This prompted our consideration of cluster analysis techniques for this purpose. We selected the k-means clustering algorithm as it maximizes separation between clusters, thereby offering the greatest scope for identifying distinct groups within the population. Both familiar and previously uncharacterized asthma subgroups were identified that are more representative of multidimensionality. The identification of early-onset atopic asthma, an established asthma phenotype, validates the method for identifying the other subgroups against an accepted reference (15). Discriminant function analysis demonstrated the majority of the clustering parameters to be significant for cluster modeling, supporting multidimensionality. Atopic status was not identified as a significant discriminator influencing cluster membership in either primary care or secondary care. However, the prevalence of atopy did differ significantly between clusters and its inclusion to describe the phenotypes is therefore appropriate.

We chose to consider the two asthma population datasets independently when performing cluster analysis. This enabled clearer identification of factors that are specifically associated with refractory asthma, a condition that is sufficiently disparate to be considered a distinct disease entity by several authors (16).

The early-onset atopic asthma phenotype was common to both asthma populations, differing only in the severity of disease expression. We identified significantly higher rates of

nonattendance for clinic appointments in the refractory subgroup, which has been associated with poorer therapeutic compliance (17). Our finding of uncontrolled eosinophilic airway inflammation was in keeping with this. Our failure to identify the same phenotype in the recruited prospective study cohort may be because poor compliance was an exclusion criterion for the study. Although equivalent measures of compliance were not obtained in our primary-care population, it may be an important factor distinguishing this phenotype between the two populations. Strategies for improving compliance may therefore have a greater role in the management of this subgroup of refractory asthma. The obese, noneosinophilic phenotype common to both populations was characterized by symptoms that were not associated with eosinophilic airway inflammation. Given the recognized association between eosinophilic airway inflammation and steroid responsiveness in airway disease (18), the reported steroid resistance of asthma in obese patients (19) may in part be explained by the general pattern of airway inflammation seen with this phenotype.

The traditional paradigm of a direct relationship between eosinophilic inflammation and symptoms underpins present therapeutic guidelines that recommend symptom-led titration of corticosteroid therapy (20). Our analysis suggests that a symptom-led approach would be effective for mild to moderate asthma in primary care for patients with early-onset atopic asthma and benign asthma, where concordance was observed between inflammation and symptoms. However, discordance between these domains is a prevalent characteristic of refractory asthma and is also a feature of the obese, symptom-predominant, noneosinophilic phenotype seen in primary care (Figure 1). This may be a significant factor predisposing to failure with a conventional protocol and supports a role for measuring eosinophilic airway inflammation in these subgroups. For symptom-predominant phenotypes, the etiology of symptoms is multifactorial and not closely related to underlying eosinophilic airway inflammation. Overtreatment with corticosteroids may therefore occur. In keeping with this, a recent study using exhaled nitric oxide (F_{ENO}) as a measure of eosinophilic airway inflammation in asthma showed that F_{ENO}-guided management resulted in lower inhaled corticosteroid use without compromising asthma control (21). In contrast, the inflammation-predominant phenotype will be undertreated, leading to uncontrolled eosinophilic inflammation that is associated with a greater risk of future severe asthma exacerbations (22). Our hypothesis is supported by the results of the longitudinal cluster-specific analysis that demonstrated a 10-fold reduction in exacerbation frequency for this phenotype with a management strategy that measures eosinophilic airway inflammation to titrate therapy.

This study has several limitations. Principal among these is the cluster analysis methodology. Although we have used the k-means clustering algorithm, it is well recognized that populations of both disease and health have a continuous spectrum of expression. The use of an algorithm that separates the population into discrete clusters may not be realistic. Alternative clustering techniques that use a probabilistic approach for cluster structure and membership within a dataset may provide additional information and should be explored (23). Nevertheless, our analysis supports the hypothesis that subgroups of clinical relevance exist within asthma populations and can be revealed using cluster analysis. Despite our efforts to be objective, there were several areas of subjectivity, including our selection of variables for clustering and our decisions on the number of clusters for each population. Although our choice of clustering parameters was broad, we cannot exclude the

Europe PMC Funders Author Manuscripts

possibility that other variables may be of greater significance in developing meaningful phenotypes. In addition, the possible association between specific cluster profiles and well-recognized etiologic factors such as nasal polyps and aspirin sensitivity could not be explored. An advantage of multivariate techniques is that no single variable should be critical for determining the model. One of the drawbacks of using a nonhierarchical clustering technique is the need to prespecify the number of expected clusters. There are no well-validated techniques for predicting the number of clusters within a given population. We estimated this from dendrogram plots obtained using the hierarchical Ward's method. The study also does not address the question of stability in cluster membership over time and with changes in treatment. Within each population, there was no significant difference in treatment regimens and doses between clusters. Thus, differences observed between clusters may be considered a product of differences in the underlying disease profile together with differences in the response to therapy. These two factors are likely to be closely related. Although longitudinal change in cluster membership has not been explored, our analysis indicates cluster profiling at baseline is predictive of response to a management strategy prospectively for at least 12 months. It is also notable that four of the parameters we used for clustering (age of onset, sex, atopic status, and body mass index) are relatively invariant and not generally affected by time and therapy.

In summary, this study supports a role for the use of multivariate techniques in the classification of asthma populations. Clinically important prognostic differences identified between the phenotypes within this model may provide a reliable framework for exploratory molecular and genetic studies, presently undermined by population heterogeneity.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Prof. M. Silverman for his comments and Prof. J. Thompson for his review and advice of the statistical methodology.

References

1. Holgate ST, Polosa R. The mechanisms, diagnosis, and management of severe asthma in adults. Lancet. 2006; 368:780–793. [PubMed: 16935689]
2. Wenzel SE, Schwartz LB, Langmack EL, Halliday JL, Trudeau JB, Gibbs RL, Chu HW. Evidence that severe asthma can be divided pathologically into two inflammatory subtypes with distinct physiologic and clinical characteristics. Am J Respir Crit Care Med. 1999; 160:1001–1008. [PubMed: 10471631]
3. Wardlaw AJ, Silverman M, Siva R, Pavord ID, Green R. Multidimensional phenotyping: towards a new taxonomy for airway disease. Clin Exp Allergy. 2005; 35:1254–1262. [PubMed: 16238783]
4. Wenzel SE. Asthma: defining of the persistent adult phenotypes. Lancet. 2006; 368:804–813. [PubMed: 16935691]
5. Heaney LG, Robinson DS. Severe asthma treatment: need for characterising patients. Lancet. 2005; 365:974–976. [PubMed: 15767000]
6. Everitt, BS.; Landau, S.; Leese, M. Cluster analysis. 4th ed. Arnold; London: 2001.
7. Hartigan JA. Clustering. Annu Rev Biophys Bioeng. 1973; 2:81–101. [PubMed: 4583660]

8. Sneath PH, Sokal RR. Numerical taxonomy. Nature. 1962; 193:855–860. [PubMed: 13914561]
9. Everitt BS, Landau S. The use of multivariate statistical methods in psychiatry. Stat Methods Med Res. 1998; 7:253–277. [PubMed: 9803525]
10. American Thoracic Society Refractory Asthma Workshop Committee. Proceedings of the ATS workshop on refractory asthma: current understanding, recommendations and unanswered questions. Am J Respir Crit Care Med. 2000; 162:2341–2351. [PubMed: 11112161]
11. Green RH, Brightling CE, McKenna S, Hargadon B, Parker D, Bradding P, et al. Asthma exacerbations and sputum eosinophil counts: a randomised controlled trial. Lancet. 2002; 360:1715–1721. [PubMed: 12480423]
12. Haldar P, Berry MA, Wardlaw AJ, Pavord ID, Green RH. Refractory asthma phenotypes and the response to sputum eosinophil guided therapy. Thorax. 2006; 61:ii31.
13. Shaw DE, Berry MA, Thomas M, Green RH, Brightling CE, Wardlaw AJ, Pavord ID. The use of exhaled nitric oxide to guide asthma management: a randomized controlled trial. Am J Respir Crit Care Med. 2007; 176:231–237. [PubMed: 17496226]
14. Ball GH, Hall DJ. A clustering technique for summarising multivariate data. Behav Sci. 1967; 12:153–155. [PubMed: 6030099]
15. Rackemann FM. Studies in asthma: a clinical survey of 1074 patients with asthma followed for 2 years. J Lab Clin Med. 1927; 12:1185–1197.
16. Wenzel SE, Busse WW. Severe asthma: lessons from the Severe Asthma Research Program. J Allergy Clin Immunol. 2007; 119:14–21. [PubMed: 17208583]
17. Smith JR, Mildenhall S, Noble M, Mugford M, Shepstone L, Harrison BD. Clinician-assessed poor compliance identifies adults with severe asthma who are at risk of adverse outcomes. J Asthma. 2005; 42:437–445. [PubMed: 16293538]
18. Pavord ID, Brightling CE, Woltmann G, Wardlaw AJ. Non-eosinophilic corticosteroid unresponsive asthma. Lancet. 1999; 353:2213–2214. [PubMed: 10392993]
19. Peters-Golden M, Swern A, Bird SS, Hustad CM, Grant E, Edelman JM. Influence of body mass index on the response to asthma controller agents. Eur Respir J. 2006; 27:495–503. [PubMed: 16507848]
20. British Thoracic Society. Scottish Intercollegiate Guidelines Network. British guideline on the management of asthma. Thorax. 2003; 58:i1–i94. [PubMed: 12653493]
21. Smith AD, Cowan JO, Brassett KP, Herbison GP, Taylor DR. Use of exhaled nitric oxide measurements to guide treatment in chronic asthma. N Engl J Med. 2005; 352:2163–2173. [PubMed: 15914548]
22. Jatakanon A, Lim S, Barnes PJ. Changes in sputum eosinophils predict loss of asthma control. Am J Respir Crit Care Med. 2000; 161:64–72. [PubMed: 10619799]
23. Fraley C, Raftery AE. How many clusters? Which clustering method? Answers via model-based cluster analysis. The Computer Journal. 1998; 41:578–588.
24. Vansteenkiste J, Rochette F, Demedts M. Diagnostic tests of hyperventilation syndrome. Eur Respir J. 1991; 4:393–399. [PubMed: 1855568]
25. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale: an updated literature review. J Psychosom Res. 2002; 52:69–77. [PubMed: 11832252]

Europe PMC Funders Author Manuscripts

AT A GLANCE COMMENTARY

Scientific Knowledge on the Subject

Although several models of asthma classification have been proposed, a system defining the phenotypes of clinical asthma that incorporate the different aspects of the disease has not been developed.

What This Study Adds to the Field

Cluster analysis may be used to classify patients with asthma into phenotypic groups that exhibit clinically relevant differences in outcome with a management strategy using a measure of eosinophilic inflammation for titrating corticosteroid therapy.

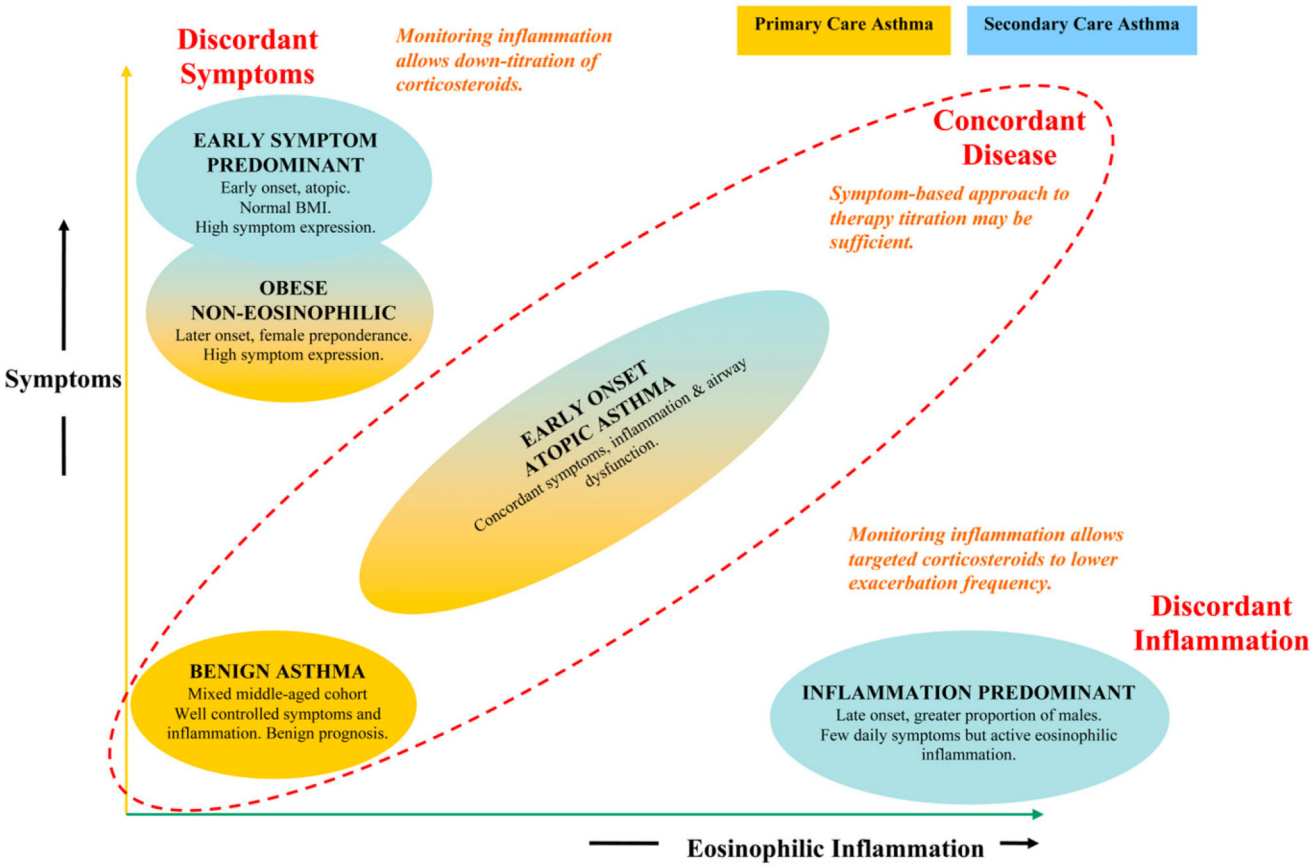

Figure 1. Clinical phenotypes of asthma.

A summary of phenotypes identified using cluster analysis in primary- and secondary-care asthma populations. The clusters are plotted according to their relative expression of symptoms and inflammation because these are the two clinically pertinent and modifiable dimensions of the disease. The plot highlights greater discordance to be a feature of secondary-care asthma. Although reasons for this dissociation are unclear, the use of measures of airway inflammation in these subgroups is clinically informative. BMI = body mass index.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

TABLE 1
Comparison of Baseline Characteristics in the three Asthma Populations

Variable	Primary Care (*n = 184*)	Secondary Care (*n = 187*)	Longitudinal Cohort (*n = 68*)	*P* Value*
Sex, % female	54.4	65.8	47.1	0.082
Age, yr (SD)	49.2 (13.9)	43.4 (15.9)	52.4 (14.6)	<0.001
Age of onset, yr (SD)	24.7 (19)	20.3 (18.4)	31.1 (23.7)	<0.001
Atopic status, % positive	72.8	73.8	57.4	0.365
Body mass index, kg/m^2 (SD)	27.5 (5.4)	28.5 (6.5)	28.0 (5.9)	0.55
PC_{20} methacholine†, mg/ml	1.04 (1.13)	†	0.67 (0.68)	0.19
Peak flow variability, amp % mean	17 (0.38)	32.2 (0.48)	13.8 (0.29)	<0.001
FEV_1 change with bronchodilator, %	1.63 (1.16)	12.8 (0.41)	3.2 (1.04)	<0.001
Post-bronchodilator FEV_1, % predicted	91.4 (21)	82.1 (21.1)	80.2 (20.6)	0.013
Sputum eosinophil count, %	1.32 (0.62)	2.9 (0.99)	2.4 (0.81)	0.08
F_{ENO}‡, ppb	31.6 (0.33)	43 (0.32)	4.32 (0.64)‡	<0.001
Sputum neutrophil count, %	55.09 (0.31)	46.7 (0.32)	41.1 (0.35)	0.04
Modified JACS§ (SD)	1.36 (0.74)	2.02 (1.16)	1.42 (1.26)	<0.001
Dose of inhaled corticosteroid, BDP equivalent/μg (SD)	632 (579)	1,018 (539)	1,821 (1,239)	<0.001
Long-acting bronchodilator use, %	40.2	93	86.7	<0.001

Definition of abbreviations: amp = amplitude; BDP = beclomethasone dipropionate; JACS = Juniper Asthma Control Score; SD = standard deviation.

*Significance figures are derived using one-way analysis of variance between the three populations for continuous variables or χ^2 test for proportions.

†Bronchial challenge testing is not routinely performed in secondary care for refractory asthma. The comparison given is between the primary-care asthma population and the longitudinal study cohort.

‡F_{ENO} was measured using the NIOX (Aerocrine, Solna, Sweden) analyzer at 50 ml/second in the primary-care population and secondary-care population. The Logan (Logan Research, Ltd., Rochester, Kent, UK) analyzer was used at a flow rate of 250 ml/second in the longitudinal study cohort. A strong linear correlation of 0.97 exists between the two measurement protocols. The statistical comparison is between Fe_{no} levels in primary and secondary care using NIOX.

§The Juniper Asthma Control Score, modified to include the symptom domains only (*see* the online supplement).

TABLE 2

Clusters in Primary Care

Variable	Primary Care (*n* = *184*)	Cluster 1 Early-Onset Atopic Asthma (*n* = *61*)	Cluster 2 Obese Noneosinophilic (*n* = 27)	Cluster 3 Benign Asthma (*n* = 96)	Significance (*P* Value)*
Sex†, % female	**54.4**	45.9	81.5	52.1	0.006
Age, yr (SD)	**49.2 (13.9)**	44.5 (14.3)	53.9 (14)	50.8 (13)	0.003
Age of onset†, yr (SD)	**24.7 (19)**	14.6 (15.4)	35.3 (19.6)	28.2 (18.3)	<0.001
Atopic status†, % positive	**72.8**	95.1	51.9	64.6	<0.001
Body mass index†, kg/m^2 (SD)	**27.5 (5.4)**	26.1 (3.8)	36.2 (5.5)	26 (3.6)	<0.001
PC_{20} methacholine†‡, mg/ml	**1.04 (1.13)**	0.12 (0.86)	1.60 (0.93)	6.39 (0.75)	<0.001
PC_{20} >8 mg/ml, n (%)	**64 (34.7)**	2 (3.3)	6 (22.2)	56 (58.3)	<0.001
Peak flow variability†‡, amp % mean	**17 (0.38)**	20 (0.47)	21.9 (0.32)	14.8 (0.32)	0.039
FEV_1 change with bronchodilator‡, %	**1.63 (1.16)**	4.5 (0.91)	1.82 (1.16)	0.83 (1.22)	<0.001
Post-bronchodilator FEV_1, % predicted	**91.4 (21)**	86.9 (20.7)	91.5 (21.4)	94.2 (20.7)	0.107
Sputum eosinophil count†‡, %	**1.32 (0.62)**	3.75 (0.64)	1.55 (0.51)	0.65 (0.44)	<0.001
F_{ENO}‡§, ppb	**31.6 (0.33)**	57.5 (0.27)	25.8 (0.29)	22.8 (0.27)	<0.001
Sputum neutrophil count‡, %	**55.09 (0.31)**	45.87 (0.24)	72.71 (0.13)	57.56 (0.36)	0.038
Modified JACS† (SD)	**1.36 (0.74)**	1.54 (0.58)	2.06 (0.73)	1.04 (0.66)	<0.001
Dose of inhaled corticosteroid, BDP equivalent/μg (SD)	**632 (579)**	548 (559)	746 (611)	653 (581)	0.202
Long-acting bronchodilator use, %	**40.2**	34.4	48.2	41.7	0.442
Previous hospital admission or emergency attendance, no. per patient	**0.60 (1.57)**	1.04	0.26	0.20	0.037
Previous outpatient attendance, % attended	**15%**	22%	19%	6%	0.121
Severe asthma exacerbations (requiring oral corticosteroids) in past 12 mo, no. per patient	**1.25 (1.94)**	1.86 (0.32)	1.07 (0.32)	0.39 (0.18)	0.002

For definition of abbreviations, *see* Table 1.

Boldface type denotes population statistics. The column headed “Cluster 3” represents a cluster not observed in the secondary-care asthma population.

*Comparison between clusters using analysis of variance for continuous variables and χ^2 test for proportions. Significance values for variables included in the cluster analysis are a product of the cluster algorithm and are provided for illustrative purposes only.

†Variables included in the cluster analysis.

‡Geometric mean ($\log_{10}$ SD)

§Measured with NIOX at a flow rate of 50 ml/second.

TABLE 3

Clusters in Secondary Care

Variable	Secondary Care (*n* = *187*)	Cluster 1 Early Onset, Atopic (*n* = *74*)	Cluster 2 Obese, Noneosinophilic (*n* = *23*)	Cluster 3 Early Symptom Predominant (*n* = *22*)	Cluster 4 Inflammation Predominant (*n* = *68*)	Significance (*P* Value)*
Sex†, % female	**65.8**	75.7	87	68.2	47.1	<0.001
Age, yr (SD)	**43.4 (15.9)**	39.4 (15.7)	42.7 (11.1)	35.5 (15.5)	50.6 (15.1)	<0.001
Age of onset†, yr (SD)	**20.3 (18.4)**	12.7 (12.9)	15.4 (15.2)	12.6 (15)	32.6 (19.1)	<0.001
Atopic status†, % positive	**73.8**	83.8	65.2	81.8	63.2	0.024
Body mass index†, kg/m^2 (SD)	**28.5 (6.5)**	27.6 (4.5)	40.9 (6.5)	23.6 (3.1)	27 (3.9)	<0.001
Peak flow variability‡, amp % mean	**32.2 (0.48)**	46.1 (0.35)	21.2 (0.76)	24.2 (0.65)	27.6 (0.36)	0.002
FEV_1 change with bronchodilator‡, %	**12.8 (0.41)**	24.5 (0.31)	9.3 (0.35)	4.5 (0.33)	9.8 (0.34)	<0.001
Post-bronchodilator FEV_1, % predicted (SD)	**82.1 (21.1)**	79.0 (21.9)	79.0 (18.5)	79.5 (26.1)	87.2 (18.5)	0.093
Sputum eosinophil count†‡, %	**2.9 (0.99)**	4.2 (0.76)	1.3 (1.01)	0.1 (0.9)	8.4 (0.64)	<0.001
FE_{NO}‡§, ppb	**43 (0.32)**	51.2 (0.36)	24.2 (0.27)	22.6 (0.30)	53.1 (0.32)	<0.001
Sputum neutrophil count, %‡	**46.7 (0.32)**	45.4 (0.39)	49.3 (0.22)	51.3 (0.23)	45.9 (0.29)	0.892
Modified JACS† (SD)	**2.02 (1.16)**	2.63 (0.93)	2.37 (1.09)	2.11 (1.11)	1.21 (0.95)	<0.001
Dose of inhaled corticosteroid, BDP equivalent/μg (SD)	**1,018 (539)**	1,168 (578)	1,045 (590)	809 (396)	914 (479)	0.008
Long-acting bronchodilator use, %	**93.0**	91.9	95.4	90.9	94.1	0.999
Maintenance oral corticosteroid use, %	**31.7**	32.4	22.7	22.7	36.8	0.604
Median Nijmegen score (IQR) (% with score >23)‖	**16 (7–26.5)**	20.5 (12–30.25) (44.6)	23 (12–33) (52.2)	16.5 (4.5–27.5) (31.8)	9 (1–17) (19.1)	0.004
Median anxiety score (IQR) (% with score ≥11)‖	**7 (4–10)**	7.5 (4.75–10.25) (24.3)	8 (3–14) (34.8)	6 (3.75–8.25) (13.6)	6 (3–9) (19.1)	0.34
Median depression score (IQR) (% with score ≥11)‖	**4 (2–7)**	4.5 (2–8) (13.5)	5 (2–7) (4.3)	4 (2–7) (4.5)	3 (1–6) (7.4)	0.104
Courses of oral corticosteroids for asthma exacerbations, n/case/yr	**4.05 (2.33)**	4.62 (0.27)	3.90 (0.38)	3.57 (0.49)	3.43 (0.27)	0.02
Hospital admissions for asthma, n/case/yr	**1.54**	1.64	1.61	1.54	1.23	0.703
Failed clinic appointments, % total appointments to DAC/yr	20.0	26.2	15.7	19.0	14.8	0.027

Definition of abbreviations: amp = amplitude; BDP = beclomethasone diproprionate; DAC = difficult asthma clinic; IQR = interquartile range; JACS = Juniper Asthma Control Score; SD = standard deviation.

Anxiety and depression scores are obtained from the Hospital Anxiety and Depression Scale, a validated 14-point screening questionnaire. Scores of greater than 11 for either domain are suggestive of clinically important symptoms (25). Boldface type denotes population statistics. Columns headed “Cluster 3” and “Cluster 4” represent clusters not identified in the primary care asthma population.

*Comparison between clusters using analysis of variance for continuous variables and χ^2 test for proportions. As for the other tables, significance values for variables included in the cluster analysis are a product of the cluster algorithm and should not be further interpreted.

†Variables included in the cluster analysis.

‡Geometric mean (log_{10} SD).

§Measured with NIOX at a flow rate of 50 ml/second.

‖The Nijmegen score is obtained from responses to the Nijmegen questionnaire, a 16-point screening questionnaire for hyperventilation syndrome. Scores of greater than 23 are suggestive of clinically significant hyperventilation (24).

Europe PMC Funders Author Manuscripts

TABLE 4
Cluster Specific Outcomes for Longitudinal Study

Cluster	Outcomes	Study Group		Significance
		Clinical (*n* = *10*)	**Sputum (*n* = *8*)**	
1: Obese female	Δ Inhaled corticosteroid dose*/μg per day (SEM)	−400 (328)	−462 (271)	0.89
	Severe exacerbation frequency over 12 mo (SEM)	1.40 (0.78)	1.50 (0.80)	0.93
	Number commenced on oral corticosteroids	2	1	0.59
		Clinical (*n* = *15*)	Sputum (*n* = *24*)	
2: Inflammation predominant	Δ Inhaled corticosteroid dose*/μg per day (SEM)	+753 (334)	+241 (233)	0.22
	Severe exacerbation frequency over 12 mo (SEM)	3.53 (1.18)	0.38 (0.13)	0.002
	Number commenced on oral corticosteroids	2	9	0.17
		Clinical (*n* = *7*)	Sputum (*n* = *4*)	
3: Early symptom predominant	Δ Inhaled corticosteroid dose*/μg per day (SEM)	+1,429 (429)	−400 (469)	0.022
	Severe exacerbation frequency over 12 mo (SEM)	5.43 (1.90)	2.50 (0.87)	0.198
	Number commenced on oral corticosteroids	6	0	Undefined

A comparison of prespecified asthma outcomes between the two management protocols analyzed according to cluster allocation of subjects at study entry.

* Expressed as equivalent dose of beclomethasone.