

Dimensions of Respiratory Symptoms in Preschool Children

Population-based Birth Cohort Study

Jaclyn A. Smith¹, Richard Drake², Angela Simpson¹, Ashley Woodcock¹, Andrew Pickles³, and Adnan Custovic¹

¹Respiratory Research Group, ²Divisions of Psychiatry and ³Biostatistics, Health Methodology Research Group, University of Manchester, Manchester, United Kingdom

Rationale: A focus on distinctive collections of symptoms may be more informative of the probability of respiratory disease than individual and possibly transient phenotypes. However, such collections or components of overall symptomatology need to be valid, and their relationship established with the known risk factors and physiologic measures.

Objectives: To analyze detailed parentally reported respiratory symptoms by principal components analysis and derive symptom components; to examine the relationship of such components with measures of lung physiology and atopy.

Methods: An unselected, population-based birth cohort (n = 946). **Measurements and Main Results:** Interviewer-administered questionnaires, lung function (specific airway resistance [sRaw]), airway reactivity (dry air challenge), and atopic status were obtained at ages 3 and 5 years; principal components analysis and multivariate analysis of variance were used to analyze the data. The four-component solution (wheeze, cough, colds, chronic symptoms) explained 53.2% of the variance in symptoms at age 3, and the five-component solution (wheeze, wheeze with irritants, wheeze with allergens, cough, chest congestion) explained 49.8% of variance at age 5. The multivariate analysis revealed novel relationships between symptoms, risk factors for asthma, and measures of lung function. At age 3, sRaw and the interaction between maternal asthma and child's atopy were not only related to wheeze but also independently to the cough component. At age 5, overall wheeze and allergic wheeze were related to lung function and airway reactivity; child's atopy was only related to symptoms when considered as a continuous trait.

Conclusions: Our analysis supports the need to move beyond the presence or absence of individual symptoms. Syndromes of coexisting symptoms more likely reflect underlying pathophysiologic processes.

Clinical trial registered with <http://www.controlled-trials.com> (ISRCTN72673620).

Keywords: principal components analysis; wheeze phenotypes; childhood asthma; cough

Childhood asthma is a clinical syndrome in which the diagnosis is often based on a history of wheezing and/or other associated respiratory symptoms (e.g., cough), coupled with the objective measurements of lung function and/or airway inflammation (1). In contrast, in epidemiologic studies, children are often assigned a phenotype based on the response to a single question (e.g.,

AT A GLANCE COMMENTARY

Scientific Knowledge on the Subject

In epidemiologic studies of respiratory symptoms, children are often assigned a phenotype based on the response to a single question (e.g., presence/absence of wheeze).

What This Study Adds to the Field

This study demonstrates the need to move beyond the presence of individual symptoms and demonstrates that, in children aged 3 and 5 years, syndromes of coexisting symptoms are significantly related lung function and risk factors for asthma.

presence/absence of wheeze or doctor-diagnosed asthma), despite the fact that more detailed information has been collected using extensive questionnaires (2).

Reshaping the wealth of epidemiologic data into a valid but comprehensible form to inform the clinical decision-making process would be of great value. One means to do this, which has been used for over 20 years in other branches of medicine and psychology, is principal components analysis (PCA) (3). PCA is a quantitative method for data reduction to extract simplified patterns from correlated multi-item data, and has relatively established criteria for deciding what may be common underlying components that best explain a dataset. It distills the mass of information from many questionnaire items into a small number of components, which explain the overall variance in item scores, and can generate coherent subsets of items. These components should be more reliable and valid than single items. Subsequent analysis may also reveal relationships between components and objective measures of disease.

On the few occasions that PCA has been used in asthma and allergy research, symptoms formed a single component separate from lung physiology (4), atopy (5), systemic/airway inflammation (6), and quality of life (7), providing evidence of a disjunction between etiologic factors, symptoms, and consequences of symptoms. The fact that most cases of persistent wheezing and asthma begin in early childhood provides an added level of complexity. Our understanding of the nature of childhood wheezing illness has been augmented by the characterization of distinct wheeze phenotypes, based on the temporal patterns of wheezing (early, late onset, persistent) (8). Given this fragmentation of the trajectory and pathophysiology of asthma symptoms, it is not surprising that there have been calls to abandon the concept of asthma as a single disease (9).

We hypothesized that, rather than relying on the presence or absence of an individual symptom (e.g., wheezing) or physician diagnosis (e.g., asthma), more useful information may be obtained by examining syndromes of coexisting symptoms derived from the collected data. Such syndromes may better reflect the underlying

(Received in original form September 25, 2007; accepted in final form March 19, 2008)

Supported by Asthma UK grant 04/014 and Moulton Charitable Trust.

Correspondence and requests for reprints should be addressed to Dr. Jaclyn A. Smith, M.D., Ph.D., University of Manchester, ERC Building, Second Floor, Wythenshawe Hospital, Manchester M23 9LT, UK. E-mail: jacky.smith@manchester.ac.uk

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org

Am J Respir Crit Care Med Vol 177, pp 1358–1363, 2008

Originally Published in Press as DOI: 10.1164/rccm.200709-1419OC on March 20, 2008

Internet address: www.atsjournals.org

pathophysiologic processes and predict not only the presence but also the persistence of chronic respiratory disease, and hence be useful in both diagnosis and treatment decision making.

Therefore, in a setting of a population-based birth cohort, we used PCA to analyze detailed parentally reported respiratory symptoms at ages 3 and 5 years and derived representative symptom components. We tested the validity of these components by examining their relations to measures of lung physiology and child's atopy at each age, and between components from one time point to the other. These data have been presented in abstract form (10).

METHODS

Design

The Manchester Asthma and Allergy Study is an unselected, population-based birth cohort (11). The study was approved by the South Manchester Local Research Ethics Committee, Manchester, UK. Written, informed consent was obtained from subjects' parents/guardians.

We recruited participants during the first trimester of the pregnancy; children were monitored prospectively and reviewed at age 3 and 5 years. For detailed description of the cohort, procedures, and analysis, *see* the online supplement.

Procedures

Symptoms. Validated respiratory questionnaires were interviewer administered (American Thoracic Society [ATS] [Reference 12], for children age 3; ATS and International Study of Asthma and Allergies in Childhood [ISAAC] [Reference 2], for children age 5).

Lung function. Specific airway resistance (sRaw) was measured using plethysmography (ages 3 and 5 yr) (13, 14). Airway reactivity was assessed by eucapnic voluntary hyperventilation challenge (dry air; age 5) (15, 16).

Atopy. Children were skin tested to common inhalant and food allergens (age 3 and 5). At age 5 years, we also measured specific IgE; sensitization was defined as any value greater than 0.2 kilo-units of allergen per liter. To quantify sensitization, we summed the levels of specific IgE to mite, cat, and dog (17).

Statistical Analysis

All analyses were performed using SPSS version 13.0 (SPSS, Inc., Chicago, IL). PCA is an exploratory technique for investigating patterns within a set of variables (in the current analysis, responses to a series of respiratory symptom questions). The large numbers of responses are reduced to a much smaller number of representative components, based on the covariance among responses. If subsets of symptoms are correlated, this suggests they are measuring aspects of a common underlying process; several components may suggest a series of underlying processes.

We used the same procedures to analyze the symptom items at ages 3 and 5 years. The steps involved in the analyses were as follows:

1. **Item selection:** Repetitive questions and questions that correlated with partial correlation coefficients greater than 0.4 to other questions were removed (i.e., not all questions were included).
2. **Component extraction:** PCA generated the component solutions, and the number of components was selected using eigenvalues greater than 1 and by examination of scree plots (18).
3. **Rotation:** The rotation of the components generates component loading scores. These are a measure of the relationship between an individual question and the underlying component: the greater the loading, the purer a measure of that component the item is. Components were rotated using standard mathematical processes to give as much clarity as possible about correlated items. We examined different rotation techniques. An oblique rotation was chosen (promax), based on the implausibility of independent components assumed by orthogonal rotations. For clarity, only items with the conventional loading of 0.4 and above were interpreted.
4. **Components scores:** For each subject included in the PCA, we calculated component scores. These scores are estimates of the

scores each individual would have received on each of the components if they had been measured directly (*see* the online supplement for details).

5. **Component validity:** To validate the components, we performed a multivariate analysis of variance (MANOVA) with PCA scores as outcomes, and lung physiology, atopy, and maternal asthma as predictors. If the predictor terms were significantly related to the set of PCA components (Pillai's test), individual associations were examined using specific tests.

RESULTS

Of 1,085 children born into the study, we reviewed 1,068 at age 3 years. Of these, 122 were prenatally randomized to an environmental control regime (11, 19, 20) and excluded from this analysis. Of the remaining 946 children, 904 attended the follow-up at age 5 years. Lung function data were available in 533 children at age 3 and 782 at age 5 years; 811 were skin tested at age 3, 798 at age 5, and 521 had IgE measured at age 5 years (Figure 1). Of the subjects, 20.2 and 27.1% were sensitized to at least one allergen at ages 3 and 5 years, respectively.

Component Generation

Age 3 years. Twenty-one items were included in the PCA (Kaiser-Meyer-Olkin measure of sampling adequacy, 0.92; Bartlett's test of sphericity, <0.001); item selection is summarized in Table E1 of the online supplement. Four components with eigenvalues greater than 1 were identified with a subsequent break in the scree plot (Figure E1a). The four-component solution explained 53.2% of the variance in symptoms. The questionnaire items loading greater than 0.4 are shown in Table 1, together with the proportion of variance explained by each component.

Age 5 years. The age 5 questionnaires contained a section on allergic and irritant triggers of wheeze, which was not included at age 3; 32 items were included (Kaiser-Meyer-Olkin measure of sampling adequacy, 0.89, Bartlett's test of sphericity, <0.001), and item selection is summarized in Table E2. Six eigenvalues greater than 1 were found. These components explained 54.5% of the variance in respiratory symptoms at this age. Because there was a break in the scree plot between five and six components (Figure E1b), we also examined the five-component solution, which explained 49.8% of the variance (eigenvalues > 1.35). The questionnaire items loading onto each component and the variance they explain are shown in Table 2. We selected the five-component solution, because the six-component solution split the chest congestion component into two, describing congestion with colds and chronic congestion. These two components were not significantly discriminated in the subsequent MANOVA, and appear to be less valid.

Different rotation techniques were examined (oblique promax and orthogonal varimax) but produced extremely similar results, demonstrating the stability of the components.

Correlations between components. At both ages, correlations among components from the promax solution were weak to moderate (Tables E3a and E3b). Each component at age 3 correlated most strongly with the equivalent component at age 5 (Table E3c). The two additional wheeze components at age 5 (wheeze with irritants and wheeze with allergens) did not correlate with the wheeze component at age 3 but did correlate with cough, colds, and chronic symptoms.

Correlations between the component scores at different ages provide a further indicator of validity, suggesting stability over time.

Symptom Components, Lung Physiology, and Atopy

Age 3. In the overall multivariate model, baseline lung function ($P = 0.001$) and the interaction between maternal asthma and

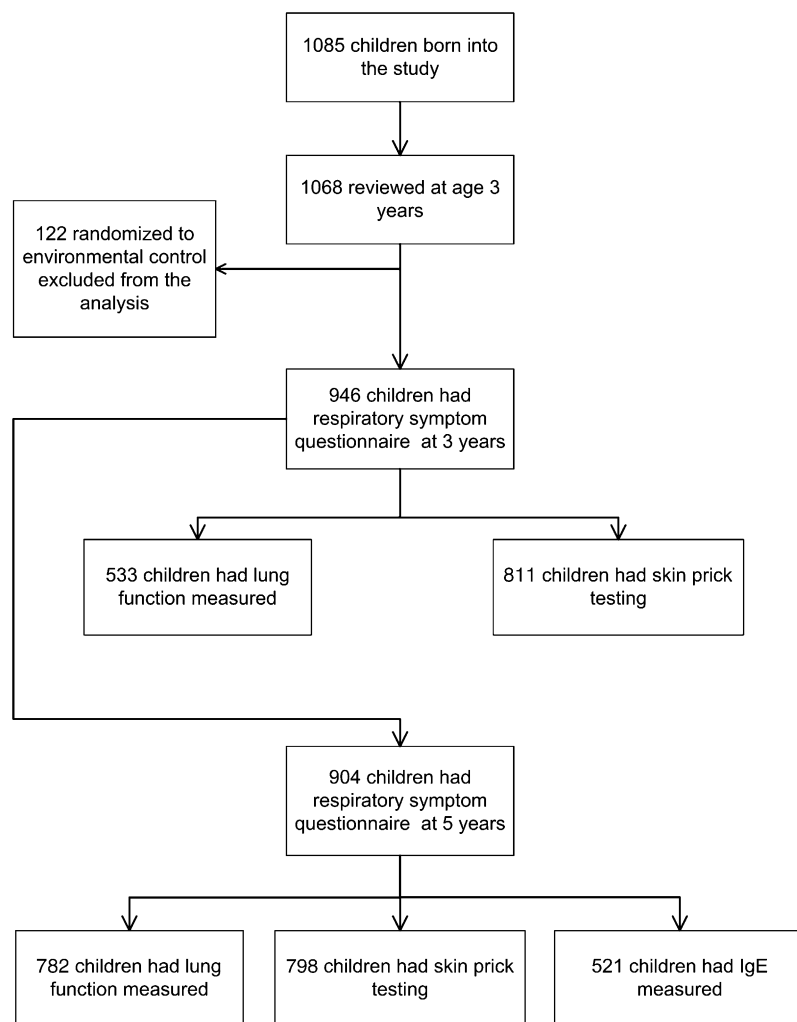


Figure 1. Consort diagram.

child's atopy ($P = 0.01$) were significantly associated with the wheeze ($P < 0.001$) and cough components ($P = 0.003$) but not colds ($p = 0.52$) or chronic symptoms ($P = 0.32$). Table 3 summarizes the significant relationships between individual predictors and the symptom components (post hoc analyses from each MANOVA).

Age 5. In a model that used atopic sensitization as a dichotomous variable, several physiologic variables were significantly related to symptom components: baseline lung function ($P \leq 0.001$), airway reactivity ($P = 0.016$), the interaction between maternal asthma and child's baseline lung function ($P = 0.007$), and maternal asthma and child's airway reactivity ($P = 0.047$). Four of five symptom components were significant in this multivariate model (wheeze, $P = 0.001$; cough, $P < 0.001$; wheeze with allergens, $P = 0.002$; chest congestion, $P = 0.001$), but not wheeze with irritants ($P = 0.29$). However, atopy (either on skin tests or IgE) was not a significant predictor of any of the symptom components.

In contrast, in a model using atopy as a continuous variable (i.e., the absolute level of specific IgE), the nature of these relationships changed. In this model, a sum of specific IgE ($P < 0.001$), baseline lung function ($P = 0.001$), the interaction between maternal asthma and sum of specific IgE ($P = 0.003$), and the interaction between maternal asthma and child's airway reactivity ($P = 0.04$) were significantly associated with the components. Again, four of five symptom components were significant in the multivariate model (wheeze, $P < 0.001$; cough,

$P < 0.001$; wheeze with allergens, $P < 0.001$; chest congestion, $P = 0.001$), but not wheeze with irritants ($P = 0.29$). Table 4 summarizes the relationships between the significant predictors and each individual symptom component (post hoc analyses from each MANOVA).

DISCUSSION

Key Results

We have demonstrated that PCA can be successfully used to produce respiratory symptom components in preschool-age children from validated questionnaires administered in a population-based birth cohort. As expected, questions pertaining to wheeze comprised a component that explained the largest proportion of the variance at both age 3 and 5 years. At age 3, other components included questions pertaining to cough, colds, and chronic symptoms. At age 5 years, in addition to the existing wheeze component, two further wheeze-related components emerged (wheeze with irritants and wheeze with allergens); these questions had not formed part of the assessment at age 3.

The components we identified have face validity (as they seem intuitively correct) and content validity (as we observed significant relationships with objective measures of respiratory function and known risk factors for asthma, e.g., maternal asthma). In addition, we observed a degree of consistency over time, suggesting their stability. Viewing atopic sensitization as a dichotomous trait in its relationship to respiratory symptom components

TABLE 1. PATTERN MATRIX: ITEM LOADINGS AT AGE 3 FOR A FOUR-COMPONENT SOLUTION

Components (% variance) and Items	Component Loadings			
	1	2	3	4
1. Wheeze (32.5%)				
Wheeze frequency*	-0.850			
Wheeze causing breathlessness	0.791			
Breathlessness frequency†	-0.768			
Wheeze with colds	0.738			
Wheeze during last 3 mo	0.731			
Wheeze apart from colds	0.643			
Wheeze after exercise	0.462			
2. Cough (8.2%)				
Cough after exertion		0.779		
Cough with cold air		0.746		
Cough when excited		0.710		
Cough mainly at night		0.564		
Wheeze with cold air		0.475		
Cough apart from colds		0.467		0.427
3. Colds (6.8%)				
Loose cough			0.685	
Chest congestion with colds			0.652	
Cough with colds			0.647	
4. Chronic symptoms (5.6%)				
Chest congestion apart from colds				0.803
Wheeze most days/nights				0.502

All items shown load over 0.40 onto a single principal components analysis component.
* Wheezy on three or more occasions.
† Breathless on three or more occasions.

obscured the nature of their relationships, which only became apparent when atopy was considered as a continuous trait.

Limitations

The validity and reliability of the results are dependent on the validity of the questionnaires and the reliability of the responses. Although we have used comprehensive, well-established, and validated questionnaires, collected prospectively in a large cohort, we cannot exclude the possibility that important symptom questions were excluded. Only about half of the variance is explained by the selected components; this is, however, comparable to other similar analyses (5, 6, 21). Furthermore, as with all studies in preschool children, we relied on parental reporting of a child's symptoms and accept the limitations of this approach. However, one strength of PCA is its capacity to embrace all the symptom information simultaneously rather than relying on a single dichotomous answer to the question "Has your child wheezed in the last 12 months?"

The different questionnaire items used at two time points may account for the differences in the symptom components identified. The majority of the additional questions at age 5 described wheezing episodes, and the circumstances in which they occurred. By age 5 years, children are more likely to have been exposed to a wider range of environmental factors that may precipitate wheezing. Moreover, the additional wheeze components identified (wheeze with allergens and wheeze with irritants) were extremely poorly correlated with the wheeze component at age 3 years, implying that these are truly independent emerging symptom components.

PCA is an exploratory analysis and the inherent assumptions and limitations of PCA should be taken into account. Not all questions were included in the PCA; excluding repetitive and correlated questionnaire items presupposes that they will contribute nothing and the subjective selection of which items to remove may affect the results. The choice of rotation may also alter the outcome of the analysis, with some studies preferring

TABLE 2. PATTERN MATRIX: ITEM LOADINGS AT AGE 5 FOR A FIVE-COMPONENT SOLUTION

Components (% variance) and Items	Component Loadings				
	1	2	3	4	5
1. Wheeze (27.0%)					
Wheeze with colds	0.778				
Wheeze limiting speech	0.766				
Wheeze frequency*	0.653				
Wheeze after exercise	0.573				
Wheeze frequency at night†	0.559				
Woken short of breath	0.538				
Wheeze with cold air	0.517				
Wheeze with pollen	0.459			0.419	
Wheeze with dust	0.458				
2. Cough (6.8%)					
Cough after exertion		0.791			
Cough with cold air		0.708			
Cough during the night		0.695			
Cough during the day		0.665			
Cough when excited		0.561			
Cough with colds		0.791			
3. Wheeze with irritants (5.8%)					
Wheeze with soap/spray/detergent			0.688		
Wheeze with wool/clothing			0.664		
Wheeze with cigarette smoke			0.628		
Wheeze with fumes			0.607		
4. Wheeze with allergens (5.4%)					
Wheeze with cats				0.753	
Wheeze with other pets				0.687	
Wheeze with dogs				0.645	
Wheeze with foods			0.405	0.536	
5. Chest congestion (4.8%)					
Loose cough					0.737
Chest congestion with colds					0.703
Chest congestion most days ≥ 3 mo per yr					0.632
Chest congestion without colds					0.520

All items shown load over 0.40 onto a single principal components analysis component.
* Wheezy on three or more occasions.
† Wheeze at night ≥ weekly.

orthogonal rotations such as the varimax (6, 21, 22). Orthogonal rotations are useful in that they maximize the differences between the components, but the resulting factors do not correlate, implying etiologies for each component involving distinct non-overlapping risk factors and no correlation in exposure to them. Oblique rotations are generally considered more natural, because they generate components that may correlate. In the current study, the choice of rotation had no material effect on the results.

Interpretation

The timing of onset and resolution of respiratory symptoms and the circumstances in which they occur may provide an indication of underlying pathologic processes. Wheeze is not a feature of normal respiratory development but is a clinical symptom caused

TABLE 3. SUMMARY OF MULTIVARIATE ANALYSIS OF VARIANCE IN CHILDREN AGED 3 YEARS SHOWING SIGNIFICANT PREDICTORS OF SYMPTOM COMPONENTS

Component	Predictor Variable	P
Wheeze	Airway resistance	<0.001
	Child's atopy and maternal asthma	0.003
Cough	Airway resistance	0.041
	Child's atopy and maternal asthma	0.011
Colds	None	
Chronic symptoms	None	

TABLE 4. POST HOC ANALYSES FROM MULTIVARIATE ANALYSIS OF VARIANCE IN CHILDREN AGED 5 YEARS SHOWING SIGNIFICANT PREDICTORS OF INDIVIDUAL SYMPTOM COMPONENTS*

Component	Model 1		Model 2	
	Atopic Status, Dichotomous (skin prick test)		Atopic Status, Continuous (sum of specific IgE)	
	Predictor Variable	P	Predictor Variable	P
Wheeze	Airway resistance	<0.001	Airway resistance	0.001
	Airway reactivity	0.038	Sum of specific IgE	<0.001
	Airway resistance and maternal asthma	0.015		
Cough	Airway resistance	0.004	Airway resistance	0.006
	Airway reactivity	0.004	Maternal asthma and airway reactivity	0.007
Wheeze with allergens	Airway reactivity and maternal asthma	0.002	Sum of specific IgE	0.009
	Airway reactivity	0.028	Maternal asthma and airway reactivity	0.015
	Airway resistance	0.035	Maternal asthma and sum of specific IgE	0.005
Wheeze with irritants	None		None	
Chest congestion	Airway resistance	<0.001	Airway resistance	<0.001
	Airway resistance and maternal asthma	0.015	Sum of specific IgE	0.032

* Model 1 uses atopic status as a dichotomous variable (positive skin prick test to at least one allergen) and model 2 uses atopic status as a quantitative variable (sum of specific IgE to mite, cat, and dog). Other predictor variables included in the analysis were airway resistance, airway reactivity, and maternal asthma.

by various pathophysiologic processes and may be difficult to interpret without considering other symptoms. In contrast, cough and congestion symptoms may occur transiently in otherwise healthy children (e.g., in the context of respiratory virus infection), but may also be features of chronic respiratory disorder (e.g., asthma). Furthermore, where a child lies on the continuum between health and disease can be difficult for a parent or clinician to judge. Distinct, clearly defined symptom components may be more informative of the probability of chronic disease than a transient phenotype. However, to use this information, the components need to be valid, and their relationship established with the known risk factors and physiologic measures.

The symptom components generated by the PCA at both age 3 and 5 years were clear in character. The age 3 components correlated little with each other, apart from the wheeze and cough components. At age 5, the wheeze and cough components again show moderate correlations, but chest congestion seemed to be a separate dimension. Similarly, there was a stability of these main components as the children aged, and the correlations over time were reassuringly specific, despite the differences in the questionnaires used (i.e., the correlations of corresponding symptom components at two time points were consistently stronger than those between the different components at the same age). Overall, the component correlation coefficients were not strong between the two time points studied, most likely due to the transient nature of symptoms in some of the children. These components all remained separate when the varimax rotation of the model was used, confirming they are distinct and probably have different etiologies.

This multivariate analysis has revealed novel relationships between symptoms, known risk factors for asthma, and objective measures of lung function and atopy. For example, at age 3 years, sRaw was not only related to wheeze but also independently to the cough component. Moreover, the interaction between maternal asthma and child's atopy was independently related to both cough and wheeze. We have previously demonstrated a significant effect of the interaction between child's atopy and maternal asthma on lung function among asymptomatic children at age 3 years (14), and suggested that these features affect the development of wheeze via their effect on lung function. However, it was apparent in the current analysis that lung function was related to symptom components independent of the interaction between child's atopy and maternal asthma, raising questions about the underlying pathophysiologic mechanisms.

Analysis at age 5 years suggests different associations, implying different pathologies for each symptom component. When atopy was considered as a dichotomous variable, the symptom components were not related to child's atopic status. However, using atopy as a continuous trait revealed the associations between symptom components, lung function, airway reactivity, and level of allergen-specific IgE. Our findings suggest that end-organ susceptibility and systemic immune responses contribute independently to distinct symptom components and that their effect differs between components.

The additional questions at age 5 years enabled characterization of three separate wheeze components (wheeze, wheeze with allergens, and wheeze with irritants), each with different relationships to physiologic hallmarks of asthma. This raises a fundamental question as to whether all early childhood wheezing should be treated by the same antiasthma drugs (23–25).

Generalizability

We wish to emphasize that the prevalence of allergic sensitization among the parents of the children in our cohort is similar to that of young adults in the United Kingdom (26). This suggests that the subjects are representative of the general population, and that our results are generalizable.

It would be of value to explore similar symptom components in the other large birth cohort studies. This could be achieved either by each study generating its own "syndromes" from its datasets and comparing those between the studies, or by testing the components described here within the other datasets (provided the questionnaires used are identical). Another approach would be to identify such components in a pooled dataset, allowing a more generalized definition of syndromes, which can be used as endpoints in etiologic analyses or identification of treatment groups.

We acknowledge the relative complexity of this approach to the analysis of datasets collected using standardized questionnaires with multiple questions. However, once determined, the components become continuous clinical outcome variables in their own right, which can be used in further analyses. The advantage of this approach is that these components retain the appropriately weighted information collected from numerous questions, while reducing the problem of multiple testing when each individual question is used as a dichotomous variable.

Conclusions

Our analysis supports the need to move beyond the presence or absence of individual symptoms; indeed, in our dataset, the questionnaire items representing the presence of wheeze or cough were redundant, because they were not discriminatory. More useful information may be gained from examining syndromes of coexisting symptoms that are more likely to reflect the underlying pathophysiologic processes. Such an approach may eventually help identify individuals who would benefit from a particular treatment/intervention. Within this context, abandoning both asthma and atopic sensitization in favor of distinct symptom components and quantitative allergy may be more useful both in diagnostics and treatment decision making.

Conflict of Interest Statement: None of the authors has a financial relationship with a commercial entity that has an interest in the subject of this manuscript.

Acknowledgment: The authors thank the Manchester Asthma and Allergy Study (MAAS) children and their parents for their continued support and enthusiasm. They also acknowledge the dedication of the MAAS study team.

References

- Global Initiative for Asthma. Global strategy for asthma management and prevention. Bethesda, MD; National Heart, Lung, and Blood Institute; 2007.
- Asher MI, Keil U, Anderson HR, Beasley R, Crane J, Martinez F, Mitchell EA, Pearce N, Sibbald B, Stewart AW, *et al.* International Study of Asthma and Allergies in Childhood (ISAAC): rationale and methods. *Eur Respir J* 1995;8:483–491.
- Liddle PF. The symptoms of chronic schizophrenia: a re-examination of the positive-negative dichotomy. *Br J Psychiatry* 1987;151:145–151.
- Juniper EF, Wisniewski ME, Cox FM, Emmett AH, Nielsen KE, O'Byrne PM. Relationship between quality of life and clinical status in asthma: a factor analysis. *Eur Respir J* 2004;23:287–291.
- Holberg CJ, Halonen M, Solomon S, Graves PE, Baldini M, Erickson RP, Martinez FD. Factor analysis of asthma and atopy traits shows 2 major components, one of which is linked to markers on chromosome 5q. *J Allergy Clin Immunol* 2001;108:772–780.
- Leung TF, Wong GW, Ko FW, Lam CW, Fok TF. Clinical and atopic parameters and airway inflammatory markers in childhood asthma: a factor analysis. *Thorax* 2005;60:822–826.
- Juniper EF, Guyatt GH, Streiner DL, King DR. Clinical impact versus factor analysis for quality of life questionnaire construction. *J Clin Epidemiol* 1997;50:233–238.
- Martinez FD, Wright AL, Taussig LM, Holberg CJ, Halonen M, Morgan WJ. Asthma and wheezing in the first six years of life. The Group Health Medical Associates. *N Engl J Med* 1995;332:133–138.
- The Lancet. A plea to abandon asthma as a disease concept. *Lancet* 2006;368:705.
- Smith JA, Drake RJ, Simpson A, Woodcock A, Pickles A, Custovic A. Dimensions of chronic respiratory symptoms in pre-school children: population-based birth cohort study [abstract]. *Allergy* 2007;62:425.
- Custovic A, Simpson BM, Murray CS, Lowe L, Woodcock A. The National Asthma Campaign Manchester Asthma and Allergy Study. *Pediatr Allergy Immunol* 2002;13:32–37.
- Ferris BG. Epidemiology standardisation project II. Recommended respiratory disease questionnaires for use with adults and children in epidemiological research. *Am Rev Respir Dis* 1978;118:7–53.
- Bisgaard H, Klug B. Lung function measurement in awake young children. *Eur Respir J* 1995;8:2067–2075.
- Lowe L, Murray CS, Custovic A, Simpson BM, Kissen PM, Woodcock A. Specific airway resistance in 3-year-old children: a prospective cohort study. *Lancet* 2002;359:1904–1908.
- Nielsen KG, Bisgaard H. Lung function response to cold air challenge in asthmatic and healthy children of 2–5 years of age. *Am J Respir Crit Care Med* 2000;161:1805–1809.
- Zach M, Polgar G, Kump H, Kroisel P. Cold air challenge of airway hyperreactivity in children: practical application and theoretical aspects. *Pediatr Res* 1984;18:469–478.
- Simpson A, Soderstrom L, Ahlstedt S, Murray CS, Woodcock A, Custovic A. IgE antibody quantification and the probability of wheeze in preschool children. *J Allergy Clin Immunol* 2005;116:744–749.
- Tabachnick B, Fidell L. Using multivariate statistics, 5th ed. Boston: Allyn and Bacon; 2007.
- Custovic A, Simpson BM, Simpson A, Kissen P, Woodcock A. Effect of environmental manipulation in pregnancy and early life on respiratory symptoms and atopy during first year of life: a randomised trial. *Lancet* 2001;358:188–193.
- Custovic A, Simpson BM, Simpson A, Hallam C, Craven M, Brutsche M, Woodcock A. Manchester Asthma and Allergy Study: low-allergen environment can be achieved and maintained during pregnancy and in early life. *J Allergy Clin Immunol* 2000;105:252–258.
- Schatz M, Mosen D, Apter AJ, Zeiger RS, Vollmer WM, Stibolt TB, Leong A, Johnson MS, Mendoza G, Cook EF. Relationships among quality of life, severity, and control measures in asthma: an evaluation using factor analysis. *J Allergy Clin Immunol* 2005;115:1049–1055.
- Rosi E, Ronchi MC, Grazzini M, Duranti R, Scano G. Sputum analysis, bronchial hyperresponsiveness, and airway function in asthma: results of a factor analysis. *J Allergy Clin Immunol* 1999;103:232–237.
- Bisgaard H, Hermansen MN, Loland L, Halkjaer LB, Buchvald F. Intermittent inhaled corticosteroids in infants with episodic wheezing. *N Engl J Med* 2006;354:1998–2005.
- Guilbert TW, Morgan WJ, Zeiger RS, Mauger DT, Boehmer SJ, Szeffler SJ, Bacharier LB, Lemanske RF Jr, Strunk RC, Allen DB, *et al.* Long-term inhaled corticosteroids in preschool children at high risk for asthma. *N Engl J Med* 2006;354:1985–1997.
- Murray CS, Woodcock A, Langley SJ, Morris J, Custovic A. Secondary prevention of asthma by the use of Inhaled Fluticasone propionate in Wheezy Infants (IFWIN): double-blind, randomised, controlled study. *Lancet* 2006;368:754–762.
- Broadfield E, McKeever TM, Scrivener S, Venn A, Lewis SA, Britton J. Increase in the prevalence of allergen skin sensitization in successive birth cohorts. *J Allergy Clin Immunol* 2002;109:969–974.

**DIMENSIONS OF RESPIRATORY SYMPTOMS IN PRE-SCHOOL CHILDREN:
POPULATION-BASED BIRTH COHORT STUDY**

Jaclyn A Smith, Richard Drake, Angela Simpson, Ashley Woodcock,

Andrew Pickles, Adnan Custovic

ONLINE DATA SUPPLEMENT

METHODS

Study design

The Manchester Asthma and Allergy Study is a prospective study which follows the development of asthma and other atopic disorders in a cohort of infants.

Setting

The setting is the maternity catchment area of Wythenshawe and Stepping Hill Hospitals, comprising of 50 square miles of South Manchester and Cheshire, UK, a stable mixed urban-rural population. Study was approved by the Local Research Ethics Committee. Informed consent was obtained from all parents.

Participants

All pregnant women were screened for eligibility at 'Booking' antenatal visits (8th-10th week of pregnancy). The study was explained to the parents, and informed consent for initial questionnaires and skin prick testing was obtained. Both parents completed a questionnaire about their and their partner's history of asthma and allergic diseases and smoking habits.

When informed consent had been obtained, pregnant ladies and their partners were skin prick tested. If the pregnant woman's partner was not present at the antenatal clinic visit, an invitation was sent for him to attend an open-access evening clinic at Wythenshawe Hospital for skin prick testing and questionnaire. Once both parents had completed questionnaires and skin prick testing, a full explanation of the proposed future follow-up for the child was given.

Of the 1499 couples who met the inclusion criteria (≤ 10 weeks of pregnancy, maternal age ≥ 18 years, questionnaire and skin test data available for both parents), 288 declined to take part in the study. Of 1211 couples who initially agreed to take part, 1085 had a successful full-term pregnancy (> 36 weeks gestation) and gave consent to a further follow-up. Of those, 128 were

prenatally randomized into an environmental control group (see below), and 957 were followed as an observational arm of the cohort.

Allocation to study groups: Children were assigned to risk groups according to parental sensitisation status. Nested within the cohort is a primary prevention study in which a subgroup of children at high risk (both parents atopic) without pets in home was randomised to either stringent environmental control (allergen impermeable covers for the parental and infant bed, hot washing of bedding weekly, HEPA vacuum cleaner, hard floor for the nursery) or a normal regime. Children randomized to the environmental control group were excluded from this analysis.

Follow-up: The children have been followed prospectively, and attended review clinics at age 3 and 5 years (± 4 weeks). We completed assessment at age 5 years in April 2003. The overall follow-up rate during the first five years of life in the observational part of the cohort was exceptionally high (874/957 [91.3%] age 3 years; 840/957 [87.8%] age 5 years).

Variables

Symptoms

Validated respiratory questionnaires were interviewer-administered at age 3 and 5 years to collect the information on symptoms, physician-diagnosed illnesses and treatments received.

Specific airway resistance

At age 3 and 5 years we carried out measurements of specific airway resistance (sR_{aw}) to assess airway function in all children who were willing to cooperate. Measurements of sR_{aw} were made using a constant volume whole body plethysmograph (Masterscreen Body 4.34; Jaeger, Würzburg, Germany). Flow and volume were measured with a heated differential pressure screen-type pneumotachograph with a resistance of $0.036 \text{ kPa}^{-1} \cdot \text{s}$ and a dead space of 160mls. Pressure measurements were made with a pressure transducer (Nr.660.99007; Hube Control AG,

Wuerenlos, Switzerland) with an input range of ± 100 Pa, a resolution of 0.05 Pa and a linear response up to 10 Hz. The plethysmograph was calibrated daily. Sensors in an ambient unit supplied with the plethysmograph recorded ambient data on temperature, humidity and barometric pressure. The pneumotachograph was volume calibrated according to the American Thoracic Society recommendations using a 2 L syringe at flow rates of 0 – 1.5, 1.5 – 5 and >5 l/s. The half value period was calibrated to ensure a specific leakage in the box of 4-7 seconds. The pressure transducer was calibrated using a 50 mL motor driven piston pump to generate sinusoidal variations of plethysmographic pressure. Electronic body temperature, pressure, and saturation (BTPS) compensation was applied throughout, using a time-shift of 60 ms. sR_{aw} is measured by a single-step procedure from the simultaneously measured changes of respiratory flow and changes of plethysmographic pressure, omitting the measurement of TGV^{E1} . Measurements were carried out during tidal breathing using a facemask, which was adapted by fitting a standard paediatric facemask with a non-compressible mouthpiece made from silicone tubing. The end of the tubing was made rigid with an aluminium splint. The purpose of this was to maintain stable airway opening, prevent nose breathing and support the cheeks. The procedure was explained to the accompanying adult and the use of the facemask demonstrated to the child. The children were encouraged to sit in the plethysmograph alone but if they refused, the accompanying adult, usually a parent, accompanied the child in the plethysmograph cabinet with the child seated on their knee. The door of the plethysmograph was closed and the subject asked to breathe through the facemask.

Children were encouraged to breathe at a rate of 30-45 breaths per minute. If a parent accompanied the child, the adult was asked to inhale and hold their breath for approximately 20 seconds. sR_{aw} measurements were made once a stable breathing pattern had been re-established. Once a stable breathing pattern was established, at least three measurements of sR_{aw} were

performed, and each was calculated from the medians of 5 consecutively measured technically acceptable loops^{E2} (each child performed at least 15 loops). The mean of these 3 measurements of *effective* sR_{aw} was used in the analysis. The measured values of sR_{aw} were corrected for the influence of the pneumotachograph screen (R_{screen}) and for the volume displacement caused by the subject (or subject + parent).

Children were asymptomatic at the time of assessment of lung function. A total of 530 children successfully performed lung function measurement at age 3 years, and 730 had sR_{aw} measured at age 5 years.

Airway reactivity

At the 5-year follow-up, airway reactivity was assessed by means of a Eucapnic Voluntary Hyperventilation (EVH) challenge. We constructed a breathing circuit consisting of a precision flow meter which directs gas from a cylinder into a large capacity reservoir bag. The subject hyperventilated, via a demand valve and large bore tubing, from a cylinder of certified medical gas containing 21% O₂, 5% CO₂, remainder N₂ with a water content <10mg per litre for 6 minutes at a ventilation rate of 75% of maximum voluntary ventilation (MVV) which was estimated pre-challenge as $FEV_1 \times 22.5$. Some children were unable to perform forced manoeuvres satisfactorily and FEV_1 could therefore not be measured. In these cases the predicted FEV_1 was used to calculate MVV.

EVH challenge was performed by a single-step procedure using a protocol established by Zach et al^{E3}. Parents were asked to withhold from giving the children short acting bronchodilators for 6 hours and long acting bronchodilators for 12 hours prior to testing. Children inhaled through a modified facemask as described previously. They were encouraged to match their ventilation to the reservoir balloon with the aim of keeping it partially inflated. This ensured that the correct level of ventilation was maintained.

Specific airway resistance was measured at baseline and 2, 5 and 10 minutes post challenge to assess the degree of airway reactivity. The highest sR_{aw} value measured at 2, 5 or 10 minutes post challenge was recorded, and the response was expressed as percent change in sR_{aw} . A bronchodilator (400 micrograms Salbutamol via a large volume spacer device) was given if the child became symptomatic or sR_{aw} increased by $>100\%$ for therapeutic purposes ($n=6$). Lung function was reassessed after 15 minutes to ascertain whether it had returned to baseline values.

Allergic sensitization

We ascertained allergic sensitisation at age 3 and 5 by skin prick testing (*D pteronyssinus*, cat, dog, mixed grasses, egg, milk, positive and negative controls). Sensitization was defined as wheal at least 3 mm greater than the negative control to at least one allergen tested.

At age 5 years specific IgE to mite, cat, dog, grasses, milk, egg, and peanut was measured in serum by ImmunoCAP (Phadia, Uppsala, Sweden). The detection limit of the assay was 0.2 kUA/L (kilo units of Allergen per litre). When IgE was used as a dichotomous variable, sensitization was defined as any value >0.2 kUA/L. In order to quantify allergic sensitization, we then summed the levels of specific IgE to mite, cat and dog as previously described.

Statistical Methods

1. Item selection: Redundant questions were removed, i.e. repetitive questions and questions which correlated with partial correlation coefficients >0.4 to other questions. The choice of which of the correlated items to remove was based on results of correlations with the other items. The few items with a continuous response were removed and those with an ordinal response were converted to binary categories.
2. Component extraction: PCA was used to generate the component solutions and the numbers of components were selected using the eigenvalues and scree plots(18).

3. Rotation: Component rotation is used to improve the interpretability of the solution. Both oblique promax and orthogonal varimax rotations were examined but produced extremely similar components (varimax solution results available on request), demonstrating stability of the components. An oblique rotation was chosen (Promax), based on the implausibility of independent components assumed by orthogonal rotations. The loading of the items onto each component is a measure of the relationship between that item and the component; the greater the loading, the purer a measure of that component the item is. Only items with the conventional loading of 0.4 and above were interpreted, for clarity.

4. Component Score Calculation: Components scores are estimates of the scores each individual would have received on each of the components had they been measured directly. Scores are calculated from the individual's responses to the items which constitute each component. The responses are weighted using the factor score coefficients, which are calculated by the regression method which adjusts the factor loadings to take into account the initial correlations between variables. There are several techniques available for calculating such scores, the regression method is one of the simplest.

5. Component Validity: To validate the components, once PCA solutions had been selected and scores generated, at each follow up point a MANOVA (multivariate analysis of variance) was run with the PCA scores as outcome variables and the lung physiology, atopy and asthma risk factor terms as the predictors. If the predictor terms were significantly related to the PCA components according to Pillai's test then individual associations between predictors and components were examined using specific tests. Correlations between the component scores at different ages are a further indicator of validity as this determines stability over time.

Supplemental Figures and tables

Table E1 Questionnaire items in PCA age 3

	Included items	Excluded items	Reason for exclusion
1	Cough with colds	Cough more than others	Correlations
2	Cough apart from colds	Wheeze ever	Correlations
3	Cough mainly at night	Age first attack	Continuous variable
4	Cough after exertion	Age last wheeze	Continuous variable
5	Cough when excited	Ever required medication for wheeze	Correlations
6	Cough with cold air	Breathing normal between attacks	Correlations
7	Loose cough	≥3 episodes of wheeze causing breathlessness	Correlations
8	Chest congested with colds		
9	Chest congested apart from colds		
10	Wheeze with colds		
11	Wheeze apart from colds		
12	Wheeze most days or nights		
13	Wheeze with cold air		
14	Wheeze during the last 6 months		
15	Wheezing causing breathlessness		
16	Wheeze after exercise		
17	Wheeze Frequency *		
18	Breathlessness Frequency [†]		
19	Breathless after exertion		

*wheezy on 3 or more occasions, [†]breathless on 3 or more occasions

Table E2 Questionnaire items in PCA age 5

	Included items	Excluded items	Reason for exclusion
1	Cough with colds	Cough more than others	Correlations
2	Cough during the day	How many years congested ≥ 3 months per year	Continuous variable
3	Cough during the night	Wheeze ever	Correlations
4	Cough after exertion	Wheeze age 4-5 years	Repetitive
5	Cough when excited	Wheeze apart from exercise	Repetitive
6	Cough with cold air	Wheeze with weather	Correlations
7	Loose cough	Wheeze apart from colds/flu	Correlations
8	Chest congested with colds	Wheeze with exercise	Repetitive
9	Chest congested apart from colds	Wheeze with colds/flu	Repetitive
10	Chest congested most days ≥ 3 months per year	Woken with chest tightness	Correlations
11	Wheeze frequency*		
12	Wheeze frequency at night [¥]		
13	Wheeze limiting speech		
14	Wheeze after exercise		
15	Wheeze with colds/flu		
16	Wheeze with cold air		
17	Wheeze with pollen		
18	Wheeze with emotion		
19	Wheeze with fumes		
20	Wheeze with dust		
21	Wheeze with cats		
22	Wheeze with dogs		
23	Wheeze with other pets		
24	Wheeze with wool/clothing		
25	Wheeze with cigarette smoke		
26	Wheeze with foods/drink		
27	Wheeze with soap/ spray/ detergent		
28	Woken short of breath		

*wheezy on 3 or more occasions, [¥]wheeze at night $\geq$ weekly

Table E3. Component Correlation Matrices

a) age 3

Components	(1) Wheeze	(2) Cough	(3) Colds	(4) Chronic Symptoms
(1) Wheeze	1.000	-	-	-
(2) Cough	0.596	1.000	-	-
(3) Colds	0.183	0.214	1.000	-
(4) Chronic Symptoms	0.272	0.294	-0.019	1.000

b) age 5

Components	(1) Wheeze	(2) Cough	(3) Wheeze with Irritants	(4) Wheeze with Allergens	(5) Chest Congestion
(1) Wheeze	1.000	-	-	-	-
(2) Cough	0.477	1.000	-	-	-
(3) Wheeze with Irritants	0.405	0.381	1.000	-	-
(4) Wheeze with Allergens	0.382	0.294	0.372	1.000	-
(5) Chest Congestion	0.178	.296	0.213	0.139	1.000

c) correlations between components age 3 and 5; correlation coefficients (p values).

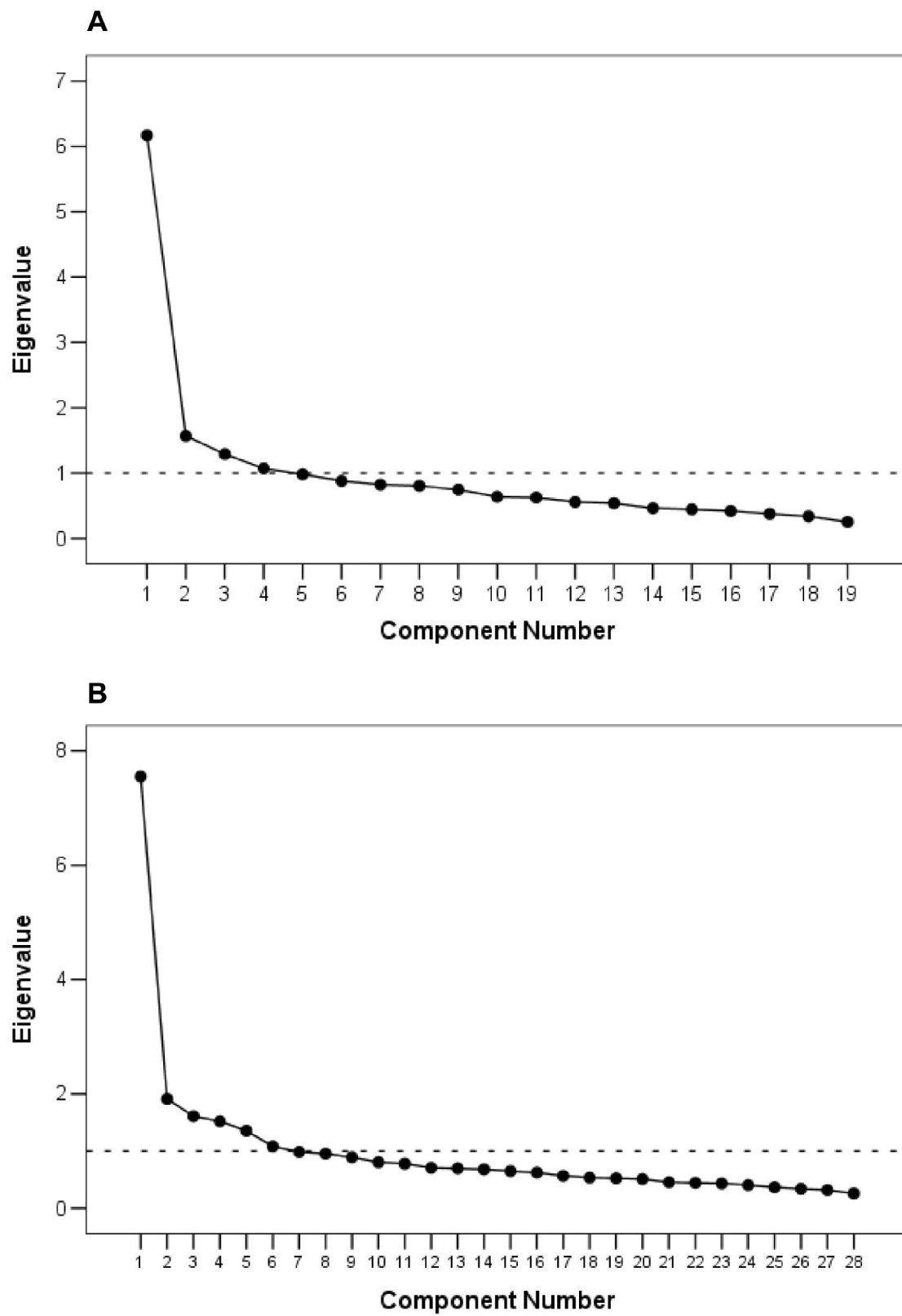
COMPONENT	Wheeze age 5	Cough age 5	Wheeze with Irritants age 5	Wheeze with Allergens age 5	Chest Congestion age 5
Wheeze age 3	0.32 (<0.001)	0.26 (<0.001)	0.02 (0.60)	-0.07 (0.06)	0.11 (0.002)
Cough age 3	0.28 (<0.001)	0.46 (<0.001)	-0.09 (0.01)	-0.14 (<0.001)	0.11 (0.001)
Colds age 3	0.23 (<0.001)	0.24 (<0.001)	-0.15 (<0.001)	-0.17 (<0.001)	0.30 (<0.001)
Chronic Symptoms age 3	-0.11 (0.002)	-0.02 (0.54)	0.12 (0.001)	0.15 (<0.001)	-0.05 (0.14)

FIGURES LEGENDS

Figure E1a Scree plot showing Eigenvalues for symptom components at age 3; reference line at Eigenvalue of 1.0

Figure E1b Scree plot showing Eigenvalues for symptom components at age 5; reference line at Eigenvalue of 1.0

Supplementary Figure E1



REFERENCES

- E1. Buhr W, Jorres R, Knapp M, Berdel D. Diagnostic value of body plethysmographic parameters in healthy and asthmatic young children is not influenced by breathing frequency. *Pediatr Pulmonol* 1990;8:23-28.
- E2. Lowe L, Murray CS, Custovic A, Simpson BM, Kissen PM, Woodcock A. Specific airway resistance in three year-old children. *Lancet* 2002;359:1904-8
- E3. Zach M, Polgar G, Kroisel P. Cold air challenge of hyperreactivity in children: practical application and theoretical aspects. *Paediatr Res* 1984;18:469-78