

METHODS

All statistical analysis was performed with R 2.10.1 software.

Model building

Variable selection. Of 48 variables initially considered, 10 were excluded because of an excess of missing data (>10% of values missing). Because of the inherent strong correlation between prebronchodilator and postbronchodilator spirometric measures, we considered only one of the 2 measurements (either prebronchodilator or postbronchodilator) for FEV₁, FVC, FEV₁/FVC ratio, and peak flow. Because of the subjective nature of many of the symptom-provoking variables, such as “animal dander worsens asthma,” these variables were excluded from further analysis. After variable selection, missing values were imputed by using a *k*-nearest neighbor algorithm from the *pamr* package of Bioconductor 2.5.^{E1} Given that clustering results might be affected by differences in scale among variables,^{E2} vector normalization was performed to scale each variable to a unit vector.

Cluster analysis. To this set of 18 variables, we applied a spectral clustering algorithm, as implemented in the *specc* function of the *kernelab* package.^{E3} We specified a range of 1 to 8 clusters for the algorithm to define. The cluster centers were randomly initialized by using the spectral clustering algorithm. We used an iterative approach, adding additional variables one at a time to the model, and used the gap statistic^{E4} to calculate the optimal number of variables for partitioning of the data.

Spectral clustering with 18 baseline phenotypic characteristics yielded many models with high gap statistics over the range of cluster numbers considered; however, the maximum statistic (gap = 0.86) was observed for 5 clusters (Fig E1, A). By using leave-out-one variable cross-validation, the importance of all 18 variables in the final model was confirmed because exclusion of any single variable resulted in substantial subgroup fragmentation and inferior model performance (data not shown). We also repeated the clustering analysis, restricting ourselves to 711 self-reported white subjects (the largest ethnic subgroup), and found no difference in cluster assignment (data not shown). Hence our final model is optimal with respect to the number of variables considered and clusters defined and does not appear to be confounded by ethnic-specific phenotypic differences.

Fig E1, B, presents a heat map of the clinical phenotypes grouped by cluster. Several variables, such as the history of hay fever, AD, prior hospitalizations, and PC₂₀, segregate discretely by cluster, suggesting that these variables were primary drivers of the clustering.

Cluster validation. We assessed the influence of including individual single variables on determining the final cluster assignments. Instead of using a multivariate model with 18 variables for cluster analysis, we used each variable separately to perform the clustering. For the continuous variables, we specified the formation of 5 clusters to allow for comparison to the multivariate model. The categorical variables led to the formation of 2 clusters. To compare the univariate and multivariate approaches to clustering, we evaluated the ability of both approaches to predict future exacerbations (ie, the time to first use of oral prednisone).

To evaluate the reproducibility of our cluster assignments, we repeated the unsupervised analysis in the CAMP cohort using the clustering algorithm used in the classification studies in the SARP adult^{E5} and childhood^{E6} cohorts, with hierarchic clustering with Ward minimum distance as an agglomeration method. We used the *hclust* function of the *stats* package in R to generate 5 clusters and compared the composition of these new clusters with our original cluster assignments. We also performed an outcomes analysis with these new clusters and compared this with our original outcomes analysis.

RESULTS

Demographic, environmental, and familial determinants of cluster grouping

Descriptions of demographic, environmental, and familial clinical variables across phenotypic clusters are presented in Table E4. Although trends for higher proportions of non-Hispanic white subjects in the mildest group and black subjects in the most

severe groups were noted, these differences were not statistically significant. In contrast, enlightening differences across clusters were observed for numerous environmental and familial factors. For example, although environmental tobacco smoke exposure was reported by subjects in all 5 clusters, the prevalence was greatest among subjects in clusters 4 and 5 (ie, those with the highest baseline exacerbation rates). However, among subjects in the less severe clusters, a direct relationship between severity and smoke exposure was not observed: those with the lowest childhood smoke exposure (cluster 2, 30.2%) had higher baseline exacerbation and greater airways hyperresponsiveness than subjects in cluster 1, who had significantly higher childhood smoke exposure (39.7%), and exacerbation rates, lung function, and airways responsiveness were markedly different between clusters 1 and 3, despite very similar childhood smoke exposure rates (39.7% vs 37.6%, respectively). Similarly, although differences in aeroallergen exposure and in familial burden of both asthma and atopy were observed across the 5 phenotypic clusters, obvious linear correlations between risk factor exposure and severity of disease were not observed.

Comparison of univariate versus multivariate cluster analysis

We compared our multivariate cluster analysis using 18 variables with a univariate approach by using each of the 18 variables by itself. To evaluate the difference in the ability of each of these methods to predict future exacerbations, as measured by the time to first use of prednisone, we performed survival analysis for each of the models. We found that the single variable with the best predictive accuracy for future exacerbations was that of history of prior hospitalization for asthma exacerbations (Fig E2). This variable was the only single variable to outperform the multivariate phenotypic clusters in terms of its ability to predict future exacerbations (as measured by *r*²). However, as a dichotomous variable, this factor provided only gross partitioning of the 3 lowest-risk and 2 highest-risk groups.

Reproducibility of cluster assignments by using different clustering algorithms

Using hierarchical clustering, we were able to generate clusters quite similar in composition to our original clusters in terms of AOE grouping (Fig E3, A). To assess whether the new cluster assignments also demonstrated longitudinal consistency similar to the original clusters, we repeated our survival analysis of time to asthma exacerbation using the 4 years of follow-up data generated as part of the CAMP clinical trial. For the survival analysis, we found that the clusters generated by using hierarchical clustering demonstrated a similar natural history to our original clusters (Fig E3, B).

Cluster grouping correlates with prospective long-term asthma control and response to specific inhaled anti-inflammatory controller medications

We performed a Cox regression analysis stratified by treatment group. The primary outcome of interest was exacerbation rate, defined as the initiation of therapy with oral prednisone. We found that for all clusters, with the exception of cluster 2, therapy with budesonide resulted in significantly fewer exacerbations when compared with placebo (Table E5). We also found that for clusters 3, 4, and 5, therapy with nedocromil resulted in fewer

exacerbations when compared with placebo. We also performed a formal test of interaction between cluster and treatment group that demonstrated a nominally significant interaction ($P = .05$) between the phenotypic cluster and the study drug for cluster 4 and both nedocromil (with placebo as reference) and budesonide (with nedocromil as reference), indicating that this cluster showed some modest response to nedocromil that was not detected in the original CAMP study when subjects were not evaluated by cluster (see [Table E6](#)).

REFERENCES

- E1. Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci U S A* 2002;99:6567-72.
- E2. de Souto MCP, De Araugo DSA, Costa IG, Soares R, Ludermir TB, Schliep A. Comparative study on normalization procedures for cluster analysis of gene expression datasets. In: *Neural Networks, 2008. IJCNN, 2008: Proceedings of the International Joint Conference on Neural Networks; 2008 June 1-8; Hong Kong. IEEE; 2008, 2792-98.*
- E3. Karatzoglou A, Hornik K, Smola A, Zeileis A. Kernlab—an S4 package for kernel methods in R. *J Stat Software* 2004;11:1-20.
- E4. Tibshirani R, Walther G, Hastie T. Estimating the number of clusters in a dataset via the gap statistic. *J R Stat Soc B* 2001;32:411-23.
- E5. 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.
- E6. 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, e1-13.

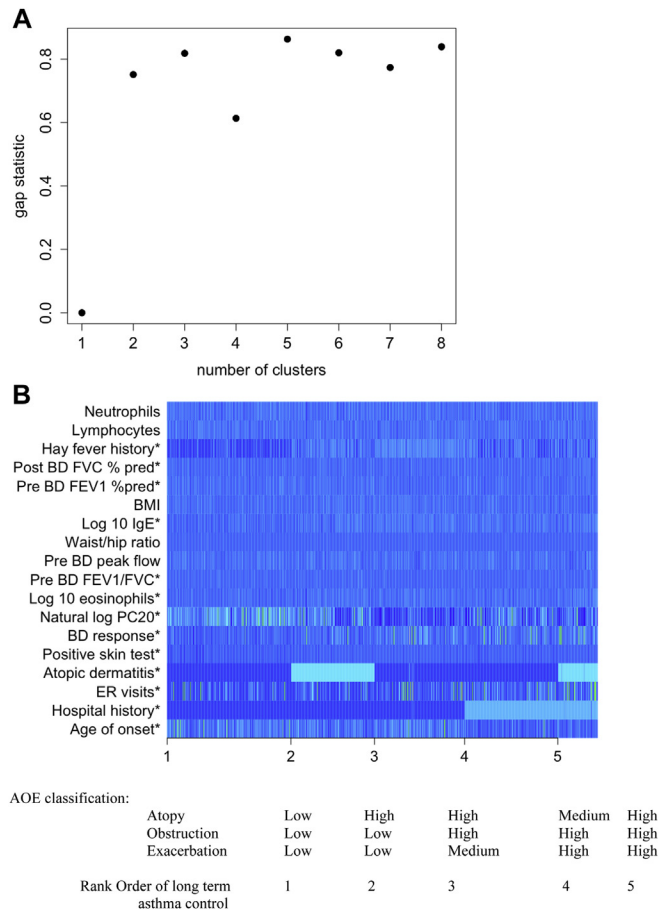


FIG E1. A, The gap statistic as a function of the number of clusters. A higher gap statistic indicates greater between-cluster separation. **B**, Heat map depicting the differences among normalized clinical variables used for clustering and the different phenotypic clusters. The cluster assignments are grouped along the *horizontal axis*, and the variables used to determine the cluster assignments appear along the *vertical axis*. The lighter shades of blue denote relatively higher magnitudes for each variable, and the darker shades denote relatively lower magnitudes. * $P < .0001$ for difference in distribution across clusters. Variables demonstrating more distinct between-cluster differences in magnitude, such as hospital history, AD, PC₂₀, and hay fever history, were the primary drivers of the cluster assignments. *BD*, Bronchodilator.

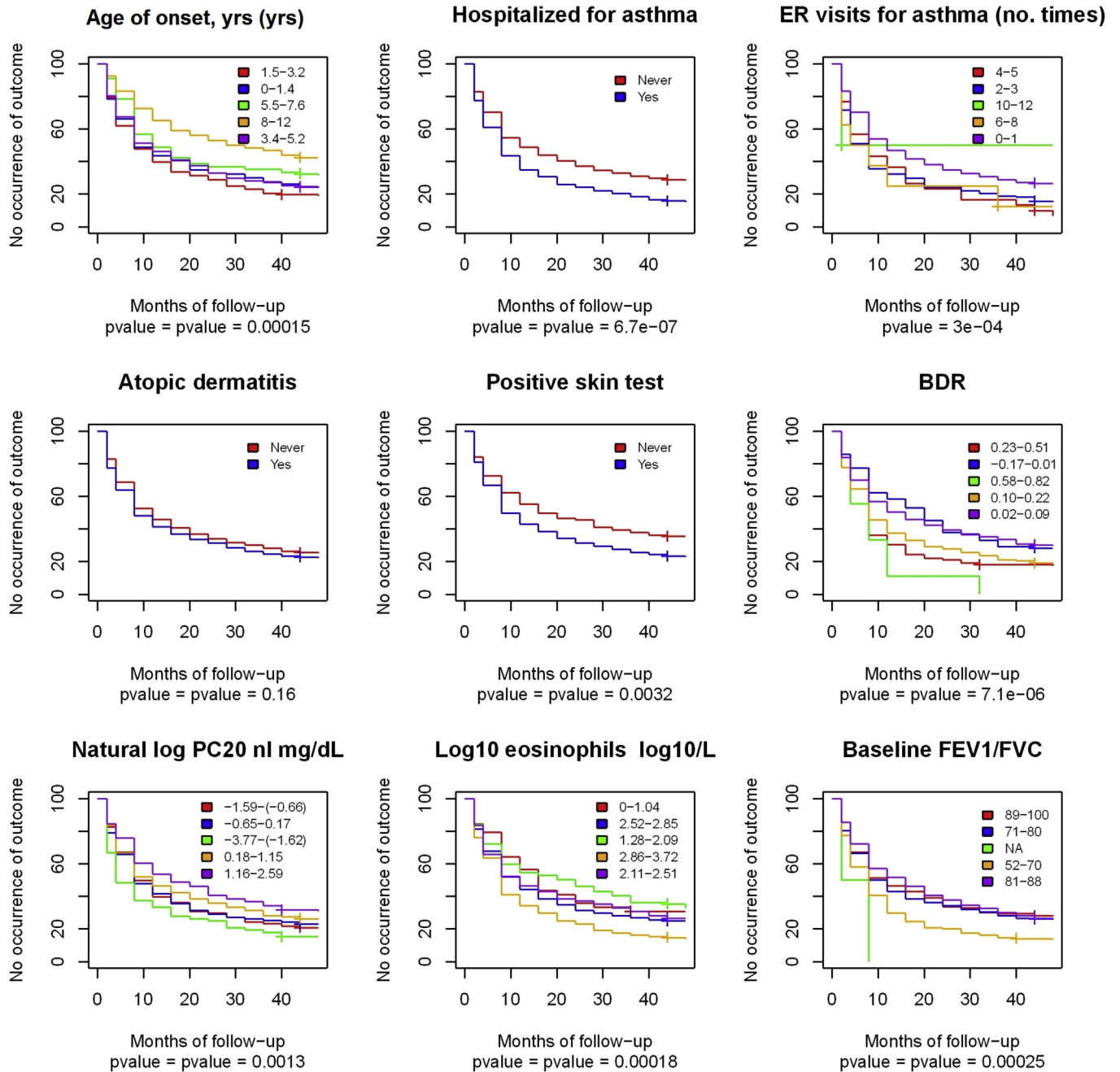


FIG E2. Kaplan-Meier plots by cluster of the cumulative probability of a first course of prednisone during the 4-year follow-up period of the CAMP trial. Clusters were determined based on a single clinical variable indicated at the top of each figure.

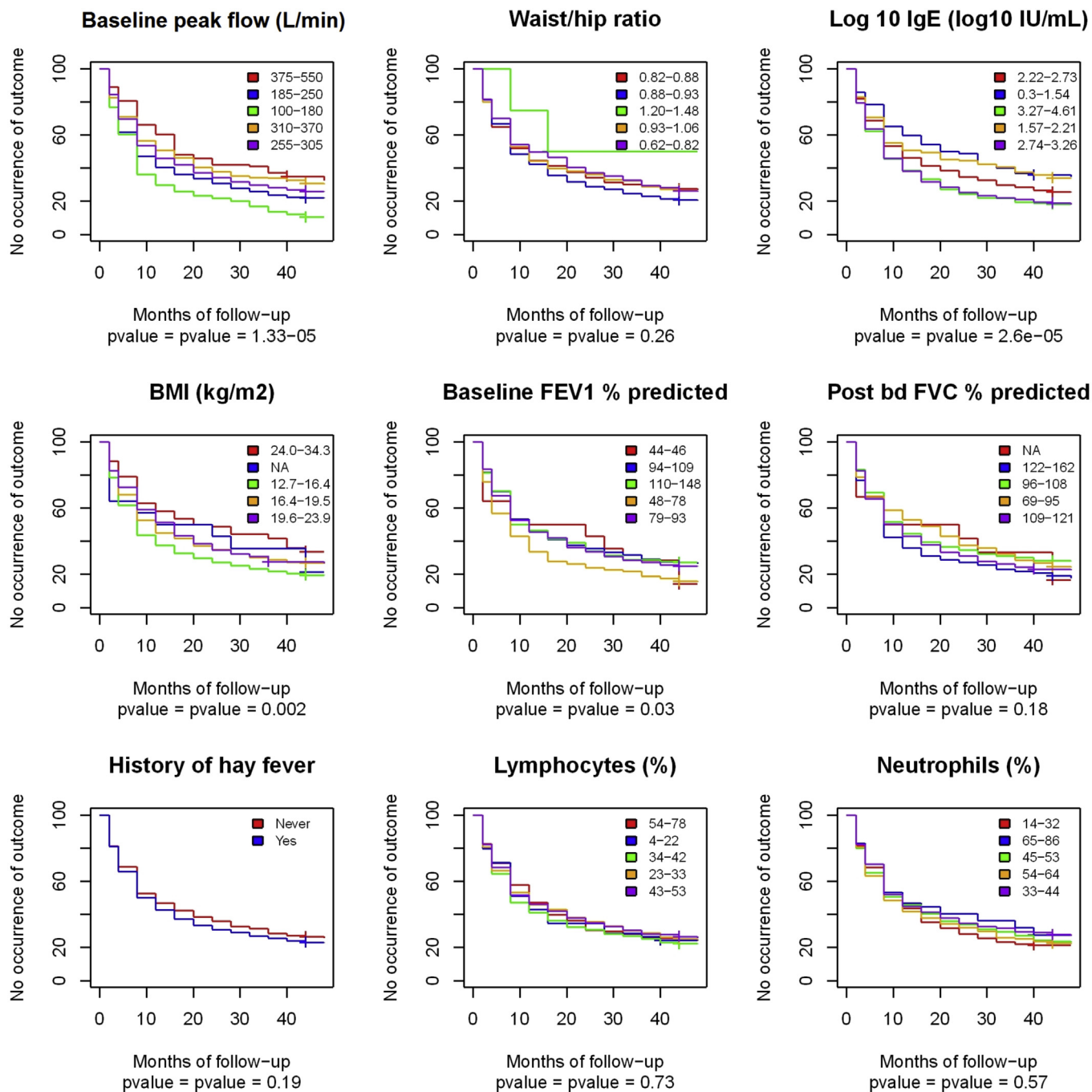


FIG E2. (Continued)

A

	New Cluster	1	2	3	4	5
Old Cluster	1	188	2	7	29	2
	2	0	164	0	37	1
	3	57	0	130	27	4
	4	3	1	1	3	217
	5	0	0	0	20	76

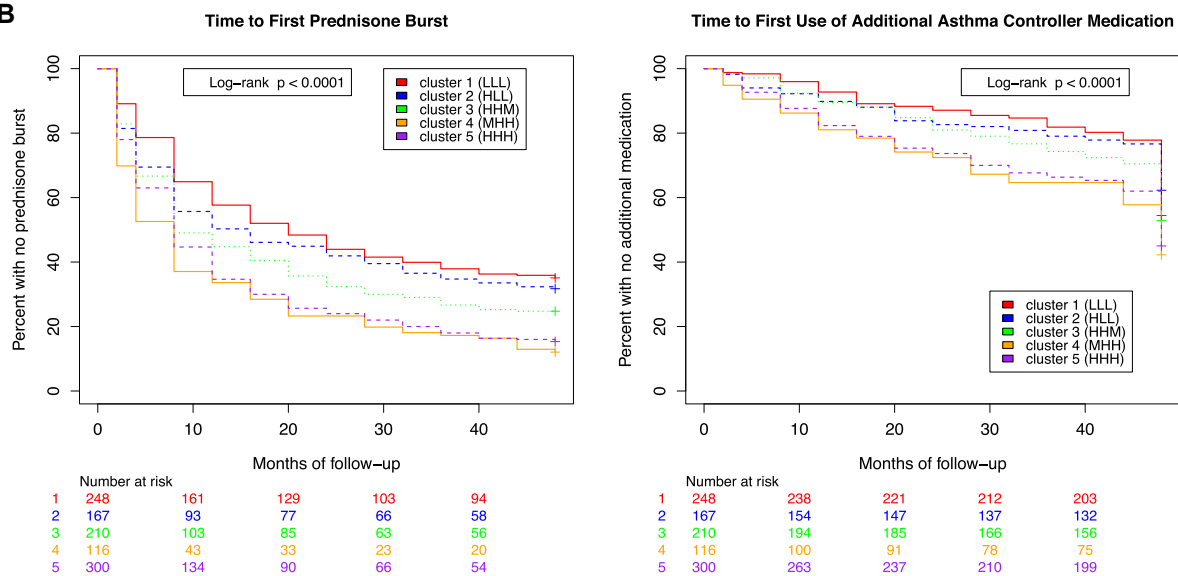
B

FIG E3. Comparison of phenotypic clusters generated by means of hierarchical clustering (new clusters) versus spectral clustering (old clusters). **A**, Comparison of study subject cluster assignments. **B**, Survival analysis for time to first use of prednisone (*left*) and time to use of additional asthma controller therapy (*right*) for hierarchical clusters.

TABLE E1. Baseline clinical variables considered for cluster analysis

Medical history	Reason for exclusion
<ul style="list-style-type: none"> • Hay fever • Positive allergy skin test result • AD • ED visits for asthma • Ever hospitalized for asthma • Age of asthma onset 	All variables were included in the model, except for caregiver's assessment, which was considered too subjective.
CAMP study coordinator's assessment of asthma	
Asthma worsened by house dust or animals or tobacco smoke or chemicals or emotional factors or exercise or certain foods or respiratory tract infections or dampness or changes in the weather or cold air or aspirin	>10% missing values (cold air and aspirin). The remainder were considered too subjective.
Clinical presentation	These variables were considered to be too subjective for use in the clustering model.
Provoked by exercise	
Provoked by allergy	
Age at first symptoms	
Symptom burden	
Age when child began wheezing with shortness of breath	>10% missing values
Prior awakening from sleep because of cough or wheeze	>10% missing values
Awakening from sleep in past 6 mo	>10% missing values
Awakening from sleep in past month	>10% missing values
Awakening from sleep in past week	>10% missing values
Cough or wheeze during the day unrelated to exercise	>10% missing values
Cough or wheeze during the day caused by exercise	>10% missing values
Cough or phlegm with or without an upper respiratory tract infection	overly subjective
Wheezing present on most days	overly subjective
Wheezing present with or without an upper respiratory tract infection	overly subjective
Wheezing present with shortness of breath	>10% missing values
Two or more episodes of wheezing with shortness of breath	>10% missing values
Received a prescription medication for wheezing with shortness of breath	>10% missing values
Normal breathing between attacks of wheezing with shortness of breath	
Anthropomorphic measurements	All variables were included in the model.
<ul style="list-style-type: none"> • BMI • Waist/hip ratio 	
Pulmonary function	All variables were included in the model.
<ul style="list-style-type: none"> • Post-BD FVC as a percentage of the predicted value (post-BD FVC percent predicted) • Pre-PD FEV₁ as a percentage of predicted value (pre-BD FEV₁ percent predicted) • Pre-BD peak expiratory flow rate • Pre-BD FEV₁/FVC ratio • Methacholine PC₂₀ (natural log) • Post-BD FEV₁ – pre-BD FEV₁/pre-BD FEV₁ (BDR) 	
Peripheral blood measures	All variables were included in the model.
<ul style="list-style-type: none"> • Absolute serum neutrophil count • Absolute serum lymphocyte count • Total serum IgE level (log₁₀) • Absolute serum eosinophils (log₁₀) 	

Baseline clinical variables were considered for cluster analysis. From an initial list of 48 variables shown in the table, we selected 18 clinical variables (denoted by *asterisks*) as inputs to the spectral clustering algorithm.

BDR, Bronchodilator response.

TABLE E2. Baseline features of 1041 CAMP asthmatic subjects

Variable	Count or mean
Sex	
Male (%)	621 (59.7)
Female (%)	420 (40.3)
Age (y)	8.94 ± 2.12
Self-reported race	
White (%)	711 (68.3)
Black (%)	138 (13.3)
Hispanic (%)	98 (9.41)
Other (%)	94 (9.03)
Family history of asthma (%)	
Yes	574 (55.1)
No	444 (42.7)
Missing	23 (2.21)
Family history of atopy (%)	724 (69.5)
History of tobacco smoke exposure (%)	439 (42.2)
Household income	
<\$30,000	242 (23.2)
≥\$30,000	758 (72.8)
Missing	41 (3.94)
Age of asthma onset (y)	3.07 ± 2.44
Hospitalized for asthma (%)	320 (30.7)
ED visits for asthma, no./100 person-years	648 ± 62.2
History of AD (%)	298 (28.6)
History of hay fever (%)	557 (53.5)
History of positive skin test result (%)	914 (87.8)
Prebronchodilator FEV ₁ (L [range])	1.65 (0.42-3.31)
Prebronchodilator FEV ₁ /FVC ratio (range)	80 (52-100)
FEV ₁ bronchodilator response (L [range])	0.11 (-0.17 to 0.82)
Methacholine PC ₂₀ (natural log mg/dL [range])	0.10 (-3.77 to 2.59)
Total serum IgE level (IU/L [range])	484 (0-5304)
Peripheral blood eosinophil count (log ₁₀ /L [range])	2.50 (0-3.72)
Waist/hip ratio	0.88 ± 0.061
BMI (kg/m ²)	18.2 ± 3.52

TABLE E3. Number of patients needing additional or replacement asthma controller medications

AOE classification	Cluster 1 (n = 300)	Cluster 2 (n = 202)	Cluster 3 (n = 218)	Cluster 4 (n = 225)	Cluster 5 (n = 96)	P value
	LLL	HLL	HHM	MHH	HHH	
Budesonide (n = 311 [29.9%])						
2 mo	0.01 ± 0.11	0.04 ± 0.21	0.03 ± 0.18	0.00 ± 0.71	0.00 ± 0.72	.32
4 mo	0.01 ± 0.11	0.05 ± 0.21	0.03 ± 0.18	0.00 ± 0.84	0.04 ± 0.69	.43
8 mo	0.01 ± 0.12	0.06 ± 0.30	0.03 ± 0.18	0.06 ± 0.96	0.04 ± 1.25	.66
12 mo	0.03 ± 0.16	0.06 ± 0.39	0.07 ± 0.26	0.11 ± 0.88	0.13 ± 1.47	.24
16 mo	0.04 ± 0.20	0.06 ± 0.40	0.11 ± 0.31	0.13 ± 0.97	0.17 ± 0.78	.21
20 mo	0.05 ± 0.28	0.10 ± 0.43	0.13 ± 0.33	0.16 ± 1.02	0.17 ± 1.72	.43
24 mo	0.06 ± 0.29	0.10 ± 0.43	0.13 ± 0.33	0.18 ± 0.93	0.17 ± 0.78	.38
28 mo	0.10 ± 0.34	0.10 ± 0.43	0.15 ± 0.36	0.25 ± 0.95	0.26 ± 0.92	.46
32 mo	0.11 ± 0.36	0.10 ± 0.43	0.19 ± 0.44	0.26 ± 0.96	0.26 ± 0.92	.43
36 mo	0.18 ± 0.52	0.11 ± 0.45	0.21 ± 0.50	0.35 ± 2.98	0.30 ± 0.70	.52
40 mo	0.24 ± 0.64	0.11 ± 0.45	0.27 ± 0.63	0.51 ± 0.65	0.36 ± 0.71	.27
44 mo	0.25 ± 0.65	0.13 ± 0.47	0.31 ± 0.67	0.54 ± 0.71	0.38 ± 0.78	.18
48 mo	0.66 ± 0.92	0.37 ± 0.64	0.50 ± 0.71	0.82 ± 0.53	0.70 ± 0.86	.27
Nedocromil (n = 312 [30.0%])						
2 mo	0.00 ± 0.0	0.02 ± 0.13	0.00 ± 0.0	0.02 ± 0.0	0.04 ± 0.0	.33
4 mo	0.00 ± 0.0	0.09 ± 0.35	0.02 ± 0.12	0.05 ± 0.0	0.17 ± 0.20	.002
8 mo	0.06 ± 0.23	0.19 ± 0.62	0.12 ± 0.38	0.11 ± 0.24	0.23 ± 0.20	.20
12 mo	0.13 ± 0.43	0.28 ± 0.94	0.14 ± 0.39	0.17 ± 0.36	0.36 ± 0.34	.17
16 mo	0.22 ± 0.60	0.41 ± 1.42	0.18 ± 0.46	0.26 ± 0.38	0.59 ± 0.48	.15
20 mo	0.30 ± 0.78	0.48 ± 1.58	0.23 ± 0.56	0.30 ± 0.55	0.68 ± 0.48	.24
24 mo	0.35 ± 0.93	0.52 ± 1.61	0.45 ± 0.94	0.40 ± 0.56	0.82 ± 0.48	.27
28 mo	0.38 ± 0.99	0.64 ± 1.80	0.62 ± 1.12	0.37 ± 0.67	0.86 ± 0.75	.29
32 mo	0.42 ± 1.11	0.69 ± 1.85	0.79 ± 1.36	0.49 ± 0.75	1.00 ± 0.75	.19
36 mo	0.48 ± 1.72	0.76 ± 1.95	0.86 ± 1.46	0.60 ± 0.88	1.15 ± 0.82	.23
40 mo	0.59 ± 1.44	0.84 ± 2.10	0.95 ± 1.59	0.69 ± 1.15	1.15 ± 1.00	.29
44 mo	0.65 ± 1.50	1.00 ± 2.45	1.05 ± 1.87	0.75 ± 1.13	1.15 ± 1.02	.48
48 mo	0.94 ± 1.66	1.27 ± 2.59	1.20 ± 1.79	1.00 ± 1.20	1.26 ± 1.22	.92
Placebo (n = 418 [40.1%])						
2 mo	0.03 ± 0.16	0.01 ± 0.12	0.05 ± 0.26	0.03 ± 0.15	0.08 ± 0.35	.82
4 mo	0.05 ± 0.27	0.10 ± 0.34	0.09 ± 0.39	0.10 ± 0.34	0.18 ± 0.46	.26
8 mo	0.10 ± 0.36	0.20 ± 0.65	0.13 ± 0.43	0.18 ± 0.44	0.25 ± 0.50	.21
12 mo	0.16 ± 0.47	0.28 ± 0.82	0.17 ± 0.49	0.26 ± 0.58	0.43 ± 0.85	.12
16 mo	0.18 ± 0.50	0.31 ± 0.85	0.20 ± 0.55	0.41 ± 0.79	0.54 ± 0.98	.03
20 mo	0.25 ± 0.67	0.39 ± 0.95	0.24 ± 0.75	0.53 ± 0.87	0.83 ± 1.27	.001
24 mo	0.38 ± 0.97	0.45 ± 0.98	0.39 ± 0.96	0.58 ± 0.97	0.91 ± 1.34	.01
28 mo	0.45 ± 1.06	0.55 ± 1.09	0.43 ± 1.11	0.65 ± 1.02	1.06 ± 1.39	.004
32 mo	0.49 ± 1.09	0.62 ± 1.21	0.52 ± 1.33	0.78 ± 1.26	1.26 ± 1.54	.002
36 mo	0.56 ± 1.24	0.75 ± 1.49	0.56 ± 1.34	0.85 ± 1.44	1.37 ± 1.66	.002
40 mo	0.64 ± 1.38	0.82 ± 1.58	0.65 ± 1.56	0.95 ± 1.61	1.43 ± 1.69	.005
44 mo	0.74 ± 1.61	0.95 ± 1.74	0.71 ± 1.79	1.04 ± 1.77	1.74 ± 1.80	.003
48 mo	1.04 ± 1.60	1.02 ± 1.67	1.03 ± 2.04	1.44 ± 1.93	2.06 ± 1.84	<.001
P values						
Budesonide vs nedocromil	.32	.78	.003	.24	.28	
Budesonide vs placebo	.81	.45	.75	.13	.97	
Nedocromil vs placebo	.56	.71	.07	.65	.49	

The cumulative number of asthma exacerbations (ie, symptom worsening requiring additional or replacement asthma controller medications) at each study time point stratified by treatment group is shown. Shown are mean ± SD cumulative number of exacerbations from the onset of the study period and P values for between-cluster differences in outcome (far right) and pairwise comparisons of within-cluster differences in outcomes (bottom level).

TABLE E4. Distribution of nonclassifying features across asthma clusters

AOE classification	Cluster 1 (n = 300)	Cluster 2 (n = 202)	Cluster 3 (n = 218)	Cluster 4 (n = 225)	Cluster 5 (n = 96)	P value
	LLL	HLL	HHM	MHH	HHH	
Sex						.45
Male	173 (57.7%)	115 (56.9%)	130 (59.6%)	146 (64.9%)	57 (59.4%)	
Female	127 (42.3%)	87 (43.1%)	88 (40.4%)	79 (35.1%)	39 (40.6%)	
Age at trial enrollment (y), mean ± SD	8.79 ± 2.05	8.83 ± 2.12	9.38 ± 2.13	9.03 ± 2.08	8.46 ± 2.25	.003
Self-reported race						.19
White	217 (72.3%)	137 (67.8%)	140 (64.2%)	150 (66.7%)	67 (69.8%)	
Black	36 (12.0%)	27 (13.4%)	26 (11.9%)	35 (15.6%)	14 (14.6%)	
Hispanic	27 (9.0%)	13 (6.4%)	31 (14.2%)	21 (9.3%)	6 (6.3%)	
Other	20 (6.7%)	25 (12.4%)	21 (9.6%)	19 (8.4%)	9 (9.4%)	
Annual household income						.31
<\$30,000	70 (23.3%)	40 (19.8%)	45 (20.6%)	66 (29.3%)	21 (21.9%)	
Highest household education						.13
Less than high school	1 (0.33%)	1 (0.50%)	2 (0.92%)	1 (0.44%)	0 (0.00%)	
High school	5 (1.7%)	5 (2.5%)	4 (1.8%)	5 (2.2%)	4 (4.2%)	
Higher education	125 (41.7%)	78 (38.6%)	87 (39.9%)	101 (44.9%)	48 (50.0%)	
Family history						
Asthma (any)	154 (51.3%)	114 (56.4%)	137 (62.8%)	110 (48.9%)	59 (61.5%)	.07
Asthma (maternal)	64 (21.3%)	46 (22.8%)	70 (32.1%)	50 (22.2%)	32 (33.3%)	.02
Asthma (paternal)	40 (13.3%)	49 (24.3%)	58 (26.6%)	46 (20.4%)	15 (15.6%)	.0009
Atopy (any)	189 (63.0%)	158 (78.2%)	162 (74.3%)	139 (61.8%)	76 (79.2%)	.0004
Atopy (maternal)	114 (38.0%)	110 (54.5%)	112 (51.4%)	98 (43.6%)	53 (55.2%)	.0009
Atopy (paternal)	87 (29.0%)	90 (44.6%)	84 (38.5%)	73 (32.4%)	37 (38.5%)	.05
Environmental exposures						
Tobacco smoke	119 (39.7%)	61 (30.2%)	82 (37.6%)	105 (46.7%)	46 (47.9%)	.01
Dust mite	60 (20.0%)	36 (17.8%)	61 (28.0%)	44 (19.6%)	16 (16.7%)	.30
Cockroach	1 (0.33%)	3 (1.49%)	1 (0.46%)	0 (0.00%)	0 (0.00%)	.04
Randomized treatment arm in CAMP clinical trial						.91
Budesonide	86 (28.7%)	68 (33.7%)	60 (27.5%)	68 (30.2%)	29 (30.2%)	
Nedocromil	94 (31.3%)	60 (29.7%)	66 (30.3%)	67 (29.8%)	25 (26.0%)	
Placebo	120 (40.0%)	74 (36.6%)	92 (42.2%)	90 (40.0%)	42 (43.8%)	
Long-term asthma control rank*	1	2	3	4	5	

*Long-term asthma control rank was subsequently determined in prospective survival analysis of time to first course of oral prednisone.

TABLE E5. Summary of *P* values for Cox proportional hazards modeling of the risk of an asthma exacerbation

Initiation of oral prednisone	Treatment with budesonide (placebo as reference)	Treatment with nedocromil (placebo as reference)	Treatment with budesonide (nedocromil as reference)
Cluster 2	.12	.12	.17
Cluster 3	.03	.03	.06
Cluster 4	<.001	<.001	.006
Cluster 5	<.001	<.001	<.001

Cox proportional hazards models for decrease in risk of future asthma exacerbations by using cluster assignment and different treatment group comparisons as predictor variables under an additive model. Cluster 1 was used as the reference group for cluster assignment. Shown are the *P* values for the degree of risk contributed by each variable to the model. For example, for cluster 4, the risk of initiation of oral prednisone is significantly decreased ($P < .001$) in the budesonide group (compared with placebo), significantly decreased ($P < .001$) in the nedocromil group (compared with placebo), and significantly decreased in the budesonide group ($P = .006$) compared with nedocromil.

TABLE E6. Summary of *P* values for Cox proportional hazards modeling of drug-by-cluster interaction

Initiation of oral prednisone	Treatment with budesonide (placebo as reference)	Treatment with nedocromil (placebo as reference)	Treatment with budesonide (nedocromil as reference)
Drug*Cluster 2	.94	.93	.99
Drug*Cluster 3	.93	.87	.97
Drug*Cluster 4	.98	.05	.05
Drug*Cluster 5	.92	.60	.50

Cox proportional hazards models for decrease in risk of future asthma exacerbations using cluster assignment and treatment group as predictor variables under an interaction model are shown. Cluster 1 was used as the reference group for cluster assignment. Shown are *P* values for the interaction terms. For example, for cluster 4, when considering the risk of initiation of oral prednisone, there is a nominally significant ($P = .05$) interaction between cluster membership and the study drug when budesonide is considered with the nedocromil group as a reference and when nedocromil is considered with the placebo group as a reference, suggesting that there is some differential response to medical therapy that exists within this phenotypic cluster compared with the other clusters.