Identifying the heterogeneity of young adult rhinitis through cluster analysis in the Isle of Wight birth cohort

Ramesh J. Kurukulaaratchy, DM, FRCP,[a] Hongmei Zhang, PhD,[c] Veeresh Patil, MRCP,[a] Abid Raza, PhD,[a] Wilfried Karmaus, MD, DrMed, MPH,[c] Susan Ewart, DVM, PhD,[d] and S. Hasan Arshad, DM, FRCP[a,b]

Isle of Wight and Southampton, United Kingdom, Memphis, Tenn, and East Lansing, Mich

Background: Rhinitis affects many young adults and often shows comorbidity with asthma.

Objective: We hypothesized that young adult rhinitis, like asthma, exhibits clinical heterogeneity identifiable by means of cluster analysis.

Methods: Participants in the Isle of Wight birth cohort (n = 1456) were assessed at 1, 2, 4, 10, and 18 years of age. Cluster analysis was performed on those with rhinitis at age 18 years (n = 468) by using 13 variables defining clinical characteristics.

Results: Four clusters were identified. Patients in cluster 1 (n = 128 [27.4%]; ie, moderate childhood-onset rhinitis) had high atopy and eczema prevalence and high total IgE levels but low asthma prevalence. They showed the best lung function at 18 years of age, with normal fraction of exhaled nitric oxide (FENO), low bronchial hyperresponsiveness (BHR), and low bronchodilator reversibility (BDR) but high rhinitis symptoms and treatment. Patients in cluster 2 (n = 199 [42.5%]; ie, mild-adolescence-onset female rhinitis) had the lowest prevalence of comorbid atopy, asthma, and eczema. They had normal lung function and low BHR, BDR, FENO values, and total IgE levels plus low rhinitis symptoms, severity, and treatment. Patients in cluster 3 (n = 59 [12.6%]; ie, severe earliest-onset rhinitis with asthma) had the youngest rhinitis onset plus the highest comorbid asthma (of simultaneous onset) and atopy. They showed the most obstructed lung function with high BHR, BDR, and FENO values plus high rhinitis symptoms, severity, and treatment. Patient 4 in cluster 4 (n = 82 [17.5%]; ie, moderate childhood-onset male rhinitis with asthma) had high atopy, intermediate asthma, and low eczema. They had impaired lung function with high FENO values and total IgE levels but intermediate BHR and BDR. They had moderate rhinitis symptoms.

Conclusion: Clinically distinctive adolescent rhinitis clusters are apparent with varying sex and asthma associations plus differing rhinitis severity and treatment needs. (J Allergy Clin Immunol 2015;135:143-50.)

Key words: *Asthma, cluster analysis, morbidity, rhinitis, wheezing*

Rhinitis is a common condition that increases in prevalence through childhood and adolescence to affect a substantial proportion of young adults.[1] Frequently under-recognized, it can be associated with considerable morbidity and impaired quality of life.[2] Comorbidity with other allergic diseases, most notably asthma, is well documented for rhinitis.[3,4] This has led to increasing acceptance of the concept of "one airway, one disease" in clinical practice.[5] In turn, rhinitis is recognized as a risk factor for asthma development,[6-8] and rhinitis treatment might help reduce asthma morbidity.[9] Asthma itself has been increasingly characterized as a heterogeneous disorder composed of different phenotypes.[10] Recently, cluster analysis has further enhanced our understanding of the diversity of wheezing disorders. Cluster analysis has identified wheeze clusters with strong[11,12] or little[13] association with atopy, impaired lung function,[11,12] severe bronchial hyperresponsiveness (BHR),[11,14] high[11,13] or low[12] disease morbidity, and obesity.[13] Different forms of rhinitis are also being identified,[4,15,16] and although rhinitis is regarded as an allergic disorder, both atopic and nonatopic forms exist.[1] However, a deeper understanding of the potential diversity of rhinitis is yet to emerge, and to date, no cluster analysis of rhinitis has been published in the literature.

Adolescence is a crucial developmental phase that represents a period of dynamic physiologic change. It is also a period that sees substantial rhinitis incidence.[1] We hypothesized that a cluster analysis in young adulthood would demonstrate distinctive rhinitis clusters with potential clinical relevance. In this article we describe a cluster analysis of young adults with rhinitis to determine phenotypes without observer bias from the Isle of Wight Birth Cohort.

METHODS

An unselected whole-population birth cohort (n = 1456) was established on the Isle of Wight (United Kingdom) in 1989 to study the natural history of allergic disease. Participants were assessed at 1, 2, 4, 10, and 18 years of age. The methodology for the first decade of follow-up has been published previously.[17-20]

Eighteen-year follow-up methodology

The local research ethics committee (06/Q1701/34) approved follow-up at 18 years of age. Participants provided informed consent and information on respiratory, nasal, and dermatologic symptoms. Study-specific plus International Study of Asthma and Allergies in Childhood[21] questionnaires were used. Questions about specific exposures are provided in the Methods section in this article's Online Repository at www.jacionline.org.

From [a]the David Hide Asthma and Allergy Research Centre, St Mary's Hospital, Newport, Isle of Wight; [b]Clinical and Experimental Medicine, Faculty of Medicine, University of Southampton; [c]the Division of Epidemiology, Biostatistics, and Environmental Health, School of Public Health, University of Memphis; and [d]the Large Animal Clinical Sciences, Michigan State University, East Lansing.

Supported by the National Institutes of Health (grant 5 R01 HL082925) and the National Institute of Allergy and Infectious Diseases (award no. R01 AI091905).

Disclosure of potential conflict of interest: The authors received research support from the National Institutes of Health.

Received for publication February 27, 2014; revised May 11, 2014; accepted for publication June 13, 2014.

Available online July 29, 2014.

Corresponding author: S. Hasan Arshad, DM, FRCP, MP 810, South Block, Southampton General Hospital, Tremona Road, Southampton, United Kingdom. E-mail: sha@soton.ac.uk.

0091-6749/$36.00

http://dx.doi.org/10.1016/j.jaci.2014.06.017

Abbreviations used
BDR: Bronchodilator reversibility
BHR: Bronchial hyperresponsiveness
CCC: Cubic clustering criterion
DRS: Dose-response slope
FEF_{25-75}: Forced expiratory flow at 25% to 75%
FENO: Fraction of exhaled nitric oxide
FVC: Forced vital capacity
SPT: Skin prick test

Participation at age 18 years was in person, by telephone, or by post. Participants attending in person also performed spirometry, fraction of exhaled nitric oxide (FENO) measurement, methacholine challenge tests, and skin prick tests (SPTs) and provided blood for total IgE measurement. Identical methodology published previously[20] was used for spirometry and methacholine challenge testing at 10 and 18 years of age. FENO (NIOX MINO; Aerocrine AB, Solna, Sweden) measurement and SPTs to 14 common aeroallergens and food allergens (ALK-Abelló, Hørsholm, Denmark) were performed, as reported previously.[20] Relevant methodology at 10 and 18 years is summarized in the Methods section in this article's Online Repository.

Definitions

Rhinitis at 10 and 18 years of age was defined by the question "Have you ever had a problem with sneezing, runny or blocked nose in the absence of cold or flu?" plus the presence of "symptoms in the last 12 months." Seasonality of rhinitis was defined as perennial (with or without seasonal exacerbation) or seasonal only. Asthma at 10 and 18 years of age was defined as an answer of yes to "ever had asthma" and either of "wheezing in the last 12 months" or "asthma treatment in the last 12 months." Eczema at 10 and 18 years was defined by the question "Have you ever been diagnosed with eczema?" plus "having an itchy rash in the past 12 months." Diagnostic definitions used at early-life follow-up are presented in the Methods section in this article's Online Repository. Atopy was defined by a positive SPT response (mean wheal diameter, 3 mm ≥ negative control) to at least 1 allergen. Age of rhinitis onset was defined as "childhood" (<12 years of age) or "adolescence" (≥12 years of age).

Bronchodilator reversibility (BDR) was defined as a percentage change in FEV_1 after inhaling 600 μg of salbutamol. Not all subjects undergoing methacholine challenge testing demonstrate a 20% decrease in FEV_1 that enables PC_{20} calculation to indicate BHR. Therefore at both 10 and 18 years of age, a continuous dose-response slope (DRS) measure of BHR was estimated by using least-square regression of the percentage change in FEV_1 on the cumulative methacholine dose for each child. The DRS was transformed as $\log_{10}$ (DRS+10) to satisfy normality and homoscedasticity, with higher values inferring greater BHR.

Statistical methods

All statistical measures were performed with the SAS statistical package, version 9.3 (SAS Institute, Cary, NC). Cluster analysis was performed on the population reporting rhinitis within the past 12 months at age 18 years (n = 468). Cluster variables were selected that defined clinical characteristics at 18 years of age. These included both questionnaire-derived data and variables from objective testing. Thirteen variables were selected for the cluster analyses: atopic status at age 18 years, asthma at age 18 years, eczema at age 18 years, age at which rhinitis appeared, seasonality of rhinitis symptoms, total IgE level ($\log_{10}$), BDR, BHR DRS, mean FENO value ($\log_{10}$), FEV_1, forced vital capacity (FVC), FEV_1/FVC ratio, and forced expiratory flow at 25% to 75% (FEF_{25-75}).

As in linear regressions, clustering methods are sensitive to the scale of the variables. Each variable was standardized by subtracting its mean and dividing by its SD to put every continuous variable on a common scale. To remove the effect of sex and height on FEV_1, FVC, and FEF_{25-75}, we regressed these variables on sex and height; the residuals with sex and height effects excluded were used in the analyses. We used this approach in preference to using percent predicted lung function values. The use of percent predicted values is dependent on values estimated from existing formulas based on reference populations. Those reference ranges might not apply to all populations and would not necessarily reflect the relevance of individual lung function measures at the age studied in our cohort. Because of the skewness of total IgE and FENO measures in the original scale, those variables were $\log_{10}$ transformed to improve normality and homoscedasticity. Correlations were calculated among the continuous variables (see Table E1 in this article's Online Repository at www.jacionline.org), between continuous and binary variables (see Table E2 in this article's Online Repository at www.jacionline.org), and between binary variables (see Table E3 in this article's Online Repository at www.jacionline.org).

The method of K-means (PROC FASTCLUS in SAS) was used for cluster analysis. To determine cluster number, we used 2 criteria: the cubic clustering criterion (CCC)[22] and the pseudo F statistic.[23] The CCC criterion is a measure of cluster deviation from the distribution expected if data points were drawn from a uniform distribution. The pseudo F statistic captures cluster homogeneity and is a ratio of the mean sum of squares between groups to the mean sum of squares within group. Larger CCC and pseudo F values indicate a better cluster solution. In our analysis we considered different numbers of clusters and then for each case obtained values for CCC and pseudo F. In general, the patterns of CCC and pseudo F values were quadratic with respect to the number of clusters. The final choice for number of clusters was determined by an overall evaluation of CCC and pseudo F combined with an R^2 statistic measuring between-cluster variations (the larger the better, see Table E4 in this article's Online Repository at www.jacionline.org).

ANOVA was used for continuous variables and χ^2 tests were used for binary variables to evaluate each clustering variable across the clusters. Pairwise *t* tests with Bonferroni multiple testing correction were applied to continuous variables, and pairwise proportion tests were applied to binary variables, with the overall significance level set at an α value of .05. Having defined distinct clusters, we sought to further characterize those clusters by assessing morbidity parameters for them. These included rhinitis symptom frequency and severity and rhinitis therapy with assessment by using χ^2 analysis for each parameter across the clusters. Potential associations with risk factors recorded prospectively during the lifetime of the cohort were then assessed. These included male sex, family history (parent or sibling) of rhinitis, cord blood IgE level, low (<2.5 kg) birth weight, maternal smoking in pregnancy, exclusive breast-feeding in the first 3 months of life, recurrent chest infections in infancy, past or current personal smoking, paracetamol use at 18 years of age, and body mass index at 18 years of age. Identical methods as for morbidity variables were used to compare risk factors between different clusters.

RESULTS

High cohort follow-up was achieved at 10 (94%, n = 1373) and 18 (90%, n = 1313) years of age. Not all subjects participated at all visits. Of the overall initial cohort of 1456 subjects, 16.8% (n = 210) were seen at 10 but not 18 years of age, and 5.4% (n = 80) were seen at 18 but not 10 years of age. Of subjects seen for a center visit at 18 years of age (n = 864), 90.5% (n = 762) were also seen at 10 years of age. Previously published data[24] demonstrated that participants attending the center for a "full visit" (n = 864) at 18 years of age did not differ significantly from the overall cohort participation (see Table E5 in this article's Online Repository at www.jacionline.org). Statistical summary of standardized variables is shown in Table E6 in this article's Online Repository at www.jacionline.org. At 18 years of age, whole population prevalence of diagnosed rhinitis was 35.8% (468/1309), that of asthma was 17.9% (234/1306), that of eczema was 12.3% (161/1306) and that of atopy was 41.3% (352/853).