

Phenotypes determined by cluster analysis in severe or difficult-to-treat asthma

Michael Schatz, MD, MS,^a Jin-Wen Y. Hsu, PhD,^a Robert S. Zeiger, MD, PhD,^a Wansu Chen, MS,^a
Alejandro Dorenbaum, MD,^b Bradley E. Chipps, MD,^c and Tmirah Haselkorn, PhD^b *San Diego, Pasadena, South San Francisco, and Sacramento, Calif*

Background: Asthma phenotyping can facilitate understanding of disease pathogenesis and potential targeted therapies.

Objective: To further characterize the distinguishing features of phenotypic groups in difficult-to-treat asthma.

Methods: Children ages 6-11 years (n = 518) and adolescents and adults ages ≥12 years (n = 3612) with severe or difficult-to-treat asthma from The Epidemiology and Natural History of Asthma: Outcomes and Treatment Regimens (TENOR) study were evaluated in this *post hoc* cluster analysis. Analyzed variables included sex, race, atopy, age of asthma onset, smoking (adolescents and adults), passive smoke exposure (children), obesity, and aspirin sensitivity. Cluster analysis used the hierarchical clustering algorithm with the Ward minimum variance method. The results were compared among clusters by χ^2 analysis; variables with significant ($P < .05$) differences among clusters were considered as distinguishing feature candidates. Associations among clusters and asthma-related health outcomes were assessed in multivariable analyses by adjusting for socioeconomic status, environmental exposures, and intensity of therapy.

Results: Five clusters were identified in each age stratum. Sex, atopic status, and nonwhite race were distinguishing variables in both strata; passive smoke exposure was distinguishing in children and aspirin sensitivity in adolescents and adults.

Clusters were not related to outcomes in children, but 2 adult and adolescent clusters distinguished by nonwhite race and aspirin sensitivity manifested poorer quality of life ($P < .0001$), and the aspirin-sensitive cluster experienced more frequent asthma exacerbations ($P < .0001$).

Conclusion: Distinct phenotypes appear to exist in patients with severe or difficult-to-treat asthma, which is related to outcomes in adolescents and adults but not in children. The study of the therapeutic implications of these phenotypes is warranted. (J Allergy Clin Immunol 2013;■■■:■■■-■■■.)

Key words: TENOR, severe asthma, difficult-to-treat asthma, phenotype, demographics, atopy, cluster analysis

From the ^aDepartments of Allergy and of Research and Evaluation, Kaiser Permanente Southern California, San Diego and Pasadena; ^bUS Medical Affairs, Genentech Inc, South San Francisco; and ^cCapital Allergy and Respiratory Disease Center, Sacramento.

TENOR (The Epidemiology and Natural History of Asthma: Outcomes and Treatment Regimens) was funded by Genentech Inc and Novartis Pharmaceuticals Corporation. This analysis was partially funded by the Department of Research and Evaluation within the Southern California Permanente Medical Group with funds from the Medical Group and the Kaiser Permanente Community Benefit Fund. Editorial support was provided by CircleScience, funded by Genentech Inc.

Disclosure of potential conflict of interest: M. Schatz has received research support and provision of writing assistance from Genentech; has received consultancy fees from Amgen, GlaxoSmithKline, Merck, and Boston Scientific; and has received research support from Aerocrine, Merck, Genentech, GlaxoSmithKline, and MedImmune. R. S. Zeiger has received consultancy fees from Aerocrine, AstraZeneca, Genentech, GlaxoSmithKline, MedImmune, Schering Plough, Sunovion, Novartis, and the National Heart, Lung, and Blood Institute/Penn State; and has received research support from Genentech, GlaxoSmithKline, Aerocrine, Merck, MedImmune, and Thermofisher. A. Dorenbaum is employed by and has stock in Genentech; and has an appointment at Stanford University in Allergy Immunology. B. E. Chipps has received consultancy fees from Genentech, AstraZeneca, GlaxoSmithKline, SRx, Novartis, Sunovion, Merck-Schering, ISTA, and Dey; has received speakers' bureau honoraria from Genentech, AstraZeneca, GlaxoSmithKline, Novartis, Sunovion, Merck-Schering, and ISTA; has received payment for the service on speakers' bureaus for Genentech, AstraZeneca, GlaxoSmithKline, Novartis, Sunovion, Merck-Schering, and ISTA. T. Haselkorn has received consultancy fees for writing/reviewing the manuscript from Genentech. The rest of the authors declare that they have no relevant conflicts of interest.

Received for publication December 4, 2012; revised August 16, 2013; accepted for publication October 8, 2013.

Corresponding author: Michael Schatz, MD, MS, Department of Allergy, Kaiser Permanente Medical Center, 7060 Clairemont Mesa Blvd, San Diego, CA 92111.

E-mail: michael.schatz@kp.org.

0091-6749/\$36.00

© 2013 American Academy of Allergy, Asthma & Immunology

<http://dx.doi.org/10.1016/j.jaci.2013.10.006>

Asthma is a heterogeneous disease that presents with diverse symptom profiles and variable response to medications.^{1,2} Identification of asthma phenotypes may improve our understanding of disease pathogenesis and facilitate potential targeted therapies.³ Several studies have used statistical techniques, particularly cluster analysis, to identify asthma phenotypes.⁴⁻¹⁴ Results have varied, in part, based on different patient populations and differences in the phenotyping variables assessed. Patients with difficult-to-treat asthma have high rates of asthma exacerbations and poor asthma control despite the use of multiple inhaled and oral control medications.¹⁵ The Severe Asthma Research Program (SARP) has examined phenotype clusters in adults⁷ and children¹⁰ with severe asthma, but no study has yet characterized the clinical features of potential phenotypic groups in patients identified as being "difficult to treat." The Epidemiology and Natural History of Asthma: Outcomes and Treatment Regimens (TENOR) study is the largest prospective, observational cohort of patients with high medication or health care system use, and more than 95% of the patients were classified by their physicians as difficult to treat. The TENOR database contains a broad spectrum of patient-reported and validated objective measures.¹⁶ The primary objective of this study was to identify and characterize clusters that represent potential phenotypes in the unique TENOR cohort. The secondary objective was to examine the relationships between these potential asthma phenotypes and asthma-related health outcomes.

METHODS

Overall TENOR study design

Full study design details of the TENOR study have been reported elsewhere.^{16,17} TENOR was a 3-year (2001-2004), multicenter, prospective

Abbreviations used

ATAQ:	Asthma Therapy Assessment Questionnaire
BMI:	Body mass index
FVC:	Forced vital capacity
IgE:	Immunoglobulin E
QoL:	Quality of life
SARP:	Severe Asthma Research Program
TENOR:	The Epidemiology and Natural History of Asthma: Outcomes and Treatment Regimens
VPC:	Very poorly controlled

cohort study of 4756 patients ages 6 years or older with severe or difficult-to-treat asthma.¹⁶ Asthma severity was based on the physician's subjective clinical assessment and was considered difficult to treat based on a number of prespecified parameters, including the following: the need for multiple asthma medications, severe exacerbations, frequent exacerbations, unable to avoid triggers, and a complex medication regimen (multiple reasons were allowed). Patients with mild or moderate asthma were eligible for enrollment if their asthma was considered difficult to treat by the evaluating physician and they met the additional inclusion criteria.¹⁶

Patients selected by their physicians for the TENOR study had either high use of the health care system (2 or more unscheduled care visits for asthma or 2 or more oral corticosteroid bursts for exacerbations) or high medication use (3 or more medications to control asthma, high-dose inhaled corticosteroids, or regular oral prednisone) in the past 12 months. Patients were enrolled from specialty allergist and pulmonologist sites that represented diverse geographic areas in the United States, including managed care organizations, academic institutions, and community or group practices. No treatment intervention was involved; all enrolled patients continued to receive asthma medication as directed by their asthma specialist.¹⁶ Data were collected at semiannual visits. The TENOR study design and protocol were approved by a central institutional review board and, as necessary, by individual institutional review boards. Written consent was obtained from all patients.

Clinical variables

As applied to asthma, the term "phenotype" has been defined as "the visible characteristics of an organism that result from the interaction between the genetic makeup and the environment."³ We thus have used the following strategy for the current analysis. Characteristics likely related to genetic or epigenetic factors were used as potential baseline cluster-defining variables; follow-up control or severity markers have been evaluated as outcome variables; and other exposure or treatment factors that could affect the outcomes, but that would not represent inherent phenotypic characteristics, have been considered as potential confounders. We did include baseline pulmonary function as a cluster variable because fixed airway obstruction may occur early in life,¹⁸ which suggests that it may be an inherent rather than an acquired characteristic of asthma.

Cluster-defining variables were chosen from the TENOR baseline variables based on clinical meaningfulness and minimal missing data. To reduce collinearity and the number of variables for the cluster analysis, both exploratory factor analysis and variable clustering were performed to divide the variables into different domains. One representative variable from each domain was chosen. The final set of variables used in cluster analyses includes (1) sex; (2) race (white vs nonwhite); (3) markers of atopy, including the number of self-reported allergic triggers (pollen, animals, mold, dust), history of allergic rhinitis or atopic dermatitis, and serum total IgE level; (4) age at asthma onset; (5) smoking status (adults) or passive smoke exposure (children); (6) body mass index (BMI); (7) self-reported history of aspirin sensitivity; and (8) ratio of FEV₁ to forced vital capacity (FVC) (FEV₁:FVC) from spirometry performed according to American Thoracic Society guidelines.¹⁹ Sex, race, smoking status, passive smoke exposure, and history of allergic rhinitis, atopic dermatitis, and aspirin sensitivity were binary variables (0-1). BMI, age of onset, serum IgE, and FEV₁:FVC were used as

continuous variables in the cluster analyses but were later grouped into categories for the cluster profiling to facilitate easier clinical interpretation. Cluster analysis variables were standardized by using the range standardization method with a scale between 0 and 1, so that the scale for continuous variables was the same as the scale for binary variables. Thus, the minimum value for transformed continuous variables was 0 and the maximum was 1.

Baseline confounding variables evaluated were the following: (1) health care coverage (commercial vs other), (2) median income by ZIP code, (3) education (some college vs high school or less in adults only), (4) cats or dogs in home (1 or more vs none), (5) moisture or dampness in home (yes vs no), (6) visible mold (yes vs no), (7) cockroaches seen (yes vs no), (8) secondhand smoke (one or more smokers in household vs none in adults only), and (9) the number of long-term asthma controllers (≥ 3 vs 2 vs 0-1).

Outcome variables evaluated were the following: (1) asthma exacerbations (any hospitalization, emergency department visit, or corticosteroid burst in the past 3 months reported at month 6 or month 12); (2) risk of very poorly controlled (VPC) asthma at month 12, determined by the study questionnaire based on Expert Panel Report 3 criteria²⁰; (3) number of asthma control problems from the Asthma Therapy Assessment Questionnaire (ATAQ)²¹ at month 12, stratified for the current analysis as 0-1 versus 2 or more; and (4) asthma-related quality of life (QoL) from the Juniper Mini-Asthma Quality of Life Questionnaire²² obtained at month 12, in which overall scores range from 1 to 7, with higher scores indicating better asthma-related QoL.

Statistical analysis

Cluster analysis was performed by using a hierarchical clustering algorithm with the Ward minimum-variance method.²³ At each generation of clusters, the Ward method joined the clusters so that the within-cluster variation was minimized. Cluster analysis was performed separately in children (ages 6-11 years at baseline) and adolescents and adults (ages 12 years and older at baseline). Plots of fit statistics (including CCC, pseudo-F Statistic, Pseudo-T2 Statistic) and dendrograms were produced and were examined to help to determine the numbers of clusters. If statistics suggested several options, a decision was made based on clinical considerations. To compare differences among clusters, ANOVA and χ^2 tests were used for continuous and categorical variables, respectively, with a Tukey-Kramer adjustment for multiple comparisons. Variables with significant ($P < .05$) differences among clusters were considered as distinguishing feature candidates.

Different modeling techniques were used to examine the cluster effect for each of the study outcomes, with and without adjustment for potential confounding variables. Multivariable logistic regression analysis was used to examine the risk of asthma exacerbations, VPC asthma, and the number of ATAQ control barriers (dichotomized at 0-1 vs 2 or more). C-index and Hosmer-Lemeshow tests were used to show goodness of fit for the logistic regression models. Linear regression analyses were performed to evaluate QoL outcome variables. For the multivariable analyses, a reference cluster was chosen based on the lowest proportion of exacerbations and VPC asthma, the highest proportion of no ATAQ control barriers, and the highest mean QoL. In the absence of a cluster meeting all of these criteria, the cluster with the highest number of these features was chosen as the reference group so that the same cluster would be the reference group for all models. From each model, an odds ratio or least-square mean difference, 95% CI, and Tukey-Kramer adjusted P value were computed for each cluster compared with the reference cluster. A nominal 2-tailed statistical significance for each analysis was set at an overall α level of .05. All analyses were conducted by using SAS version 9.2 (SAS Institute Inc, Cary, NC).

RESULTS

From 4756 patients enrolled in TENOR, a total of 4130 (86.8%) with no missing variables were included in the study: 518 children ages 6-11 years and 3612 adolescents and adults ages 12 years and older (Table I). Approximately two-thirds of the pediatric cohort were male patients, whereas two-thirds of the adolescent and adult cohort were female patients. The majority

TABLE I. Baseline characteristics of study subjects

Characteristic	Children, no. (%) (n = 518)	Adolescents and adults, no. (%) (n = 3612)
Sex		
Female	169 (32.6)	2402 (66.5)
Male	349 (67.4)	1210 (33.5)
Race		
White	316 (61.0)	2789 (77.2)
Nonwhite	202 (39.0)	823 (22.8)
Age of onset		
<3 y	290 (56.0)	600 (16.6)
3-11 y	228 (44.0)	923 (25.6)
12-17 y	—	260 (7.2)
≥18 y	—	1829 (50.6)
Obese		
Yes*	159 (30.7)	2071 (57.3)
No	359 (69.3)	1541 (42.7)
Passive smoke exposure		
Yes	133 (25.7)	—
No	385 (74.3)	—
Active smoking		
Yes	—	1232 (34.1)
No	—	2380 (65.9)
Atopy		
≥3 Allergic triggers	220 (42.5)	2084 (57.7)
% Allergic rhinitis	356 (68.7)	2778 (76.9)
% Atopic dermatitis	172 (33.2)	793 (22.0)
Total IgE		
≥100 IU/mL	145 (28.0)	784 (21.7)
<100 IU/mL	373 (72.0)	2828 (78.3)
Aspirin sensitivity		
Yes	—	432 (12.0)
No	—	3180 (88.0)
Abnormal FEV ₁ :FVC ratio†		
Yes	342 (66.0)	1828 (50.6)
No	176 (34.0)	1784 (49.4)

*Defined as BMI >30 kg/m² in adolescents and adults, >95th percentile in children.

†Based on predicted normal values for age (from the National Asthma Education and Prevention Program²⁰).

of patients were white, but the proportion of nonwhites in the pediatric cohort was almost twice as high compared with the adolescents and adults. The age of asthma onset was <3 years in more than half of children and ≥18 years in half of adults. The majority of adolescents and adults were obese (58%), which was almost double the percentage of children who were obese (31%). One-fourth of the children reported passive smoke exposure, and one-third of adults were smokers. The majority of patients reported a history of allergic rhinitis (69% to 77%). A minority reported atopic dermatitis, which was more common in children (33%) than in adolescents and adults (22%), and high IgE levels (≥100 IU/mL) were observed in 22% to 28% of the patients. Aspirin sensitivity was reported by 12% of adolescents and adults. Children were more likely to manifest abnormal FEV₁:FVC ratios (66.0%) than were adolescents and adults (50.6%).

Children clusters

Cluster analysis of the children revealed 5 clusters (Table II). Cluster 1 included atopic white male patients with no smoke exposure. Other clusters were distinguished by female sex (cluster 2), nonatopic status (cluster 3), passive smoke exposure

(cluster 4), and nonwhite race (cluster 5). Clusters 1 and 2 manifested less airway obstruction than the remaining 3 clusters. All variables except age of onset and obesity were significantly different among the clusters, although mean BMI was higher in cluster 4 than in cluster 1. For children, the reference group for the multivariable analyses was chosen to be cluster 1 based on lowest exacerbation rate and highest proportion of no ATAQ control barriers (see Table E1 in this article's Online Repository at www.jacionline.org). No significant relationships were demonstrated among cluster membership and outcomes in children (see Table E2 in this article's Online Repository at www.jacionline.org).

Adolescent and adult clusters

Cluster analysis in adolescents and adults also revealed 5 clusters (Table III). The most common cluster (approximately 35% of patients [cluster 1]) included white female patients most likely to have adult onset, without aspirin sensitivity, and who had lower total IgE levels. Other clusters were distinguished by the highest atopy, especially atopic dermatitis (cluster 2), male sex (cluster 3), nonwhite race (cluster 4), and aspirin sensitivity (cluster 5). All variables tested were significantly different in some pairwise comparisons among clusters, although the magnitudes of the differences were smaller for smoking and obesity than for the other variables. Cluster 3 had the highest proportion of patients with abnormal baseline pulmonary function and the lowest mean FEV₁:FVC ratio.

Among adolescents and adults, the reference group for the multivariable analyses was cluster 3, based on the lowest exacerbation rate, lowest proportion of VPC asthma, highest proportion of no ATAQ barriers, and highest mean QoL (Table E1). Patients in cluster 5 had a 75% increased risk of exacerbations (Fig 1) compared with patients in cluster 3. Patients in clusters 4 and 5 had >50% increased risks of poorly controlled asthma (Fig 2, A), but these results did not reach the statistical significance level required for multiple comparisons. The cluster 3 reference group demonstrated better asthma control based on the ATAQ than all other clusters except cluster 2 (Fig 2, B) and better asthma-related QoL than all other clusters (Fig 3). Clusters 4 and 5 had worse QoL than any of the other clusters (Fig 3). Although cluster 5 membership was a significant predictor of exacerbations and cluster 3 membership was a significant predictor of a lower number of control barriers, the relatively low c-statistics of those models (0.566-0.585) demonstrated relatively low discrimination of the models overall.

DISCUSSION

Identifying asthma phenotypes can facilitate increased understanding of disease pathogenesis and the identification of more targeted therapies.³ Understanding these factors in the difficult-to-treat asthmatic population is particularly important due to the high humanistic and financial cost of their disease.¹⁵ The current study used cluster analysis to show that clinical phenotypes can be defined within the population of patients with severe or difficult-to-treat asthma based on sex, race, atopy, aspirin sensitivity (adolescents and adults), and passive smoke exposure (children). Sex, race, and atopy are clearly inherent characteristics with a strong genetic basis. Aspirin sensitivity may have a genetic component and reflects intrinsic altered

TABLE II. Characteristics of 5 clusters in 518 children ages 6-11 years

Feature	Cluster*				
	1	2	3	4	5
No. patients (%)	115 (22.2)	81 (15.6)	162 (31.3)	87 (16.8)	73 (14.1)
% Boys	100†	0†	64.2	65.5	100†
% White race	100†	58.0	58.0	69.0	0†
% Onset <3 y old	46.1	54.3	58.0	57.5	67.1
Obese					
% BMI > 95th percentile	20.0	28.4	34.6	35.6	35.6
Mean (SD) BMI	18.8 ± 3.7‡	19.9 ± 5.2	20.4 ± 5.2	22.2 ± 10.4§	20.2 ± 4.5
% Passive smoke exposure	0†	0†	28.4	100†	0†
Atopy					
≥3 Allergic triggers	49.6	39.5	25.3‡§¶	57.5	54.8
% Allergic rhinitis	100†	100†	0†	100†	100†
% Atopic dermatitis	40.0	37.0	2.8‡§¶#	36.8	37.0
IgE					
% Total IgE ≥100	73.0	64.2	60.5	73.6	80.8
Mean (SD) IgE	505 ± 792	393 ± 520	469 ± 722	787 ± 1593	757 ± 1180
FEV ₁ :FVC ratio					
% <normal for age	57.4	50.6	73.5	72.4	72.6
Mean (SD) FEV ₁ :FVC	82.2 (9.5) ¶	83.4 (9.3) ¶	78.0 (10.6)§#	79.2 (11.0)	77.7 (10.9)§#

*Features that most distinguish one cluster from others in boldface.

†When a cluster has a boundary value (0, 100) for an item, this item is different by definition, and it is not a question of whether it may be different due to random chance, thus, a *P* value is not applicable.

‡Statistically significantly different from cluster 4 at the *P* < .05 level.

§Statistically significantly different from cluster 1 at the *P* < .05 level.

||Statistically significantly different from cluster 3 at the *P* < .05 level.

¶Statistically significantly different from cluster 5 at the *P* < .05 level.

#Statistically significantly different from cluster 2 at the *P* < .05 level.

TABLE III. Characteristics of 5 clusters in 3612 adolescents and adults ages ≥12 years

Feature	Cluster*				
	1	2	3	4	5
No. patients (%)	1262 (34.9)	659 (18.2)	664 (18.4)	596 (16.5)	431 (11.9)
% Girls and women	100†	67.5‡	0†	61.1‡	76.8§
% White race	100†	83.2‡	100†	0†	73.1§
% Onset ≥18 years old	61.0§ ¶	38.5‡#	43.7‡#	44.5‡#	58.0§ ¶
BMI					
% Obese (BMI > 30 kg/m ²)	57.5	52.1	54.8	62.6	61.7
Mean (SD) BMI	30.0 ± 7.9	29.0 ± 8.0‡¶	28.3 ± 6.6‡¶#	30.4 ± 8.4§	30.6 ± 8.0§
% Smokers	35.7	27.5	36.8§	36.7	31.8
Atopy					
≥3 Allergic triggers	56.7‡§	68.9 ¶#	47.9‡§#	51.5‡§	67.1 ¶#
% Allergic rhinitis	75.9‡§¶	85.1 ¶#	74.0‡§	68.6‡§#	83.3 ¶#
% Atopic dermatitis	0†	100†	0†	5.5‡	23.4¶
IgE					
% Total IgE ≥100	25.3	42.2	41.6	48.7	30.4
Mean (SD) IgE	204 ± 437§ ¶	512 ± 1302‡ ¶#	349 ± 619‡§ #	440 ± 1168‡#	289 ± 705§¶
% Aspirin sensitive	0†	0.2	0†	0†	100†
FEV ₁ :FVC ratio					
% <normal for age	43.0	49.5	61.8	54.2	52.7
Mean (SD) FEV ₁ :FVC	72.5 ± 11.7‡	72.9 ± 12.1‡	69.1 ± 12.0§¶#	71.6 ± 12.5	70.1 ± 12.1§#

*Features that most distinguish one cluster from others are in boldface.

†When a cluster has a boundary value (0, 100) for an item, this item is different by definition, and it is not a question of whether it may be different due to random chance, thus, a *P* value is not applicable.

‡Statistically significantly different from cluster 5 at the *P* < .05 level.

§Statistically significantly different from cluster 2 at the *P* < .05 level.

||Statistically significantly different from cluster 3 at the *P* < .05 level.

¶Statistically significantly different from cluster 4 at the *P* < .05 level.

#Statistically significantly different from cluster 1 at the *P* < .05 level.

physiology. Although passive smoke exposure is an environmental factor, there is evidence to suggest that tobacco smoke may have epigenetic effects in promoting disease

development.^{24,25} Therefore, all of these distinguishing features can be considered to be stable individual characteristics appropriate for use in defining clinical phenotypes.

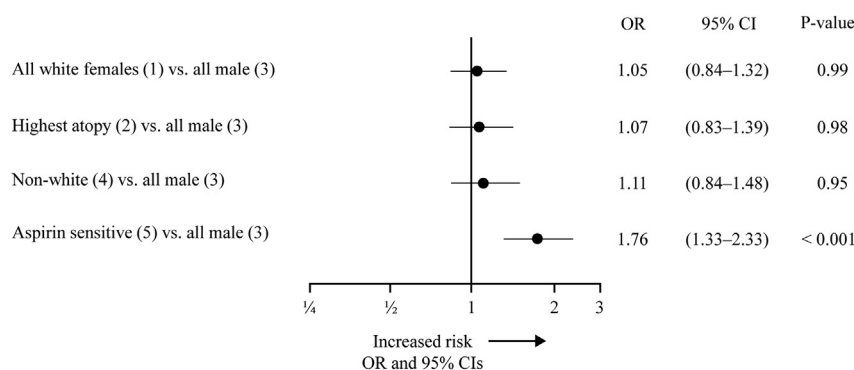


FIG 1. The relationship of cluster membership to the risk of asthma exacerbations (hospitalization, emergency department visit, or steroid burst) identified at month 6 or month 12 visits in adolescents and adults ($n = 2691$). P values were adjusted for multiple comparisons; $P < .05$ was considered statistically significant. C-index, 0.585; Hosmer and Lemeshow Test, 13.94, $P = .08$. Cluster numbers are shown in brackets after the text. *OR*, Odds ratio.

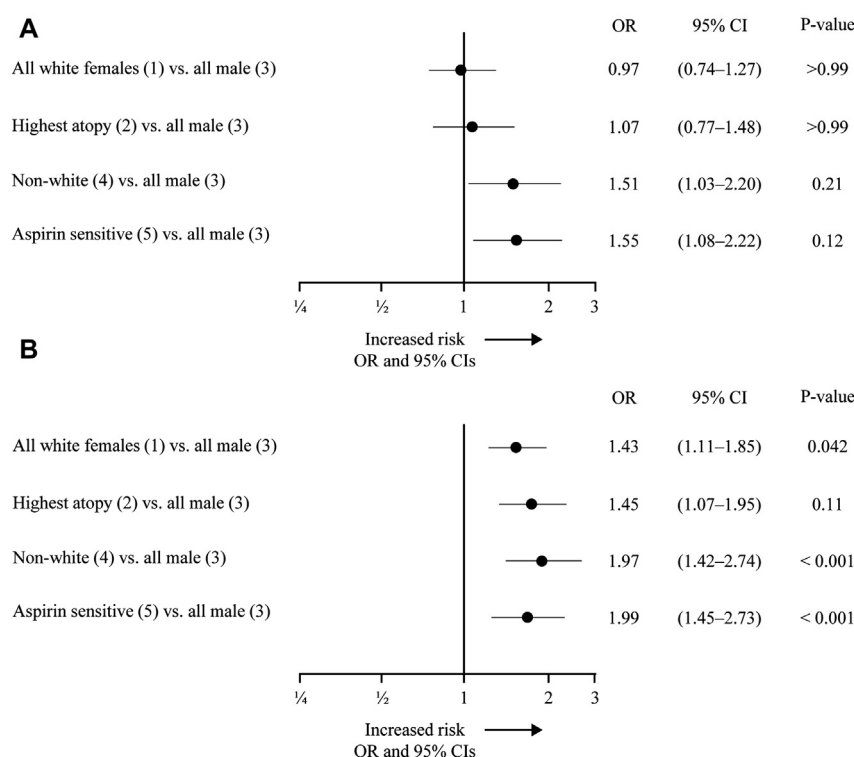


FIG 2. The relationship of cluster membership to asthma control in adolescents and adults, assessed at month 12 visit. P values were adjusted for multiple comparisons; $P < .05$ was considered statistically significant. Cluster numbers are shown in brackets after the text. **A**, Risk of VPC asthma, as defined in the Expert Panel Report 3²⁰ ($n = 1779$); C-index, 0.639; Hosmer and Lemeshow Test, 5.02, ($P = .75$). **B**, Risk of a higher number of asthma control barriers, as defined by the ATAQ²¹ ($n = 2156$); C-index, 0.566. *OR*, Odds ratio.

This study also took the novel approach of relating the baseline phenotypes to subsequent health outcomes, including asthma exacerbations, asthma control, and QoL. In adolescents and adults, the fifth cluster of patients who were aspirin sensitive, primarily white, female, and atopic with late-onset asthma were significantly more likely than any of the other clusters to experience exacerbations (Fig 1). These patients also were more likely to exhibit VPC asthma (Fig 2, A) and had worse QoL (Fig 3) than 3 of the other 4 clusters. These findings were

consistent with prior studies that showed increased asthma severity in patients with aspirin sensitivity.^{26,27} Because the fifth cluster was predominantly female patients, these results were also consistent with prior data that showed an increased risk of exacerbations and lower QoL in adult female patients versus male patients.^{28,29}

The fourth cluster, of patients who were non-white, less likely to be atopic, with high total IgE, and mostly early onset asthma, demonstrated poorer QoL (Fig 3) and a trend toward more

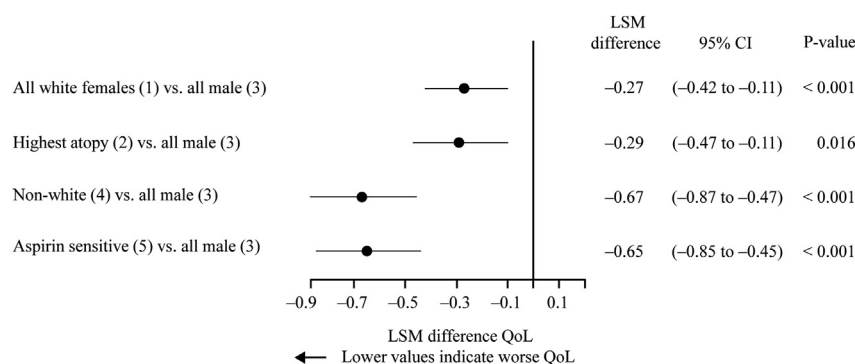


FIG 3. The relationship of cluster membership to differences in asthma-related QoL as assessed by the mini-Asthma Quality of Life Questionnaire²² at month 12 visit in adolescents/adults ($n = 2170$). P values were adjusted for multiple comparisons; $P < .05$ was considered statistically significant. $R^2 = 0.077$. Cluster numbers are shown in *brackets* after the text. *LSM*, Least squares mean.

frequent VPC asthma (Fig 2, A) compared with all of the other clusters except the patients with aspirin sensitivity. These results are consistent with the worse outcomes, including poorer asthma control and QoL, frequently reported in more general populations of nonwhite versus white patients, particularly in black patients,³⁰⁻³² who made up 65% of the nonwhite sample for this study. These outcome analyses were adjusted for socioeconomic status and environmental factors, which made it less likely that those confounding factors explained the poorer outcomes for this nonwhite cluster. We cannot differentiate reporting and/or perception differences versus disease differences as an explanation for the reduced self-reported QoL in the nonwhite cluster, but the increased exacerbations and mortality experienced by black patients^{31,32} suggest to us that the reduced QoL represents differences in actual disease impact, not just reporting or perception.

The all male adult reference cluster had the best asthma control and QoL (Table E1) but the worst pulmonary function (Table III). This is consistent with observations in the general population of adult patients with asthma, that male patients have a better QoL but lower pulmonary function than female patients.³³ It should be noted that the differences among clusters in outcomes were smaller than the differences among clusters in phenotypic features. This is most likely related to the selection of patients with prior poor outcomes for the TENOR cohort.

For children, although distinct clusters were observed with robust differences among them (Table II), cluster membership was not significantly related to subsequent health outcomes. The reasons for this are unclear. The sample size was smaller and significant differences were harder to achieve. Another explanation is that the various phenotypes may be equally severe regarding clinical outcomes in children with severe or difficult-to-treat asthma, even though the phenotypes may reflect varying pathogenetic mechanisms, as discussed below.

If phenotypes are stable and universal, then they should be reproducible in different studies and populations. Several prior studies have used cluster analysis to define phenotypes in patients with asthma.⁴⁻¹⁴ However, it is difficult to compare these studies with our study because of a number of important differences. Three studies included patients with asthma among patients with other respiratory symptoms.^{4,6,8} Most of the remaining studies included candidate phenotyping variables related to asthma severity, control, or health care utilization that

we used as outcome variables,^{5,7,10-13} and 1 study used only a history of exacerbations as the clustering variable.¹⁴ In addition, several studies in patients with asthma included markers of inflammation, such as fractional exhaled nitric oxide or sputum eosinophilia, as potential cluster-defining variables,^{5,10,12,13} and 1 study based the cluster analysis entirely on gene expression profiling of induced sputum.⁹ Finally, no prior cluster analysis studies included aspirin sensitivity in adults, and only 1 prior study¹² included passive smoke exposure in children as a phenotyping variable; both of these features were found to be distinguishing characteristics in the current study.

The SARP cluster analysis⁷ included patients similar to ours, and, in spite of the above differences, there were homologies among our findings and theirs for 4 of our 5 clusters in adults (Table IV). Our first cluster was similar to their third cluster except for obesity, which, although it was not more frequent than in other clusters in the current study, was still present in the majority of patients in this cluster. Our fourth cluster was similar to their fourth cluster regarding nonwhite race and severity, although the patients in the SARP study were more atopic than our patients in cluster 4. Our aspirin-sensitive fifth cluster was similar to their fifth cluster in several ways, although the SARP report did not describe aspirin sensitivity. It was harder to find homologies between our findings and the SARP findings in children.¹⁰ However, our most-severe fifth cluster (based on QoL) was similar to their most severe fourth cluster (based on pulmonary function and health care utilization) in nonwhite race prevalence (90% in SARP, 100% in the current TENOR analysis) as well as male predominance, early onset disease, and a high prevalence of atopic markers. Universal phenotypes can be identified across populations only if future studies use the same characteristics as candidate variables.

Defining phenotypes among patients with difficult-to-treat asthma may help gain insight into disease pathogenesis. The identification of race, sex, and atopy as defining characteristics in both age groups suggests genetic mechanisms, although the sex differences could be explained by hormonal differences as well, particularly in adolescents and adults. Pregnancy factors and early environmental exposures could also interact with genetic factors in the atopy effect. As discussed above, aspirin sensitivity also could reflect genetic influences, but the full pathogenetic meaning of the aspirin-sensitive phenotype awaits further insights into the etiology of aspirin sensitivity in patients with asthma. Finally,

TABLE IV. A comparison of the clusters identified in the current TENOR cohort analysis and those identified in adults in SARP

Cluster		Similarities	Differences
TENOR (n = 3612)	SARP (n = 726)		
1	3	Female, late onset, less atopy	Obese (SARP)
2	1, 2	Female, early onset, atopic	
3	None		
4	4	Highest nonwhite, severe	Atopic (SARP)
5	5	Female, white, late-onset, severe	Aspirin sensitive (TENOR)

passive smoke exposure may lead to epigenetic effects in children with atopy that promote the development of asthma.²⁵ In this observational cohort, we were unable to evaluate if there are differences in inflammatory markers between our fourth cluster of children with passive smoke exposure and clusters 1 and 5, which manifested similar levels of atopic manifestations but had no passive smoke exposure.

To be useful, asthma phenotypes must have clinically meaningful implications. The current phenotype definitions in adolescents and adults with severe or difficult-to-treat asthma do indicate a worse clinical prognosis regarding QoL for the nonwhite and aspirin-sensitive clusters and for asthma exacerbations in the aspirin-sensitive phenotype. Patients with these phenotypes presumably could benefit from particularly intense surveillance and therapy to improve their outcomes. As noted above, the prognostic implications of the clusters in children with severe or difficult-to-treat asthma are less clear. Although we adjusted for intensity of therapy in the outcome analyses, because this is an observational study rather than a randomized controlled trial of a specific treatment, the therapeutic implications of the current cluster results cannot be determined. Randomized controlled trials of interventions in patients whose baseline phenotype is assessed and related to therapeutic responses would shed light on the treatment implications of the phenotypes defined in the current study.

Strengths of the current study include the large sample size (especially in adolescents and adults); the unique and important population of difficult-to-treat real-world patients; the prospective design; rigorous statistical methodology; use of validated control and QoL instruments; and the ability to control for potentially important socioeconomic, environmental, and treatment confounders. This study does have some potential limitations. Because the cluster analysis was not prespecified in the development of the TENOR study, potential cluster variables could only be identified *post hoc* based on availability and apparent relevance. Atopy was assessed based on patient report of physician diagnoses of allergy-related conditions and total IgE, which would not be as rigorous as actually measuring specific IgE by skin or blood testing. Pulmonary function tests were not standardized across centers, and all other information was self-reported, although validated asthma control and QoL questionnaires were used. Aspirin sensitivity was also self-reported and was not confirmed by challenge. Self-report could be variably accurate based on age, sex, or race or ethnicity, although we are not aware of data that have specifically addressed this for a history of allergy or aspirin sensitivity. As noted above, no biologic markers other than total IgE and pulmonary function were included, but this study was meant to be an analysis of clinical phenotypes. Finally, as also noted above, the observational design did not allow us to explore the treatment

implications of these phenotypes in a randomized and controlled fashion.

Finally, it should be noted that the sharpest contrasts among clusters were found for binary (yes-no) variables (eg, sex, race, passive smoke exposure, aspirin sensitivity) as opposed to continuous variables (BMI, IgE, FEV₁:FVC). Although of presumed biologic significance, this also may reflect a limitation of the cluster algorithm used in which, due to the distributions of the continuous variables and the transformation method used, more clear-cut differences were detected in yes-no variables than in continuous variables. Only studies that showed further clinical implications of these phenotypes would confirm that these clusters reflect biologic, not just mathematical, distinct entities.

In summary, we showed that cluster analysis can identify clinical phenotypes in patients with severe or difficult-to-control asthma based on sex, race, atopy, aspirin sensitivity, and passive smoke exposure. In adolescents and adults, these clusters were related to subsequent asthma outcomes, including asthma exacerbations and QoL. Further studies will be necessary to determine the generalizability of these phenotypes, their relationships to inflammatory and other biologic markers, and their treatment implications.

We thank Dave Miller, MS, and David Mink, MS, at ICON Late Phase and Outcomes Research for statistical support.

Clinical implications: Asthma is a heterogeneous disease that includes multiple phenotypes. This study specifically defines such phenotypes in adults and children with difficult-to-treat asthma, which may have pathogenetic and therapeutic implications.

REFERENCES

1. Wenzel SE. Severe asthma in adults. *Am J Respir Crit Care Med* 2005;172:149-60.
2. Busse WW, Banks-Shlegel S, Wenzel SE. Pathophysiology of severe asthma. *J Allergy Clin Immunol* 2000;106:1033-42.
3. Wenzel S. Asthma: defining of the persistent adult phenotypes. *Lancet* 2006;368:804-13.
4. 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;31:974-81.
5. Haldar P, Pavord ID, Shaw DE, Berry MA, Thomas M, Brightling CE, et al. Cluster analysis and clinical asthma phenotypes. *Am J Respir Crit Care Med* 2008;178:218-24.
6. Weatherall M, Travers J, Shirtcliffe PM, Marsh SE, Williams MV, Nowitz MR, et al. Distinct clinical phenotypes of airways disease defined by cluster analysis. *Eur Respir J* 2009;34:812-8.
7. Moore WC, Meyers DA, Wenzel SE, Teague WG, Li H, Li X, et al. Identification of asthma phenotypes using cluster analysis in the Severe Asthma Research Program. *Am J Respir Crit Care Med* 2010;181:315-23.

8. Simpson A, Tan VYF, Winn J, Svensén M, Bishop CM, Heckerman DE, et al. Beyond atopy: multiple patterns of sensitization in relation to asthma in a birth cohort study. *Am J Respir Crit Care Med* 2010;181:1200-6.
9. Baines KJ, Simpson JL, Wood LG, Scott RJ, Gibson PG. Transcriptional phenotypes of asthma defined by gene expression profiling of induced sputum samples. *J Allergy Clin Immunol* 2011;127:153-60.e9.
10. Fitzpatrick AM, Teague WG, Meyers DA, Peters SP, Li X, Li H, et al. Heterogeneity of severe asthma in childhood: confirmation by cluster analysis of children in the National Institutes of Health/National Heart, Lung, and Blood Institute Severe Asthma Research Program. *J Allergy Clin Immunol* 2011;127:382-9.e13.
11. Siroux V, Basagaña X, Boudier A, Pin I, Garcia-Aymerich J, Vesin A, et al. Identifying adult asthma phenotypes using a clustering approach. *Eur Respir J* 2011;38:310-7.
12. Mahut B, Peyrard S, Delclaux C. Exhaled nitric oxide and clinical phenotypes of childhood asthma. *Respir Res* 2011;12:65.
13. McGrath KW, Icitovic N, Boushey HA, Lazarus SC, Sutherland ER, Chinchilli VM, et al. A large subgroup of mild-to-moderate asthma is persistently noneosinophilic. *Am J Respir Crit Care Med* 2012;185:612-9.
14. Ortega H, Miller DP, Li H. Characterization of asthma exacerbations in primary care using cluster analysis. *J Asthma* 2012;49:158-69.
15. Chipps BE, Yegin A, Borish L, Wenzel S, Weiss S, Hayden ML, et al. Key findings and clinical implications from The Epidemiology and Natural History of Asthma: Outcomes and Treatment Regimens (TENOR) study. *J Allergy Clin Immunol* 2012;130:332-42.e10.
16. Dolan CM, Fraher KE, Bleecker ER, Borish L, Chipps B, Hayden ML, et al. Design and baseline characteristics of The Epidemiology and Natural History of Asthma: Outcomes and Treatment Regimens (TENOR) study: a large cohort of patients with severe or difficult-to-treat asthma. *Ann Allergy Asthma Immunol* 2004;92:32-9.
17. Chipps BE, Szeffler SJ, Simons FER, Haselkorn T, Mink DR, Deniz Y, et al. Demographic and clinical characteristics of children and adolescents with severe or difficult-to-treat asthma. *J Allergy Clin Immunol* 2007;119:1156-63.
18. Pascual RM, Peters SP. The irreversible component of persistent asthma. *J Allergy Clin Immunol* 2009;124:883-90.
19. Standardization of spirometry, 1994 update. American Thoracic Society. *Am J Respir Crit Care Med* 1995;152:1107-36.
20. National Asthma Education and Prevention Program. Expert Panel Report 3 (EPR-3): guidelines for the diagnosis and management of asthma—summary report 2007. *J Allergy Clin Immunol* 2007;120:S94-138.
21. Vollmer WM, Markson LE, O'Connor E, Frazier EA, Berger M, Buist AS. Association of asthma control with health care utilization: a prospective evaluation. *Am J Respir Crit Care Med* 2002;165:195-9.
22. Juniper EF, Guyatt GH, Cox FM, Ferrie PJ, King DR. Development and validation of the Mini Asthma Quality of Life Questionnaire. *Eur Respir J* 1999;14:32-8.
23. Ward JH Jr. Hierarchical grouping to optimize an objective function. *J Am Stat Assoc* 1963;58:236-44.
24. Chang J-C, Wang L, Chen R-F, Liu C-A. Perinatal gene-gene and gene-environment interactions on IgE production and asthma development. *Clin Dev Immunol* 2012;2012:270869.
25. Maccani MA, Knopik VS. Cigarette smoke exposure-associated alterations to non-coding RNA. *Front Genet* 2012;3:53.
26. Mascia K, Haselkorn T, Deniz YM, Miller DP, Bleecker ER, Borish L, et al. Aspirin sensitivity and severity of asthma: evidence for irreversible airway obstruction in patients with severe or difficult-to-treat asthma. *J Allergy Clin Immunol* 2005;116:970-5.
27. Chang JE, White A, Simon RA, Stevenson DD. Aspirin-exacerbated respiratory disease: burden of disease. *Allergy Asthma Proc* 2012;33:117-21.
28. Schatz M, Camargo CA Jr. The relationship of sex to asthma prevalence, health care utilization, and medications in a large managed care organization. *Ann Allergy Asthma Immunol* 2003;91:553-8.
29. Schatz M. Outcome measures in asthma management. In: Pawanker R, Holgate ST, Rosenwasser LJ, editors. *Allergy frontiers: diagnosis and health economics*. Vol 4. Tokyo: Springer; 2009. p. 507-41.
30. Lieu TA, Lozano P, Finkelstein JA, Chi FW, Jensvold NG, Capra AM, et al. Racial/ethnic variation in asthma status and management practices among children in managed Medicaid. *Pediatrics* 2002;109:857-65.
31. Grant EN, Lyttle CS, Weiss KB. The relation of socioeconomic factors and racial/ethnic differences in US asthma mortality. *Am J Public Health* 2000;90:1923-5.
32. Mosen DM, Schatz M, Gold R, Mularski RA, Wong WF, Bellows J. Medication use, emergency hospital care utilization, and quality-of-life outcome disparities by race/ethnicity among adults with asthma. *Am J Manag Care* 2010;16:821-8.
33. Osborne ML, Vollmer WM, Linton KL, Buist AS. Characteristics of patients with asthma within a large HMO: a comparison by age and gender. *Am J Respir Crit Care Med* 1998;157:123-8.

**Phenotypes Determined by Cluster Analysis in Patients with Severe or Difficult-to-Treat
Asthma**

Michael Schatz, Jin-Wen Hsu, Robert S. Zeiger, Wansu Chen, Alejandro Dorenbaum,
Bradley E. Chipps, Tmirah Haselkorn

Online Data Supplement

TABLE EI. Outcome results by cluster in children and adolescents/adults

Cluster (n)	Outcome			
	%	% Very		% ATAQ = Mean (SD) AQLQ
		Poorly Controlled	0	
<i>Children</i>				
1 (115)	36.9	47.6	45.6	6.17 (0.80)
2 (81)	44.4	21.2	23.8	6.22 (0.92)
3 (162)	43.4	50.5	27.7	6.40 (0.66)
4 (87)	53.3	36.7	15.2	5.69 (1.21)
5 (73)	51.7	61.8	13.2	6.19 (0.78)
<i>Adolescents/Adults</i>				
1 (1262)	43.8	54.3	32.3	5.17 (1.28)
2 (659)	43/8	54.5	28.6	5.19 (1.30)
3 (664)	39.9	52.0	41.9	5.54 (1.17)
4 (596)	49.2	69.2	24.7	4.78 (1.45)
5 (431)	59.2	69.0	26.2	4.71 (1.40)

Definition of abbreviations: AQLQ, Asthma Quality of Life Questionnaire; ATAQ, Asthma

Therapy Assessment Questionnaire; FEV₁, forced expiratory volume in 1 second; SD, standard deviation.

Mildest result for each outcome in each age group is bolded.

TABLE EII. Relationships of outcomes to cluster membership in children

Outcome	Results by Cluster (vs. Reference Cluster 1)			
	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Exacerbations				
OR (CI)	1.1 (0.5–2.2)	0.9 (0.5–1.6)	1.2 (0.6–2.3)	1.1 (0.5–2.2)
<i>P</i> -value	>0.99	>0.99	0.99	>0.99
Very poorly controlled				
OR (CI)	0.3 (0.1–0.8)	1.3 (0.6–2.7)	0.8 (0.3–1.9)	1.8 (0.7–4.6)
<i>P</i> -value	0.15	0.95	0.98	0.77
ATAQ barriers				
OR (CI)	1.8 (0.8–4.3)	1.1 (0.6–2.3)	2.3 (0.9–5.6)	1.6 (0.6–4.1)
<i>P</i> -value	0.64	>0.99	0.37	0.86
AQLQ				
LSM diff (CI)	0.07 (–0.36 to 0.49)	–0.12(–0.46 to 0.21)	–0.34(–0.75 to 0.07)	–0.41(–0.85 to 0.03)
<i>P</i> -value	>0.99	0.95	0.47	0.36

AQLQ, Asthma Quality of Life Questionnaire; ATAQ, Asthma Therapy Assessment

Questionnaire; CI, 95 % confidence interval; *LSM diff*, least squares mean difference; *OR*, odds ratio; *pred*, predicted.