ERJ Express. Published on May 26, 2011 as doi: 10.1183/09031936.00102110

PATTERNS OF AIRWAY DISEASE AND THE CLINICAL DIAGNOSIS OF ASTHMA IN THE BUSSELTON POPULATION

AW (Bill) Musk[1,2,4], Matthew Knuiman[1,2], Michael Hunter[1], Jennie Hui[1,3,5,6], Lyle Palmer[1,5], John Beilby[3], Mark Divitini[2], Siobhain Mulrennan[4], Alan James[1,4]

(1) Busselton Population Medical Research Foundation, Sir Charles Gairdner Hospital, Nedlands, Western Australia, 6009

(2) School of Population Health, University of Western Australia, Nedlands, Western Australia, 6009

(3) PathWest, Queen Elizabeth II Medical Centre, Nedlands, Western Australia, 6009

(4) Sir Charles Gairdner Hospital, Nedlands, Western Australia, 6009

(5) Centre for Genetics, Epidemiology and Biostatistics, UWA, Crawley, Western Australia 6009.

(6) School of Pathology and Laboratory Medicine, University of Western Australia, Nedlands, Western Australia

ABSTRACT

Introduction: Common airway diseases are phenotypically heterogeneous syndromes which do not fit closely into the conventional diagnostic categories of asthma and COPD. Identification of more homogeneous phenotypes of airflow limitation may lead to more appropriate treatments and better understanding of their pathophysiology and genetic basis.

Aim: To examine how objective measures related to lung function cluster in the general population and how the resulting patterns relate to "asthma" and "bronchitis" as diagnosed by a doctor, recent wheezing, cough and sputum production and tobacco smoking.

Subjects: Age-stratified random general population sample of 1,969 adults identified from the electoral register of Busselton, Western Australia.

Methods: Cross-sectional survey 2005-7 comprising questionnaire on respiratory symptoms, doctor-diagnosed asthma (DDA)-ever, doctor-diagnosed bronchitis (DDB)-ever, recent wheezing and smoking history together with anthropometric measurements, spirometry (FEV1, FVC), methacholine challenge or bronchodilator response to diagnose airway hyper-responsiveness (AHR), exhaled nitric oxide (eNO), prick skin tests to common allergens, and peripheral blood eosinophil and neutrophil counts. Cluster analysis (SAS version 9.2) with variables sex, age, atopy, FEV1 %predicted, FEV1/FVC, AHR, eNO, log eosinphil count, log neutrophil count and body mass index (BMI) was used to identify phenotypic patterns.

Results: Seven clusters (number of subjects, % with DDA, % with recent wheeze, % with DDB) were identified: "normal males" (467, 7%, 15%, 13%), "normal females" (477, 12%, 13%, 18%), "obese females" (250, 16%, 26%, 28%), "atopic younger adults" (330, 21%, 27%, 17%), "atopic adults with high eNO" (130, 30%, 34%, 25%), "atopic males with reduced FEV1" (103, 33%, 54%, 32%), and "atopic adults with BHR" (212, 40%, 38%, 26%). Overall 51% of subjects had never smoked but this varied from 20% to 67% across clusters.

Conclusion: The clinical diagnosis of asthma (ever) (and also recent wheeze and bronchitis-ever) is not specific for any of the clustering patterns of airway abnormality in the general population underlining the heterogeneity of the entity of “asthma” clinically and for understanding the genetic basis of the asthma syndromes.

INTRODUCTION

Since the International Ciba Guest Symposium in 1969 (Ciba 1969) and the statement of the American Thoracic Society (1972), asthma has generally been defined in physiological terms as variable airflow obstruction, emphysema has been defined in morphological terms and bronchitis in clinical terms. The conventional clinical entities of asthma and chronic obstructive pulmonary disease (COPD) are heterogeneous conditions with a large degree of overlap and encompass virtually all patients with obstructive spirometry who have no other demonstrated specific airway disease processes. This heterogeneity exists with respect to clinical, physiological and pathological factors and is observed both between individuals and within individuals over the course of their lives. The conditions are complex disorders thought to be due to the interaction of multiple genetic and environmental factors whose relative contributions to disease expression are likely to vary with age and between different populations and different geographical locations. This degree of heterogeneity has led to the idea that multiple distinct patho-physiological entities underlie the clinical entities [1-3].

The conflation of diverse processes into clinical labels may have contributed to the difficulties that have been encountered in understanding these diseases, especially studies of genetic epidemiology [4] and of novel treatments that may only target one molecular pathway and therefore only be effective in those patients with a particular sub-phenotype [5]

A variety of asthma phenotypes have been recognized [6-8] and it has been suggested that better definitions of asthma sub-types should form the basis of treatment options for patients that have, in the past, been classified mainly by clinical severity [9]. This approach is yet to be widely adopted for clinical use, since accurate characterisation of different phenotypes is

difficult and not standardised. Similar concerns regarding the diagnosis and management of COPD have been expressed [10].

In an effort to improve phenotype characterisation, statistical analyses of patient populations have been explored [3, 11]. This approach has been used in patients diagnosed with asthma or with COPD [9, 12-15]. To a large extent these populations have already been selected based on clinical severity, clinical setting or physiological abnormality based on lung function testing. Such selection introduces bias (such as diagnostic and recall bias) in the definition of phenotypes and ignores the crossover of symptoms and abnormalities of airway function that occur between various airway diseases.

The aim of the present study was to examine the clustering of objective measures of airway disease in the general population and determine how well these patterns relate to the conventional diagnostic entities of asthma and bronchitis (as diagnosed by a doctor), tobacco smoking and respiratory symptoms. The research hypothesis was that objective measures of measurable airway-related attributes in the general population define discrete phenotypic clusters all of which include "doctor-diagnosed asthma" and are more homogeneous than conventional diagnostic labels of asthma, chronic obstructive lung disease (COPD) and bronchitis.

SUBJECTS

An age-stratified random sample of 2,932 adults (63% of those contacted) identified from the electoral register (registration to vote is compulsory in Australia) for the district of Busselton in Western Australia participated in a prevalence study of respiratory disease conducted between July 2005 and July 2007. Due to missing data on some clinical measurement

variables (each variable had 100 to 300 missing values) there were a total of 1,979 that had all required data for this analysis. The participants were predominantly Caucasian with some having Asian ancestry, there being no Aboriginal people taking part in the survey.

METHODS

Written informed consent was obtained and the study was approved by the Human Research Ethics Committee of the University of Western Australia. A mailed questionnaire included respiratory symptoms, smoking and previous illnesses based on the British Medical Research Council Respiratory Questionnaire [16]. The diagnosis of asthma was taken from an affirmative response to the question: "Has a doctor ever told you that you have asthma?" (doctor-diagnosed asthma – DDA-ever). Doctor-diagnosed bronchitis (DDB-ever) was defined as a positive response to the question, "Has a doctor ever told you that you have bronchitis?" Recent wheeze was defined as a positive response to "Has your chest ever made a wheezing or whistling sound? If yes, in the last 12 months?". Height (cm) and weight (kg) were measured in stockinged feet without overgarments. Body mass index (BMI) was calculated as weight/height^2. Ever-smokers had smoked at least 1 cigarette per day or 50g of tobacco per month for at least 1 year. For analysis smoking history was scored as: never smoked, ex-smoker, current smoker and with the latter two groups split according to the pack-years of smoking (<5, > 5 pack-years) with a pack-year defined as smoking the equivalent of one pack (20 cigarettes) per day for a year. Cough and phlegm were recorded if subjects reported cough and phlegm production on most days for at least 3 months of the year.

Forced expiratory volume in 1 second (FEV_1) and forced vital capacity (FVC) were measured using Medgraphics CPFS/D USB pneumotach spirometers and recorded with BreezeSuite 6.2 software (Medical Graphics Corporation, St Paul, USA). Spirometers were calibrated daily

using a 3 litre syringe and corrected for room temperature, humidity and barometric pressure. Forced expiratory manoeuvres were performed with the subject seated and wearing a nose clip and were repeated up to a maximum of 8 efforts until 3 readings of FEV_1 and FVC, reproducible to at least 200 ml [17] were obtained, although attempts were made to achieve reproducibility of less than 75 ml [18]. The highest values for FEV_1 and FVC were recorded. Best FEV1 and FVC were expressed as percentages of the predicted values [19].

Airway responsiveness was measured by calculating the dose of inhaled methacholine chloride that provoked a 20% fall in FEV_1 (PD_{20}). Participants inhaled normal saline and then increasing doses of methacholine (0.03-4 mg/ml) via a hand-held dosimeter-driven nebuliser (Model 45, DeVilbiss, USA) and FEV_1 was measured at 30 and 90 seconds. The test was stopped if there was a >20% fall in FEV_1 from the post-saline FEV_1, a maximal provocation concentration of 4 μmol was delivered or the subject was unwilling to continue. Participants were excluded from the methacholine challenge if they were pregnant or breast-feeding, were taking beta-adrenergic blocker medication or had used a long acting beta-2 agonist aerosol in the preceding 48 hours. Because of the risk of causing severe/unpleasant symptoms participants with an initial FEV_1 of less than 60% of the predicted normal value did not receive methacholine but were given salbutamol (200 μg) administered from a metered dose inhaler via a spacer and spirometry was repeated after 15 minutes. Airway hyper-responsiveness (AHR) was defined as a PD_{20} of less than 4 μmol of methacholine or more than a 15% increase in FEV_1 following salbutamol as both have been shown to similarly define hyper-responsiveness [20].

Forearm skin prick responses to inhaled common allergens (rye grass pollen, grass mix, house dust mite (Dermatophagoides pteronyssinus and D. farinae), cat pelt, dog dander, Alternaria

tenius, Aspergillus fumigatus, and Mould mix (Hollister-Stier, USA)) were assessed. Histamine (10mg/ml) and glycerine were used as positive and negative controls, respectively. Weal size (average of the long axis and its perpendicular in mm) was measured after 10 minutes for the positive control and after 15 minutes for the allergens. Atopy was defined as any allergen weal of $\geq$ 3 mm greater than the negative control. Exhaled nitric oxide (eNO) was measured prior to spirometry and methacholine challenge using a calibrated chemiluminescence analyser with on-line measurement of single exhalations (Niox, Aerocrine, Solna, Sweden) according to a standard protocol [21]. Absolute neutrophil (neut) and eosinophil (eosin) counts were performed on a venous blood sample collected at the time of the survey.

K-means cluster analysis (SAS Version 9.2, SAS Institute Inc Carey NC) was used to divide participants into homogeneous groups. The quantitative variables in the cluster analysis were: sex, age, atopic status, BHR, FEV1%pred and FEV1/FVC, eNO, blood eosinophils, blood neutrophils and body mass index. The logarithm of the neutrophil and eosinophil counts were used in the analysis as they showed skewed distributions. Initially a cluster model with 40 clusters was fitted to the data to identify outliers (subjects who were allocated to a cluster of less than 5 people) and after removing 10 outlying subjects there were a total of 1,969 subjects in the main cluster analysis. The number of clusters was based on examining the change in the ratio of the between- to the within-cluster variation as more clusters were added, consideration of the size of the resulting clusters (at least 100 subjects) and the stability and consistency of the cluster patterns when analysed separately for gender and smoking groups. Based on these considerations the 7 cluster model emerged as most appropriate. For the standardised clustering variables the total standard deviation was 1.0 (ie for the whole sample as a single cluster) and the within cluster standard deviation declined steadily from 0.94 for

the two cluster model to 0.79 for the 7 cluster model, and thereafter declined more slowly as further the number of clusters was increased. The independent variables selected for analysis were chosen to reflect putative mechanisms and measurable clinical characteristics associated with common airway disease phenotypes avoiding those that tend to measure the same thing [14]. It was not possible to determine if some variables represented epiphenomena given the poor understanding of the fundamental mechanisms of airway diseases.

RESULTS

About half of the participants were female (table 1). The average age was 54 years, the average BMI was 27 kg/m^2, about half were atopic, 12% had BHR, 51% had never smoked, 38% were ex-smokers (with two-thirds having smoked more than 5 pack-years) and 12% were current smokers (with three-quarters having already smoked more than 5 pack-years), 18% reported doctor-diagnosed asthma, 24% reported recent wheeze, and 20% reported doctor-diagnosed bronchitis.

The 7 clusters that emerged appeared naturally to fit groups recognizable in the general population (Table 1). Cluster 1 (labelled “normal males”) with 467 subjects consisted of males with normal levels of all measured attributes of whom 7% had ever been told by a doctor that they had asthma, 15% had recent wheeze, and 13% bronchitis. Cluster 2 (labelled “normal females”) with 477 subjects consisted of females with normal levels of all measured attributes of whom 12% had been told by a doctor that they had ever had asthma, 13% had recent wheeze, and 18% bronchitis. Cluster 3 (labelled “obese females”) with 250 subjects were mainly overweight females (mean BMI 33.7, 86% female) with near-normal measured attributes of whom 16% had ever been told by a doctor that they had asthma, 26% had recent

wheeze, and 28% bronchitis. Cluster 4 (labelled "atopic younger") with 330 subjects comprised mainly younger adults (mean age 40 years) who were atopic (70%) but otherwise had normal measured attributes and 21% had ever been told by a doctor that they had asthma, 27% had recent wheeze, and 17% bronchitis. Cluster 5 (labelled "atopic with high eNO") with 130 subjects was characterized primarily by being atopic (86%) with elevated levels of eNO (mean 56.1) but also with a higher prevalence of BHR (16%) and eosinophlia of whom 30% had ever been diagnosed with asthma by a doctor, 34% had recent wheeze, and 25% with bronchitis. Cluster 6 (labelled "atopic males with poor FEV1") with 103 subjects contained mainly older males (83% male, mean age 70 years), 80% of whom had ever smoked and whose lung function was impaired (mean FEV1% predicted 66.7%) with 33% having been told by a doctor that they had asthma, 54% had recent wheeze, and 32% having been told that they have bronchitis. Cluster 7 (labelled "atopic with BHR") with 212 subjects comprised exclusively adults with BHR (100%) and with a high prevalence of atopy (70%) but near normal lung function and 40% of whom had been told by a doctor that they had asthma, 38% had recent wheeze, and 26% bronchitis.

Although we have assigned labels to clusters based on the principal distinguishing variables the other variables are also an integral part of the definition of the clusters as every cluster variable was significantly different across the 7 clusters ($p<0.0001$) and was still significantly different ($p<0.0001$) across the subset of clusters which did not have that variable as a principal distinguishing variable. To illustrate, BMI is a principal distinguishing variable for cluster 3 ('obese females') and BMI still varies significantly across the other 6 clusters.

The prevalence rates of doctor diagnosed asthma were not evenly distributed across the groups ($p<0.0001$) with Clusters 1 and 2 significantly less than Clusters 3 and 4 which were

also significantly less than clusters 5,6 and 7. The prevalence of recent wheeze also varied across clusters ($p<0.0001$) with a similar ordering to doctor diagnosed asthma with Clusters 1 and 2 having the lowest prevalence, clusters 3 and 4 next lowest, and 5,6 and 7 the highest prevalence. However, in contrast to doctor diagnosed asthma, cluster 6 had significantly higher prevalence of recent wheeze than clusters 5 and 6. Thus, it can be seen that doctor diagnosed asthma and recent wheeze did not segregate clearly with any particular cluster profile. The prevalence of cough/phlegm and doctor diagnosed bronchitis, and current smoking also varied significantly between groups (Table 1) but again did not segregate with any particular cluster. However, smoking (and in particular ever having smoked at least 5 pack-years) clearly segregated with Cluster 6 ($P<0.001$) in which one third reported being told by their doctor that they had asthma, a similar number had bronchitis, and over half had recent wheeze.

The prevalence rates of DDA and DDB appeared closely related in the various clusters (Figure 1) and the prevalence of DDA was also closely related to that of cough and phlegm (Figure 2). The same was true for recent wheeze (Figures 3 and 4). It is evident that there was a tendency for clusters that had a higher prevalence of diagnosed bronchitis and higher prevalence of cough/phlegm to have higher prevalence of diagnosed asthma and recent wheeze. These strong relationships suggest that none was clearly distinguishable from the others in a general population despite the extensive phenotyping that was attempted. The atopic clusters tended to have higher prevalence of all four conditions.

DISCUSSION

Previous studies of clustering of airway phenotypes have included patients with disease of clinical severity and have not comprised members of the general population. Recently, there

have been moves to dissect phenotypes of asthma and COPD that may have implications for treatment or causality. There is increasing evidence in adult patients with asthma that there are distinct groups based on inflammatory profiles of induced sputum and that these may respond to different forms of treatment [22, 23]. Using cluster analysis, Haldar et al. [9] examined asthmatics from an ambulatory care group and another from a hospital specialty clinic setting and found that both groups segregated according to inflammatory markers suggesting that it may be possible to identify groups of patients that respond to different forms of treatment. More recently, Moore et al. [24] examined patients from a severe asthma cohort using cluster analysis and identified 5 clusters based on atopy, asthma severity and health care utilisation. Notably, this study and that of Haldar et al. [9] identified a cluster of overweight women included in their asthma groups, similar to that seen in the present study. Weatherall et al. [15] used cluster analysis to examine 175 subjects with airflow obstruction from a general population sample and were able to identify 5 clusters based on degree of airflow obstruction, symptoms and markers of emphysema. Pillai et al. [13] examined features of asthma patients aged 7-35 years that could be useful in identifying important genetic and environmental contributors and found that baseline pulmonary function, allergen sensitization, self-reported allergies, symptoms of rhinitis and symptoms of asthma were important. Factors scored as quantitative traits appeared to be better phenotypes for epidemiological and genetic analyses than categories derived from the presence or absence of combinations of positive skin tests or elevated IgE. Patel et al. [25] showed that CT scans measures of airway wall thickness and emphysema independently contributed to airflow obstruction and had independent familial aggregations, suggesting that different inherited factors contribute to airway disease phenotypes. Thus these studies, undertaken in groups of patients defined by lung function, symptoms or diagnostic labels, show that several distinct phenotypes may exist within the framework of these “diagnostic labels”.

The present study was designed differently from other cluster analyses [1, 9, 10, 12] as the aim was to examine how airway disease-related physiological measures in a general population sample cluster without regard to clinical findings/disease status and to determine how the entities of doctor-diagnosed asthma or bronchitis and recent wheezing segregate within these clusters. It was therefore not constrained by preconceived notions of normality and abnormality or clinical diagnostic entities but relied on objective findings that are not dependent on clinical judgment. The results of the study demonstrate that asthma and bronchitis as diagnosed by a doctor are distributed across all seven clusters of airway-related phenotypes.

Cluster analysis was utilized for this study because it achieves segregation of people into groups with homogeneous disease manifestations and has recently been suggested as an appropriate method for determination of airway phenotypes [3]. The method used to find clusters (k-means cluster analysis) was the same as in Haldar [9]. However, other approaches such as hierarchical methods have also been suggested. For comparison purposes, we also used a hierarchical approach to best 7–cluster model and found that five of the cluster types from our k-means approach (1 = normal males, 2 = normal females and 5 = atopic with high eNO, 6 = atopic males with poor FEV, 7 = atopic with BHR) consistently appeared but with cluster types 3 (= obese females) and 4 = (atopic younger) sometimes slightly modified and with some mixing into 'normal' clusters 1 and 2. For example under Ward's hierarchical method, clusters 1 to 4 were normal males, normal females, atopic males, and atopic females. Thus the clinical or disease phenotypic clusters (5, 6, and 7) emerged consistently in both clustering approaches. Further the pattern of the prevalence of doctor-diagnosed asthma and recent wheeze was essentially unchanged across the cluster types. This suggests that the findings are robust against different statistical methods to find clusters.

We chose objective and quantitative indices related to airway behaviour in a general population sample of adults to examine the patterns of clustering. The use of relatively homogeneous clusters of airway patho-physiology rather than conventional clinical entities such as 'asthma' may assist future studies to identify aetiological factors that contribute to the specific pathways involved in these phenotypic clusters. The approach however cannot take environmental factors such as smoking or allergen exposure into account. This represents the difficulty of diagnosing asthma where the underlying pathology renders the subject susceptible to variable symptoms, probably related to environmental exposures. Therefore symptoms may vary while the "disease process" remains present. For this reason we examined the association both of a cumulative diagnosis, doctor-diagnosed asthma, and of current symptoms cough and phlegm and recent wheeze, with our cluster groups.

It is now about one hundred years since the entity of "cardiac asthma" was separated from "bronchial asthma" when it became possible to differentiate between heart failure and airway narrowing as the predominant underlying physiological abnormality in most patients with variable wheeze and breathlessness. Over the intervening years the adjective "bronchial" has been lost. However different clinical patterns of bronchial asthma as diagnosed by a doctor (ie DDA) have long been recognized including atopic and non-atopic asthma, childhood onset and adult onset asthma, reversible and irreversible asthma (as in asthma with COPD), neutrophilic and eosinophilic (and pauci-granulocytic) asthma [6], steroid-responsive and steroid-nonresponsive asthma indicating an acceptance that bronchial asthma itself is a heterogeneous disease. Similarly, COPD and its synonyms are terms of clinical convenience because the causes and management of the different patho-physiological varieties are similar despite the knowledge that emphysema with primary loss of pulmonary elastic recoil is a different entity from irreversible intrinsic airway narrowing as demonstrated by Colebatch et al [26] over 40 years ago. The process of phenotypic reclassification needs to be extended,

possibly with improved use of better/novel biomarkers [27] to redefine the processes of airflow obstruction further. Further dissection of airway phenotypes is needed to understand the molecular patho-physiology of airflow obstruction and allow coherent investigation of genotype-phenotype relationships. For example, a recent large genome-wide association study of doctor-diagnosed asthma has shown that genetic associations of adult and childhood onset asthma differ [28].

Additional phenotype analysis such as assays of interleukin-6 or TGF-beta which appear to segregate with acquired and innate immune responses may further advance the understanding of the airway phenotype-genotype relationships and may help to guide treatment with the advent of novel specific therapies in the future [29]. Similarly more precise differentiation of the underlying physiological abnormalities or the morphological information provided by (high resolution) computerized tomography could better guide therapy and genetic investigation (especially of COPD). This study of the general population confirms that the clinical entities of asthma and COPD are heterogeneous and suggests that the cluster approach may contribute towards disentangling the heterogeneity and might help to better identify specific risk factors and therapy.

ACKNOWLEDGEMENTS

The 2005-07 survey was funded by a project grant from the National Health and Medical Research Council (NHMRC) of Australia and the Sir Charles Gairdner Hospital Medical Research Foundation. The authors acknowledge the assistance of survey staff Elspeth Inglis and Ryan Fleay, local Busselton Health Study Action Committee, the many survey volunteers, and the long-standing support of the Busselton community for the surveys. A. James is supported by an NHMRC Practitioner Fellowship.

Table 1: Subject characteristics, overall and by cluster (phenotypic pattern)

Cluster	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**All**
	Normal males	Normal females	Obese females	Atopic younger	Atopic with high eNO	Atopic males with poor FEV	Atopic with BHR	
Total number	467	477	250	330	130	103	212	1,969
Clustering variables								
Sex (% female)	0.2	98.5	86.4	36.4	36.2	17.5	58.5	50.6
Age, yr (SD)	61 (15)	54 (17)	58 (15)	40 (13)	48 (19)	70 (12)	53 (18)	54 (17)
Atopic status, % positive	35.1	36.7	28.4	70.0	86.2	63.1	70.3	49.1
BHR, % positive	0.0	0.0	0.4	0.0	16.2	12.6	100.0	12.5
FEV1, % predicted	101.4	99.7	94.2	99.8	97.2	66.7	88.6	96.3
FEV1/FVC, %	76.4	79.2	80.3	81.1	76.8	56.1	74.4	77.1
eNO, ppb	19.5	15.8	16.9	16.0	56.1	22.1	23.4	20.7
Eosinophil count	0.17	0.15	0.22	0.27	0.31	0.29	0.24	0.21
Neutrophil count	3.06	2.93	3.82	4.43	3.16	4.33	3.57	3.48
BMI, kg/m^2	27.2	24.0	33.7	26.4	25.7	26.7	26.8	26.9
Non-clustering variables								
Dr Diagnosed Asthma, %	6.9	11.7	16.4	20.9	30.0	33.0	39.6	18.0
Recent wheeze, %	15.0	12.8	25.6	26.7	33.8	54.4	38.2	23.6
Cough/phlegm, %	22.7	15.7	20.0	25.8	22.3	38.8	27.4	22.5
Dr Diagnosed bronchitis, %	13.3	18.4	28.0	17.3	24.6	32.0	25.9	20.2
Never smoked, %	45.8	60.2	50.8	49.1	66.9	20.4	46.7	50.6
Ex-smoker, <5 packyears	11.3	16.1	9.6	10.6	8.5	7.8	12.7	11.9
Ex-smoker, > 5 packyears	36.0	17.2	26.0	16.4	20.0	58.3	24.5	25.7
Current, < 5 packyears	0.6	1.5	1.2	7.3	2.3	0.0	2.8	2.3
Current, > 5 packyears	6.2	5.0	12.4	16.7	2.3	13.6	13.2	9.3

Figure 1. Patterns in prevalence of doctor-diagnosed asthma and doctor-diagnosed bronchitis across clusters.

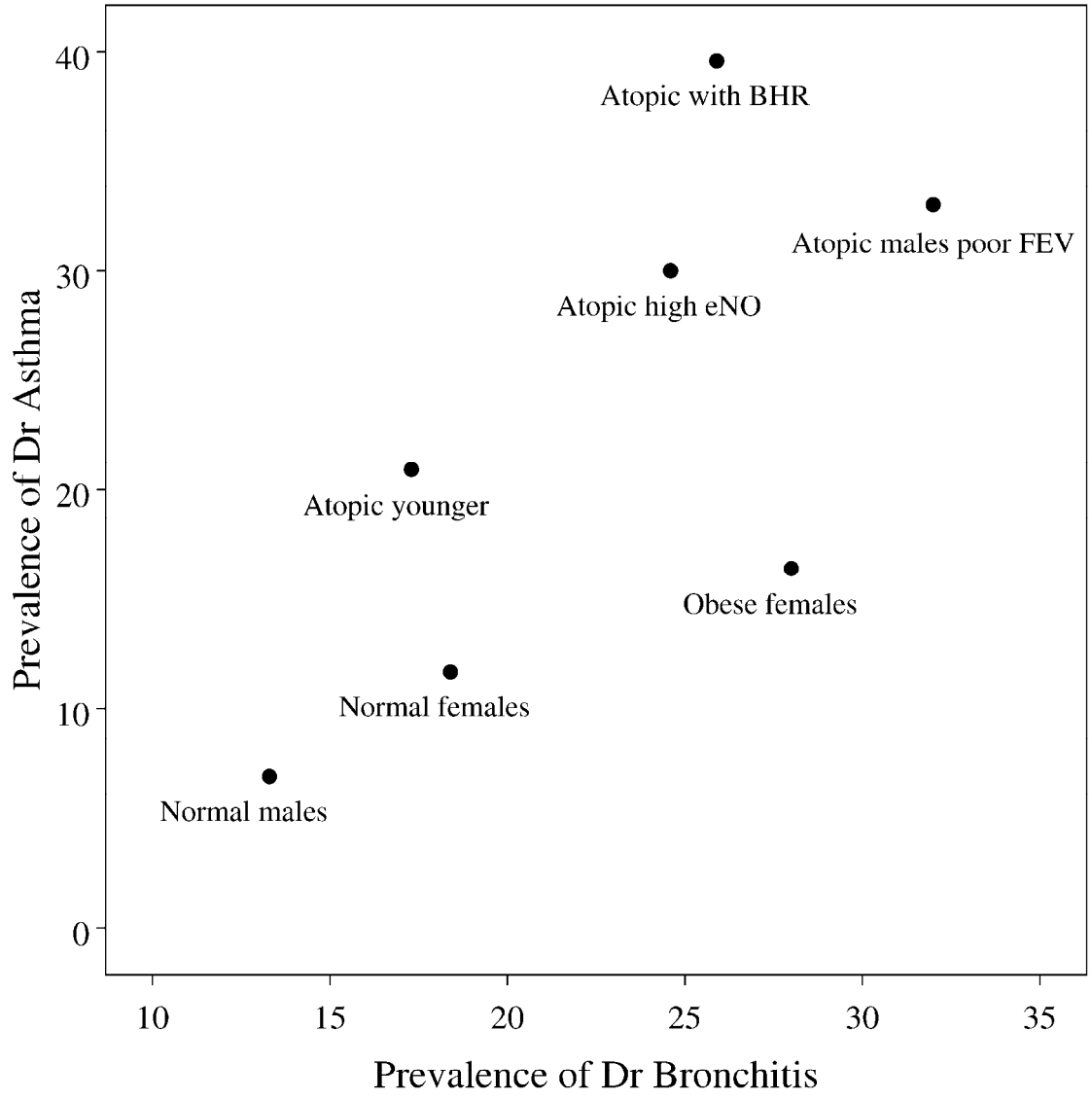


Figure 2. Patterns in prevalence of doctor-diagnosed asthma and cough/phlegm across clusters.

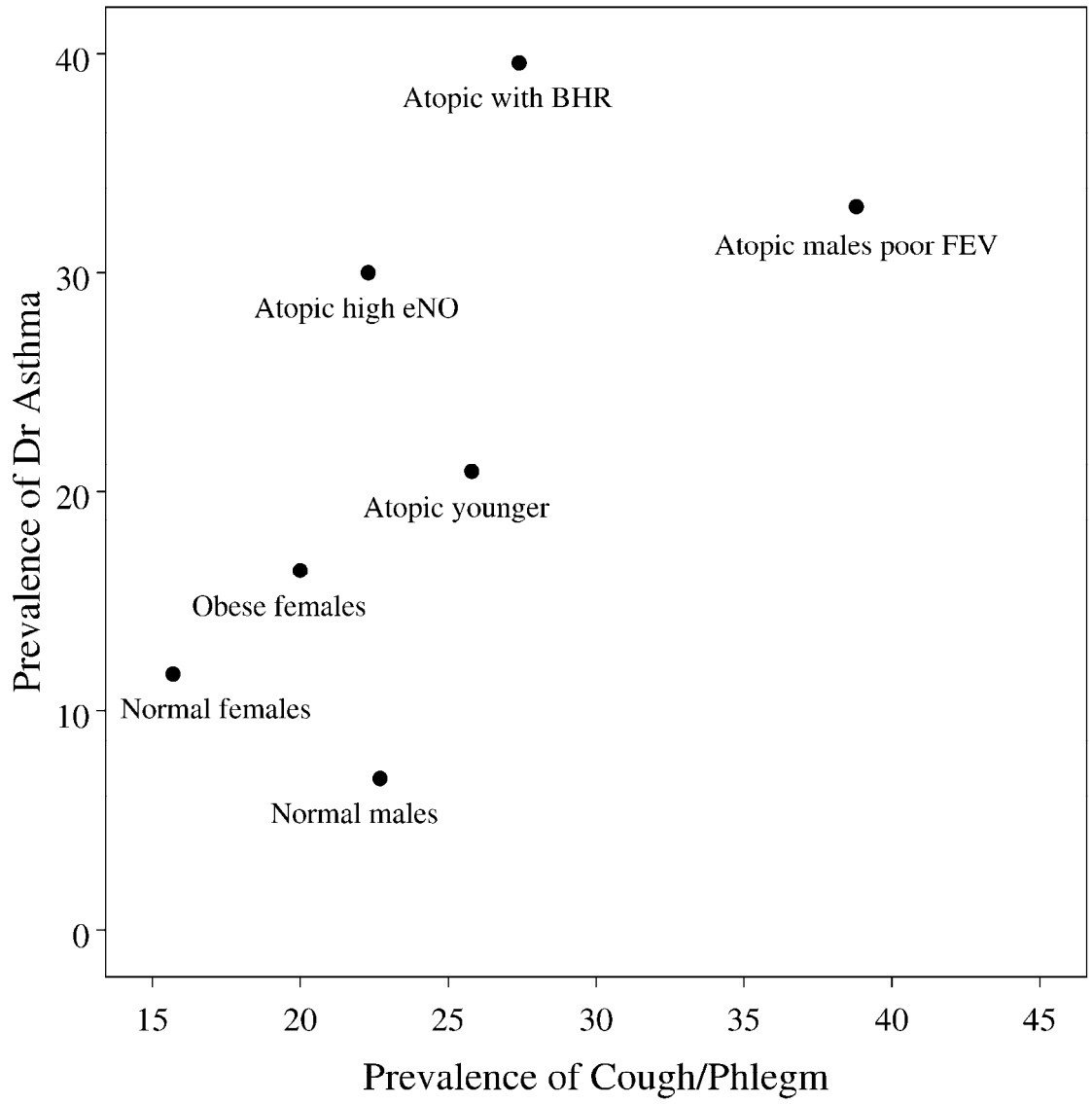


Figure 3. Patterns in prevalence of recent wheeze and doctor-diagnosed bronchitis across clusters.

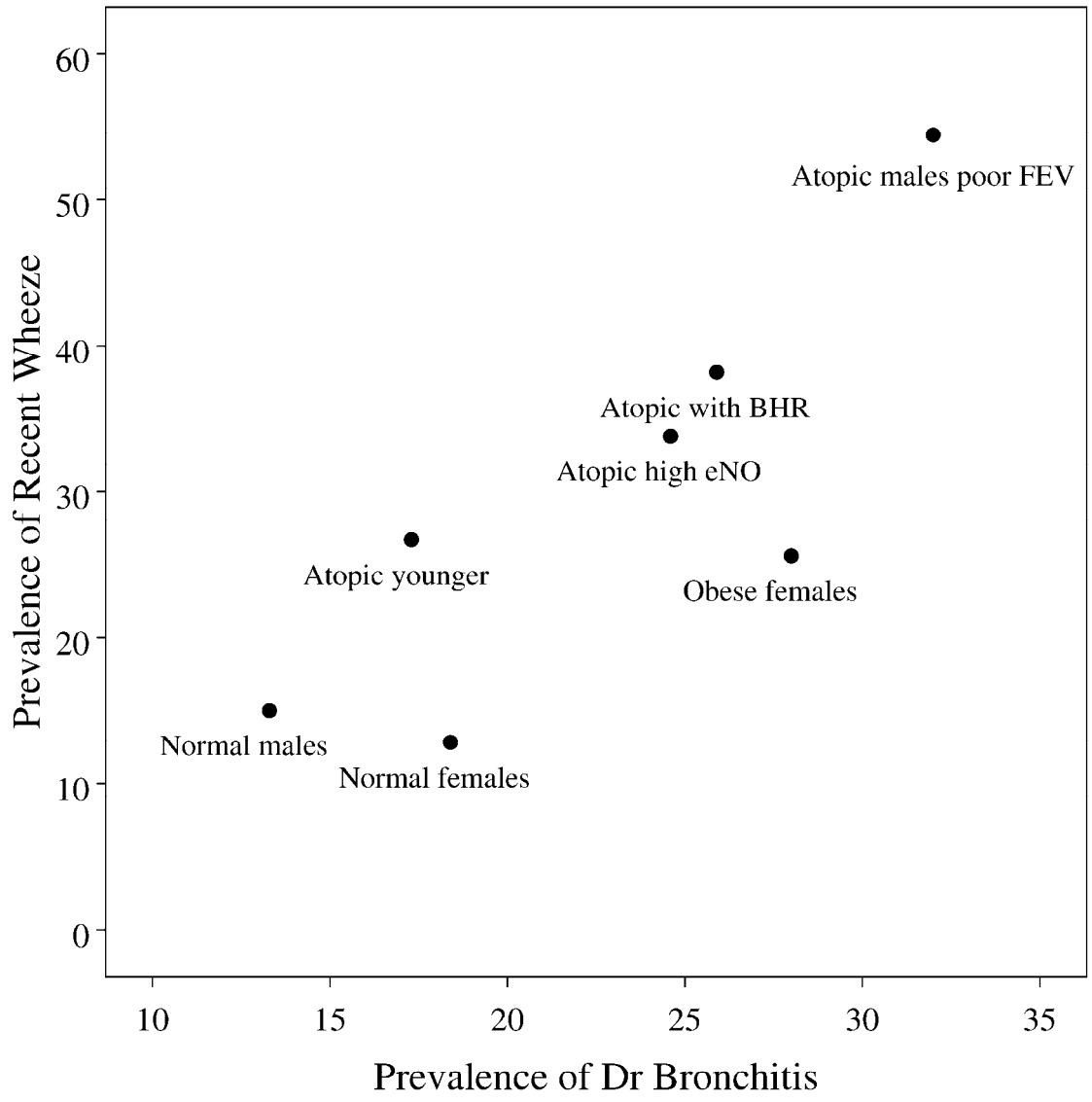
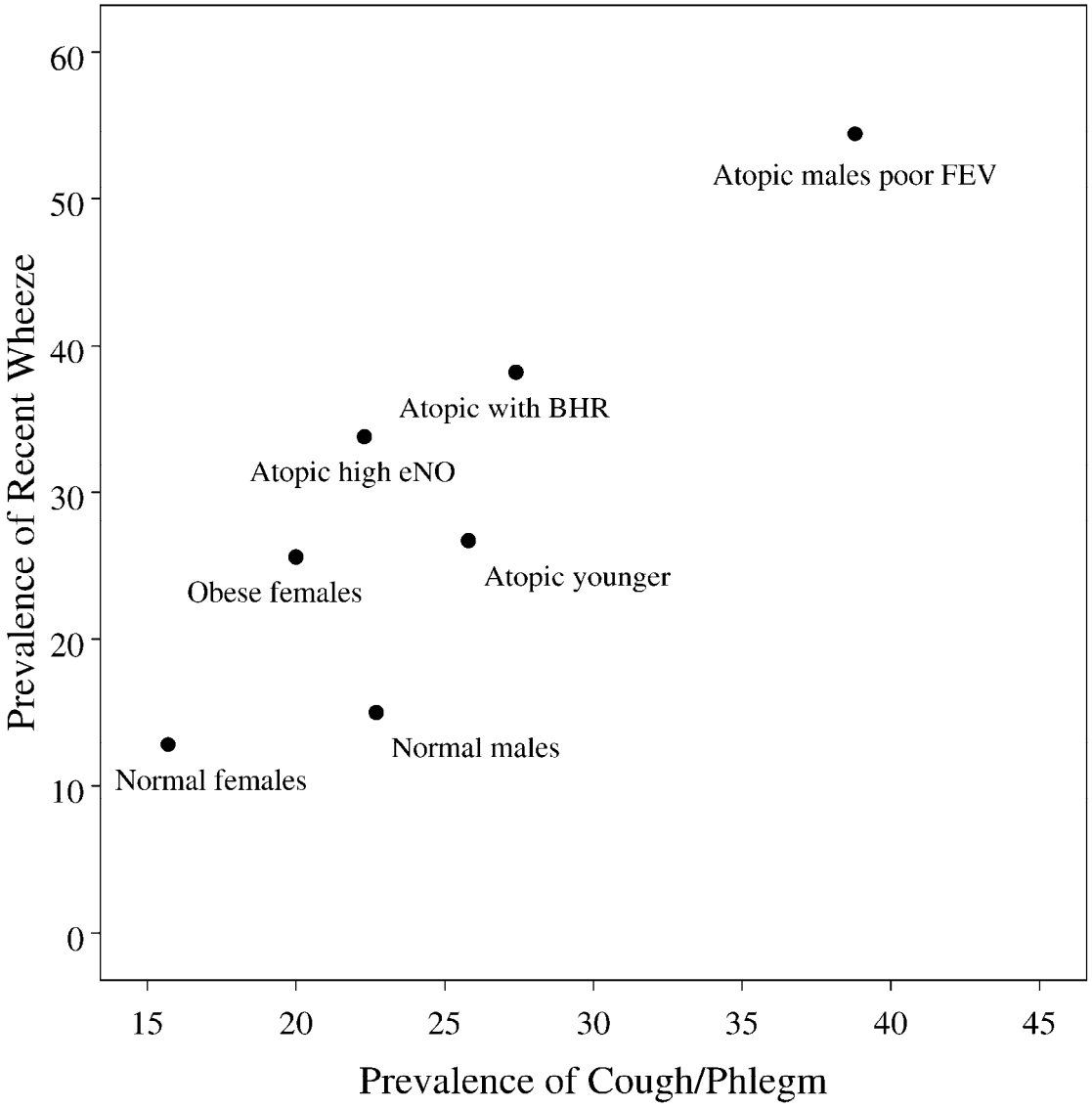


Figure 4. Patterns in prevalence of recent wheeze and cough/phlegm across clusters.



REFERENCES

1. Hargreave FE, Nair P. The definition and diagnosis of asthma. *Clin Exp Allergy* 2009: 39(11): 1652-1658.
2. Pavord ID, Wardlaw AJ. The A to E of airway disease. *Clin Exp Allergy*: 40(1): 62-67.
3. Wardlaw AJ, Silverman M, Siva R, Pavord ID, Green R. Multi-dimensional phenotyping: towards a new taxonomy for airway disease. *Clin Exp Allergy* 2005: 35(10): 1254-1262.
4. Boezen HM. Genome-wide association studies: what do they teach us about asthma and chronic obstructive pulmonary disease? *Proc Am Thorac Soc* 2009: 6(8): 701-703.
5. Haldar P, Brightling CE, Hargadon B, Gupta S, Monteiro W, Sousa A, Marshall RP, Bradding P, Green RH, Wardlaw AJ, Pavord ID. Mepolizumab and exacerbations of refractory eosinophilic asthma. *N Engl J Med* 2009: 360(10): 973-984.
6. Gibson PG, Simpson JL, Saltos N. Heterogeneity of airway inflammation in persistent asthma : evidence of neutrophilic inflammation and increased sputum interleukin-8. *Chest* 2001: 119(5): 1329-1336.
7. Spycher BD, Silverman M, Kuehni CE. Phenotypes of childhood asthma: are they real? *Clin Exp Allergy*: 40(8): 1130-1141.
8. Wenzel SE. Asthma: defining of the persistent adult phenotypes. *Lancet* 2006: 368(9537): 804-813.
9. Haldar P, Pavord ID, Shaw DE, Berry MA, Thomas M, Brightling CE, Wardlaw AJ, Green RH. Cluster analysis and clinical asthma phenotypes. *Am J Respir Crit Care Med* 2008: 178(3): 218-224.
10. Beasley R, Weatherall M, Travers J, Shirtcliffe P. Time to define the disorders of the syndrome of COPD. *Lancet* 2009: 374(9691): 670-672.
11. Spycher BD, Silverman M, Brooke AM, Minder CE, Kuehni CE. Distinguishing phenotypes of childhood wheeze and cough using latent class analysis. *Eur Respir J* 2008: 31(5): 974-981.
12. Moore WC, Meyers DA, Wenzel SE, Teague WG, Li H, Li X, D'Agostino R, Jr., Castro M, Curran-Everett D, Fitzpatrick AM, Gaston B, Jarjour NN, Sorkness R, Calhoun WJ, Chung KF, Comhair SA, Dweik RA, Israel E, Peters SP, Busse WW, Erzurum SC, Bleecker ER. Identification of asthma phenotypes using cluster analysis in the Severe Asthma Research Program. *Am J Respir Crit Care Med*: 181(4): 315-323.
13. Pillai SG, Tang Y, van den Oord E, Klotsman M, Barnes K, Carlsen K, Gerritsen J, Lenney W, Silverman M, Sly P, Sundy J, Tsanakas J, von Berg A, Whyte M, Ortega HG, Anderson WH, Helms PJ. Factor analysis in the Genetics of Asthma International Network family study identifies five major quantitative asthma phenotypes. *Clin Exp Allergy* 2008: 38(3): 421-429.
14. Weatherall M, Shirtcliffe P, Travers J, Beasley R. Use of cluster analysis to define COPD phenotypes. *Eur Respir J*: 36(3): 472-474.
15. Weatherall M, Travers J, Shirtcliffe PM, Marsh SE, Williams MV, Nowitz MR, Aldington S, Beasley R. Distinct clinical phenotypes of airways disease defined by cluster analysis. *Eur Respir J* 2009: 34(4): 812-818.
16. Lebowitz MD, Burrows B. Comparison of questionnaires: the BMRC and NHLI respiratory questionnaires and a new self-completion questionnaire. *Am Rev Respir Dis* 1976: 113(5): 627-635.
17. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med* 1999: 159(1): 179-187.
18. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, Crapo R, Enright P, van der Grinten CP, Gustafsson P, Jensen R, Johnson DC, MacIntyre N, McKay R, Navajas D, Pedersen OF, Pellegrino R, Viegi G, Wanger J. Standardisation of spirometry. *Eur Respir J* 2005: 26(2): 319-338.
19. Buist AS, Vollmer WM, Sullivan SD, Weiss KB, Lee TA, Menezes AM, Crapo RO, Jensen RL, Burney PG. The Burden of Obstructive Lung Disease Initiative (BOLD): rationale and design. *COPD* 2005: 2(2): 277-283.

20. Ryan G, Latimer KM, Dolovich J, Hargreave FE. Bronchial responsiveness to histamine: relationship to diurnal variation of peak flow rate, improvement after bronchodilator, and airway calibre. *Thorax* 1982: 37(6): 423-429.
21. Kharitonov SA, Yates D, Barnes PJ. Increased nitric oxide in exhaled air of normal human subjects with upper respiratory tract infections. *Eur Respir J* 1995: 8(2): 295-297.
22. Green RH, Brightling CE, Woltmann G, Parker D, Wardlaw AJ, Pavord ID. Analysis of induced sputum in adults with asthma: identification of subgroup with isolated sputum neutrophilia and poor response to inhaled corticosteroids. *Thorax* 2002: 57(10): 875-879.
23. Simpson JL, Scott R, Boyle MJ, Gibson PG. Inflammatory subtypes in asthma: assessment and identification using induced sputum. *Respirology* 2006: 11(1): 54-61.
24. Moore WC, Meyers DA, Wenzel SE, Teague WG, Li H, Li X, D'Agostino R, Jr., Castro M, Curran-Everett D, Fitzpatrick AM, Gaston B, Jarjour NN, Sorkness R, Calhoun WJ, Chung KF, Comhair SA, Dweik RA, Israel E, Peters SP, Busse WW, Erzurum SC, Bleecker ER. Identification of asthma phenotypes using cluster analysis in the Severe Asthma Research Program. *Am J Respir Crit Care Med* 2010: 181(4): 315-323.
25. Patel BD, Coxson HO, Pillai SG, Agusti AG, Calverley PM, Donner CF, Make BJ, Muller NL, Rennard SI, Vestbo J, Wouters EF, Hiorns MP, Nakano Y, Camp PG, Nasute Fauerbach PV, Screaton NJ, Campbell EJ, Anderson WH, Pare PD, Levy RD, Lake SL, Silverman EK, Lomas DA. Airway wall thickening and emphysema show independent familial aggregation in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2008: 178(5): 500-505.
26. Colebatch HJ, Finucane KE, Smith MM. Pulmonary conductance and elastic recoil relationships in asthma and emphysema. *J Appl Physiol* 1973: 34(2): 143-153.
27. Sin DD, Vestbo J. Biomarkers in chronic obstructive pulmonary disease. *Proc Am Thorac Soc* 2009: 6(6): 543-545.
28. Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, von Mutius E, Farrall M, Lathrop M, Cookson WO. A large-scale, consortium-based genomewide association study of asthma. *N Engl J Med*: 363(13): 1211-1221.
29. Travers J, Marsh S, Williams M, Weatherall M, Caldwell B, Shirtcliffe P, Aldington S, Beasley R. External validity of randomised controlled trials in asthma: to whom do the results of the trials apply? *Thorax* 2007: 62(3): 219-223.