

## Molecular phenotyping of severe asthma using pattern recognition of bronchoalveolar lavage-derived cytokines

Allan R. Brasier, MD,<sup>a,b</sup> Sundar Victor, MS,<sup>a,b</sup> Gary Boetticher, PhD,<sup>c</sup> Hyunsu Ju, PhD,<sup>d</sup> Chang Lee, MS,<sup>a</sup> Eugene R. Bleecker, MD,<sup>e</sup> Mario Castro, MD,<sup>f</sup> William W. Busse, MD,<sup>g</sup> and William J. Calhoun, MD<sup>a,b</sup> Galveston and Clear Lake, Tex, Winston-Salem, NC, St Louis, Mo, and Madison, Wis

**Background:** Asthma is a heterogeneous clinical disorder. Methods for objective identification of disease subtypes will focus on clinical interventions and help identify causative pathways. Few studies have explored phenotypes at a molecular level.

**Objective:** We sought to discriminate asthma phenotypes on the basis of cytokine profiles in bronchoalveolar lavage (BAL) samples from patients with mild-moderate and severe asthma. **Methods:** Twenty-five cytokines were measured in BAL samples of 84 patients (41 severe, 43 mild-moderate) using bead-based multiplex immunoassays. The normalized data were subjected to statistical and informatics analysis.

**Results:** Four groups of asthmatic profiles could be identified on the basis of unsupervised analysis (hierarchical clustering) that were independent of treatment. One group, enriched in patients with severe asthma, showed differences in BAL cellular content, reductions in baseline pulmonary function, and enhanced response to methacholine provocation. Ten cytokines were identified that accurately predicted this group. Classification methods for predicting methacholine sensitivity were developed. The best model analysis predicted hyperresponders with 88% accuracy in 10 trials by using a 10-fold cross-validation. The cytokines that contributed to this model were IL-2, IL-4, and IL-5. On the basis of this classifier, 3 distinct hyperresponder

classes were identified that varied in BAL eosinophil count and PC<sub>20</sub> methacholine.

**Conclusion:** Cytokine expression patterns in BAL can be used to identify distinct types of asthma and identify distinct subsets of methacholine hyperresponders. Further biomarker discovery in BAL may be informative.

(*J Allergy Clin Immunol* 2008;121:30-7.)

**Key words:** Asthma, inflammation, cytokines, phenotypes, ELISA, hierarchical clustering, bioinformatics, class discovery

Asthma is a chronic inflammatory disease of the airways characterized by recurrent episodes of symptomatic airflow obstruction and various degrees of airways hyperreactivity to nonspecific stimuli.<sup>1</sup> The recognition that this disease has a chronic inflammatory component has directed therapy toward early use of inhaled glucocorticoid therapy, typically producing significant reductions in inflammatory markers and improvement in pulmonary function.<sup>1</sup> However, there is a subset of patients (~5% to 7%) with severe, or refractory asthma<sup>2</sup> that does not respond to glucocorticoids. These patients account for 40% to 50% of the health costs of asthma and incur significant morbidity and decrements in quality of life.<sup>3,4</sup>

Severe asthma is a heterogeneous disorder with distinct ages of onset,<sup>2</sup> duration of disease, degree of airflow impairment, presence of modifying factors (gastroesophageal reflux disease, sinusitis), and type of underlying inflammation.<sup>2,5,6</sup> In this regard, phenotypic analysis of patients with severe asthma prospectively enrolled in the Severe Asthma Research Program (SARP) has shown that patients with severe asthma tend to be older and have a greater frequency of respiratory infection (sinusitis and/or pneumonia), suggesting that as a group, they have alterations in innate immune defenses.<sup>7</sup> In addition, at least some patients with severe asthma have been characterized as having either neutrophil-predominant inflammation or increased tissue eosinophils by endobronchial biopsy.<sup>8,9</sup> Eosinophil-positive patients, especially with early-onset disease and associated airway remodeling, have been shown to have an increased incidence of near-fatal events.<sup>5,8</sup> However, others have found no clinical differences between the eosinophilic and noneosinophilic phenotypes.<sup>10</sup> Together, these observations suggest that severe asthma is a pathologically heterogeneous disorder that still lacks an objective method for distinguishing clinically significant subtypes.<sup>6</sup>

The findings that patients with severe asthma have distinct inflammatory processes suggest that they may also express distinct airway cytokine profiles compared with patients with responsive asthma. Here we investigate this hypothesis by

From <sup>a</sup>the Department of Internal Medicine, University of Texas Medical Branch, Galveston; <sup>b</sup>the Sealy Center for Molecular Medicine, Galveston; <sup>c</sup>the Department of Computer Science, University of Houston-Clear Lake; <sup>d</sup>the Sealy Center on Aging, Galveston; <sup>e</sup>the Department of Medicine, Wake Forest University School of Medicine, Winston-Salem; <sup>f</sup>Washington University in St Louis; and <sup>g</sup>the Department of Medicine, University of Wisconsin, Madison.

Supported by National Institutes of Health grant AI062885 (A.R.B.), a National Institute of Environmental Health Sciences Pilot Project (A.R.B.), National Heart, Lung, and Blood Institute contract BAA-HL-02-04 (A. Kurosky), HL69130 US Severe Asthma Research Program (W.J.C.), and the Integrated Health Science Facility Core P30 ES06676 (to J. Halpert, UTMB).

Disclosure of potential conflict of interest: W. W. Busse has consulting arrangements with Genentech/Novartis, Isis, GlaxoSmithKline, Atlanta, Wyeth, Pfizer, Dynavax, Centocor, Johnson and Johnson, Hoffman LaRoche, Alza CV Therapeutics, Takeda, Abbott, and Millenium; has received grant support from Novartis, Dynavax, Wyeth, Centocor, GlaxoSmithKline, Medicinova, Sanofi-Aventis, Dey, Pfizer, Astellas Pharma, Inflazyme, and Biowa; and is on the speakers' bureau for Novartis, Merck, and Astra-Zeneca. W. J. Calhoun has received grant support from Sepracor. The rest of the authors have declared that they have no conflict of interest.

Received for publication May 15, 2007; revised October 10, 2007; accepted for publication October 16, 2007.

Reprint requests: Allan R. Brasier, MD, MRB 8.128, University of Texas Medical Branch, 301 University Blvd, Galveston, TX 77555-1060. E-mail: arbrasie@utmb.edu.

0091-6749/\$34.00

© 2008 American Academy of Allergy, Asthma & Immunology

doi:10.1016/j.jaci.2007.10.015

#### Abbreviations used

ATS: American Thoracic Society  
BAL: Bronchoalveolar lavage  
FVC: Forced vital capacity  
IL-1Ra: IL-1 receptor antagonist  
LR: Low responder  
MCP: Monocyte chemoattractant protein  
MIP: Macrophage inflammatory protein  
MOP: Manual of procedures  
SARP: Severe Asthma Research Program

examination of airway cytokine expression patterns in bronchoalveolar lavage (BAL) from a matched group of patients with nonsevere and severe asthma by using bead-based multiplex cytokine arrays (Luminex xMAP). The data were analyzed by using both unsupervised and supervised classification methods. Accurate definition of asthmatic phenotypes based on molecular profiles may facilitate clinical investigation on the pathogenesis and treatment of asthma.

## METHODS

### Patients

In the SARP, enrollees are categorized as (1) healthy volunteers, (2) patients with nonsevere asthma, and (3) patients with severe asthma on the basis of a standardized manual of procedures (MOP) based on an National Heart, Lung, and Blood Institute workshop.<sup>11,12</sup> All enrollees have history, physical examination, spirometry, bronchodilator reversibility, allergy skin testing, and methacholine challenge testing. Healthy volunteers have normal lung function and negative methacholine challenge, no history of asthma, and no need for any routine medications. Patients with nonsevere asthma have lung function that can be normalized by using standard doses of inhaled glucocorticoids, with or without long-acting  $\beta$ -agonists or leukotriene modifiers. Patients with severe asthma are defined according to American Thoracic Society (ATS) consensus for refractory asthma.<sup>11</sup> These patients are characterized by abnormal lung function in the face of aggressive standard inhaled glucocorticoid therapy and at least 1 additional control agent; they must in addition have at least 2 positive skin tests. Subjects performed spirometry before and after as many as 8 puffs (90  $\mu\text{g}/\text{puff}$ ) of albuterol. The baseline FEV<sub>1</sub> testing required a withhold of 4 to 6 hours of short-acting bronchodilators and a hold of 10 to 12 hours for long-acting bronchodilators. Hankinson predicted values (with race correction) were used to obtain percent predicted values.<sup>13</sup> Methacholine challenge was performed according to the SARP MOP. No testing was performed on subjects with <70% FEV<sub>1</sub>. All studies were approved by the local institutional review boards, and all subjects gave informed consent.

### BAL

Bronchoscopy and BAL were conducted according to the SARP MOP. Briefly, after topical anesthesia, bronchoscopy was performed. BAL was obtained using 2 aliquots of 50 mL each of 0.9% NaCl. Cells were separated by low speed centrifugation (400g, 20 minutes), and supernatants were frozen for subsequent analysis.

### BAL analysis

Bronchoalveolar lavage and deidentified clinical information were obtained from the SARP for 84 randomly selected patients matched for age and sex (Table I); by the SARP criteria, 43 had nonsevere asthma and 41 had severe asthma.

For each sample, 50  $\mu\text{L}$  BAL was clarified by high-speed centrifugation (10,000g for 3 minutes at 4°C). The supernatants were then analyzed for 25 human cytokines (BioSource 25-Plex panel, Invitrogen, Carlsbad, Calif).

Duplicate samples and serial dilutions of the cytokine standards (50  $\mu\text{L}$ ) were incubated with antihuman cytokine-coated beads in 96-well filtration plate (Millipore, Bedford, Mass) for 30 minutes. In this assay, panels of colored microspheres conjugated with capture antibodies are bound to the sample (each capture antibody is conjugated with a uniquely colored microsphere). This panel includes IL-1, IL-1 receptor antagonist (IL-1Ra), IL-2, IL-2R, IL-4, IL-5, IL-6, IL-7, IL-8, IL-10, IL-12p40, IL-13, IL-15, IL-17, TNF- $\alpha$ , IFN- $\beta$ , IFN- $\gamma$ , GM-CSF, macrophage inflammatory protein (MIP)-1 $\alpha$ , MIP-1 $\beta$ , IP-10, monokine induced by IFN- $\gamma$  (MIG), eotaxin, RANTES, and monocyte chemoattractant protein (MCP)-1. The plates were vacuum-washed 3 times with 100  $\mu\text{L}$  of wash buffer and incubated with 25  $\mu\text{L}$  biotinylated antibody cocktail for 30 minutes. The immune reaction is then developed by adding 50  $\mu\text{L}$  streptavidin-phycoerythrin for 10 minutes followed by 3 washes. The samples are then resuspended in 100  $\mu\text{L}$  assay buffer, and 100 beads of each cytokine are acquired and analyzed. For each cytokine, a standard curve is generated by using recombinant proteins to estimate protein concentration in the unknown sample. In our hands, these assays have a sensitivity comparable to ELISA measurements, with a detection limit of 10 to 30 ng/L (depending on cytokine), low interassay variation (<10%), and a dynamic range of as many as 3 orders of magnitude.

Validation of IL-2 measurements were performed in 10 hyperresponder and 26 low responder (LR) BALs by using QuantiGlo human IL-2 chemiluminescent ELISA (R&D Systems, Minneapolis, Minn) according to the manufacturer's instructions (see this article's Fig E4 in the Online Repository at [www.jacionline.org](http://www.jacionline.org)).

### Data analysis

Chemokine concentrations were determined on the basis of a simultaneously measured standard curve using a logistic curve fitting algorithm (BioPlex Manager 3.0 Software, BioRad, Hercules, Calif). The standard curve for each cytokine in this panel has a range between 2 and 5000 pg/mL. Sample data were used when duplicate measurements showed less than 10% difference. BAL fluid concentrations were analyzed as raw concentrations without normalization to total protein, albumin, or other markers. This strategy is consistent with the recommendation of the Bronchoalveolar Lavage Cooperative Study Group.<sup>14</sup>

### Hierarchical clustering

The data were reduced from 25 to 18 features by removing those cytokines for which more than 50% of the data were undetectable. These cytokines were IL-1 $\beta$ , IL-7, IL-10, IL-12, IL-13, IFN- $\alpha$ , and GM-CSF. Unsupervised agglomerative ("bottoms-up") hierarchical clustering was performed on the 50 percentile normalized data by using the unweighted paired group mean with correlation as the similarity measure (Spotfire Decision Site 9.0, Somerville, Mass). Cytokine values below the level of assay detection were replaced by values representing 1/10 of the lowest value measured on the standard curve.

### Classifiers and attribute reduction

A C4.5 classifier<sup>15</sup> was applied on the z-score transformed cytokine values using entropy for splitting and 10-fold cross validation (WEKA 3.5.6 software, University of Waikato, Hamilton, New Zealand).

### Statistical analysis

ANOVA with multiple comparisons and Kruskal-Wallis tests were performed by using SAS, version 9.1 (SAS, Inc, Cary, NC) and SPSS, Release 11.0.1 (SPSS, Inc, Chicago, Ill). Shrunken centroid classification and feature reduction were performed by using prediction analysis in microarray.<sup>16</sup>

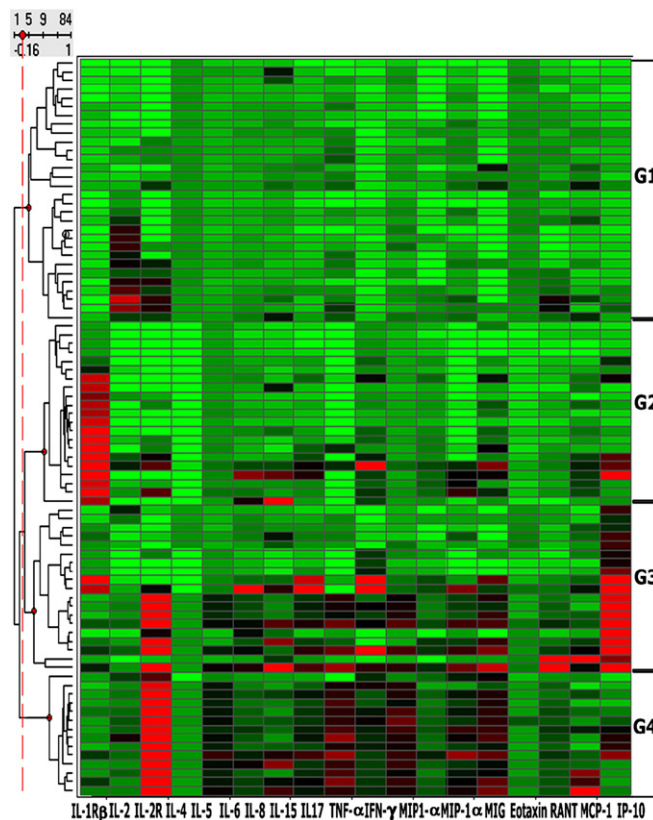
## RESULTS

The patients studied were 41 patients with severe asthma and 43 age-matched and sex-matched patients with nonsevere asthma enrolled by the SARP program whose characteristics are shown in

**TABLE I.** Patient characteristics

Phenotype	Characteristic	No. of men = 30 (36%)	No. of women = 54 (64%)	All subjects
Not severe (n = 43)		15 (34.9%)	28 (65.1%)	43 (100%)
	Age of onset (y)	11.6 ± 10.4	18.6 ± 11.5	16 ± 11.5
	Baseline FEV <sub>1</sub> (L)	3.7 ± 1	2.8 ± 0.6	3.14 ± 1§
	Baseline FEV <sub>1</sub> (% predicted)	85.4 ± 18.1	90.2 ± 16.2	88.5 ± 16.8§
	Baseline FVC (L)	5.3 ± 0.9	3.6 ± 0.7	4.2 ± 1.15§
	Baseline FVC (% predicted)	97.7 ± 12.3	99.2 ± 15.7	98.6 ± 4.5§
	Maximum FEV <sub>1</sub> reversal	10.1 ± 5.8	12.3 ± 11.4	11.5 ± 9.8*
	IgE titer (log)	2.22 ± 0.3	2.31 ± 0.49	2.28 ± 0.43
	Oral glucocorticoid	7%	4%	5%§§
	Reactive to <i>Alternaria</i>	43%	16%	26%
Severe (n = 41)		15 (36.6%)	26 (63.4%)	41 (100%)
	Age of onset (y)	10.2 ± 12.3	16.4 ± 15.4	14.1 ± 14.6
	Baseline FEV <sub>1</sub> (L)	2.3 ± 0.91	2.2 ± 0.6	2.2 ± 0.8
	Baseline FEV <sub>1</sub> (% predicted)	64.2 ± 16.7	76.1 ± 19.8	71.7 ± 19.4
	Baseline FVC (L)	3.41 ± 1	3.1 ± 0.74	3.2 ± 0.8
	Baseline FVC (% predicted)	79.7 ± 13.3	87.2 ± 18.3	84.5 ± 16.9
	Maximum FEV <sub>1</sub> reversal	24.2 ± 20.7	14.7 ± 16	18.2 ± 18.2
	IgE titer (log)	2.48 ± 0.4	2.02 ± 0.5	2.19 ± 0.5
	Oral glucocorticoid	33%	46%	41%
	Reactive to <i>Alternaria</i>	36%	19%	26%

\* $P < .05$ ; § $P < .01$ ; §§ $P < .001$  ( $t$  test for all subjects severe vs not severe).



**FIG 1.** Hierarchical clustering of 18 cytokines. Shown is a heat map of clustering cytokine values. Each row is an individual patient. *Left*, Dendrogram showing similarity of groups. *Right*, Four major groups (G) are indicated by vertical bars (G1-G4). RANT, RANTES.

**Table I.** There were no differences in the age of onset, sex distribution, serum IgE values, or positive skin test results to *Alternaria* between the 2 groups. Patients with nonsevere asthma had nearly normal FEV<sub>1</sub> (89% predicted) and forced vital capacity (FVC;

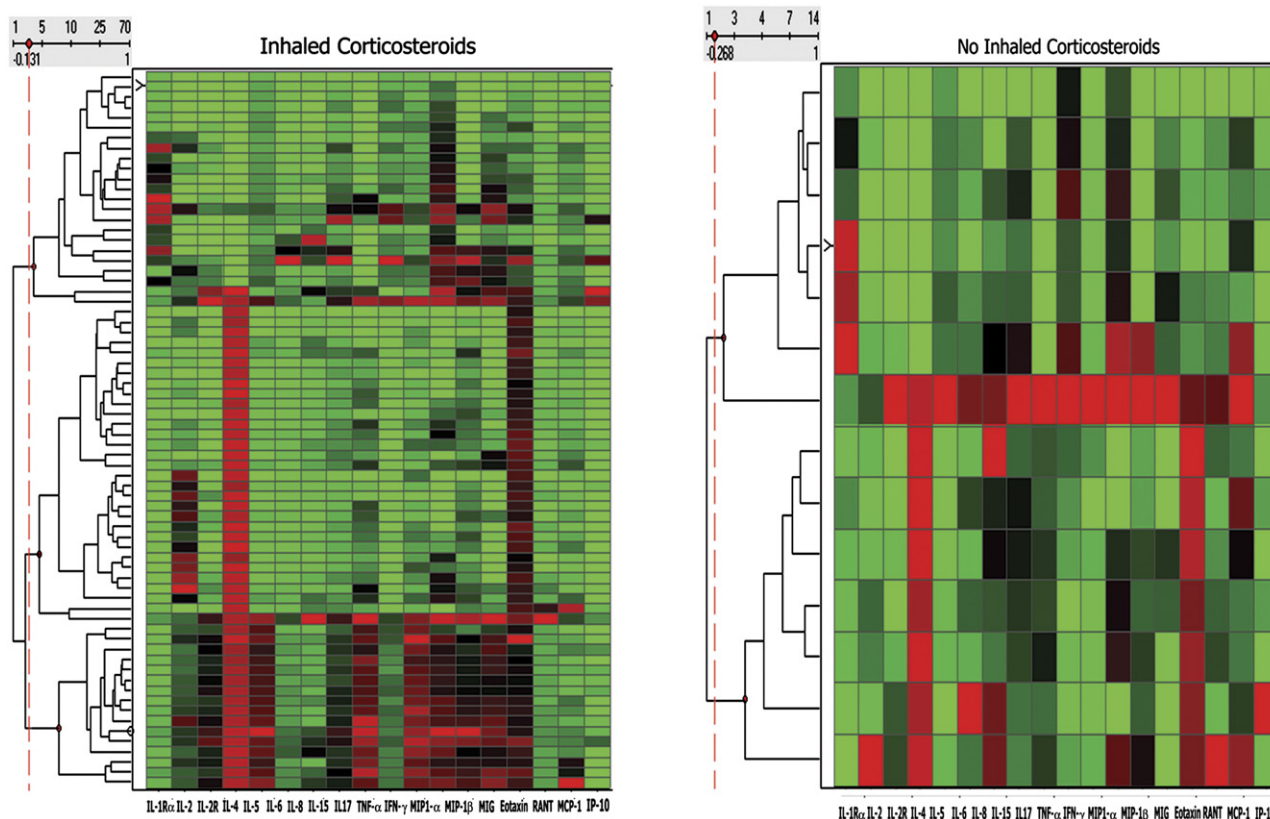
**TABLE II.** Between-group differences in groups 1, 2, 3, and 4

Function	Characteristic	P value (ANOVA)	
General	Body mass index	.02	
	BAL cellularity	Macrophages	.05
		Lymphocytes	.05
		Eosinophils	.05
Lung function (Baseline)	FEV <sub>1</sub>	<.001	
	FEV <sub>1pp</sub>	.02	
	FVC	<.001	
	FVCpp	.03	
	Maximum FEV <sub>1</sub> /MPV	<.001	
	Maximum FVC/MPV	<.001	
	Maximum FVCpp/MPVLung	.01	
Lung function (Post-treatment)	logPC <sub>20</sub> methacholine	.01	
	FEV <sub>1</sub> reversal (%)	.03	
	FEV <sub>1</sub> albuterol reversal	<.001	

pp, Percent predicted; MPV, maximum postbronchodilator value.

99% predicted). Patients with severe asthma had significant reductions in FEV<sub>1</sub> compared with patients with nonsevere asthma (72% vs 89%;  $P < .01$ ) and a greater maximal FEV<sub>1</sub> reversal after albuterol inhalation (18% vs 12%;  $P < .05$ ), and a significantly greater number were taking oral glucocorticoids (41% vs 5%;  $P < .001$ ). The differences in baseline pulmonary function are representative of the severe asthmatic population,<sup>7</sup> and the differences in treatment regimen are inherent in the classification of severe asthma by the ATS consensus criteria.<sup>12</sup>

We initially focused on molecular profiling on the basis of cytokine measurements because these molecules mediate airway inflammation by recruiting leukocyte populations, affecting T<sub>H</sub>1/T<sub>H</sub>2 balance, and promoting smooth muscle cell proliferation. We were able to detect the expression of 18 cytokines in BAL (7 cytokine measurements were not detectable in the majority of patients and were excluded) and were further analyzed.



**FIG 2.** Treatment patterns as a result of glucocorticoid therapy. Patients were separately clustered on the basis of glucocorticoid therapy at the time of BAL. *Left*, Subjects taking glucocorticoids (inhaled or oral) versus *right*, subjects not on glucocorticoids. Note similar cytokine patterns are seen in both groups.

### Identification of 4 asthma phenotypes

To reveal the natural underlying groupings, the cytokine concentration data were subjected to unsupervised agglomerative hierarchical clustering (Fig 1). Briefly, this method groups each subject on the basis of mathematical similarities of BAL cytokine concentrations to the others.<sup>17</sup> Initially, each subject is in its own cluster. At each step, the nearest 2 subjects (determined by Pearson correlation as the distance metric) are combined into a higher-level cluster. The iteration continues until all the subjects are grouped. Each row corresponds to a subject, and the individual cytokine values are shown in each column, with green being low-expressing and red being high. From this analysis, 4 groups labeled G1 to G4 could be discerned displaying different patterns of cytokine expression. For example, G1 had high levels of IL-2, G2 had high levels of IL-1Ra, G3 had high levels of IP-10, and G4 had high levels of IL-2R and many other cytokines (Fig 1).

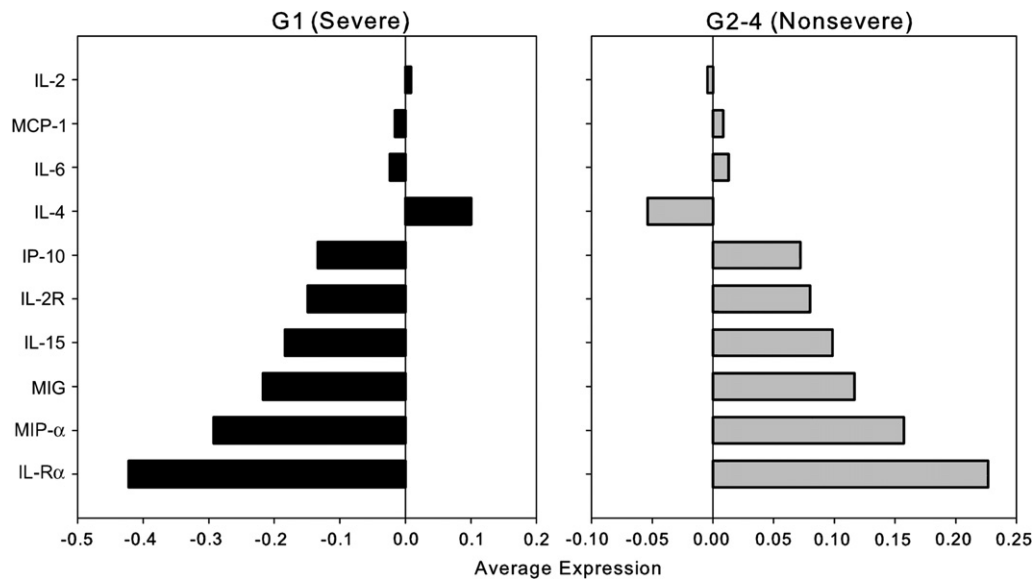
To determine whether the patients within these groups represented biologically distinct subgroups of asthma, the clinical features of the 4 groups were compared with one another. We found that more than 15 different variables were statistically different between these groups (Table II). Importantly, these included cellular features of BAL (pulmonary eosinophils, alveolar macrophages) and lung function measurements (values of lung function, FEV<sub>1</sub> response to bronchodilation and sensitivity to methacholine). To determine further how each group differed from one another, pairwise comparisons between the groups were performed by using multiple comparisons in ANOVA (with Bonferroni correction). This analysis indicated that the

patients in G1 had a significantly reduced FEV<sub>1</sub>, FVC, and FEV<sub>1</sub> improvement after bronchodilator therapy compared with other groups (for the complete pairwise comparison, see this article's Table E1 in the Online Repository at [www.jacionline.org](http://www.jacionline.org)).

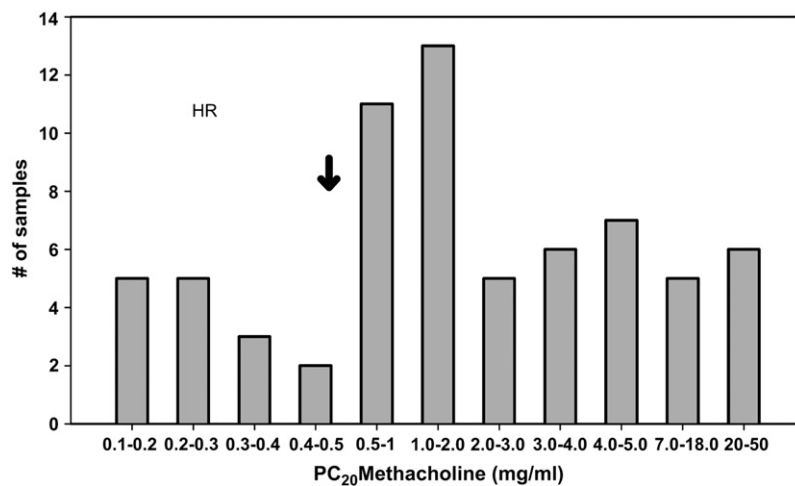
Moreover, G1 was enriched in patients classified as severe by the ATS criteria,<sup>12</sup> with 18 of the 30 patients (60%) assessed by SARP investigators as having severe asthma. G2, the group with the best preservation of lung function was enriched in patients with nonsevere asthma, with only 8 of the 13 (38%) identified as severe by ATS criteria. These findings indicated that BAL cytokine patterns were informative of disease phenotypes as determined by nonoverlapping clinical criteria.

### Analysis of treatment effect on cytokine expression patterns

Because the subjects analyzed in this data set had significant differences in glucocorticoid use, it was important to determine whether cytokine expression patterns were a result of therapy or were reflective of the underlying disease process. For this purpose, data from subjects on inhaled glucocorticoids were analyzed and compared with those not taking glucocorticoids. We compared cytokine expression patterns using both descriptive statistics and unsupervised analyses. There were no significant between-group differences in the expression levels for the 18 cytokines analyzed (see this article's Table E2 in the Online Repository at [www.jacionline.org](http://www.jacionline.org)). Hierarchical clustering performed on both groups separately produced similar patterns as



**FIG 3.** Cytokine classifiers for G1. Shown is a rank-ordered list of the 10 cytokines that minimize cross-validation error for G1 patients with asthma. *Left*, Centroid of G1; *right*, centroid of combined G2-G4 (threshold of 1.2). X-axis is deviation from the overall class centroid.



**FIG 4.** Identification of hyperresponder (*HR*) subjects. Shown is a frequency histogram of the 67 patients in whom PC<sub>20</sub> methacholine sensitivity was measured. Patients with PC<sub>20</sub> methacholine response of <0.5 mg/mL were classified as HR.

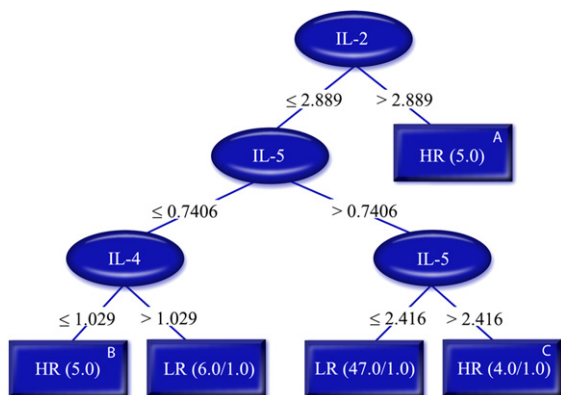
well (Fig 2). From this comparison, we concluded that glucocorticoid medication did not significantly affect the cytokine expression patterns in these subjects with chronic stable asthma.

### Identification of cytokines having greatest effect on group 1 classification

Although clustering based on 18 cytokines identified phenotypically distinct subgroups, we sought to identify the cytokines that most contributed to the clustering result. For this purpose, a feature reduction technique using a robust linear discriminant method known as *shrunk centroids* was used.<sup>16</sup> Centroids characterizes each class mathematically as a vector of its means (known as a *centroid*). Through a reiterated process of training and cross-validation, the number of features was reduced (shrunk)

to those with the smallest variation within the class while still retaining classification accuracy. This identified a smaller set of cytokines that were most important in the decision process.

We performed shrunk centroids to identify minimal features that differentiate patients in G1 from all the other asthma subtypes combined (G2, G3, and G4). This analysis identified 10 cytokines as being most important for identification of this severe group (Fig 3 and Fig E2 in the Online Repository at [www.jacionline.org](http://www.jacionline.org)). The rank order of these cytokines (most informative to least) was IL-1Ra, MIP-1 $\alpha$ , MIG, IL-15, IL-2R, IP-10, IL-4, IL-6, MCP-1, and IL-2. Using this group of cytokines, the subjects could be accurately clustered into the same groups (see this article's Figs E3 and E4 in the Online Repository at [www.jacionline.org](http://www.jacionline.org)). Importantly, reducing this panel of discriminant cytokines further to 9 or 8 significantly increased the



**FIG 5.** Classification and regression tree classification. C.4.5 decision tree was performed on the z-score normalized cytokine data. Shown is the most accurate model. For each node (rectangles), the classification and number of correctly grouped subjects are indicated. The identity of the hyperresponder (HR) class—A, B, or C—is indicated in the upper right corner of each terminal leaf.

misclassification error (see this article’s Fig E1 in the Online Repository at [www.jacionline.org](http://www.jacionline.org)).

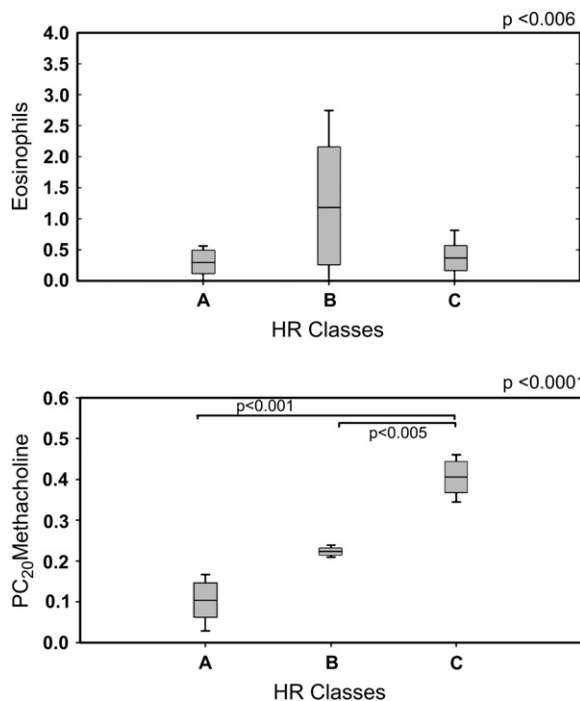
### Classification model for hyperresponsiveness to methacholine

We next sought to identify cytokine patterns that best predicted airway hyperresponsiveness. For this purpose we defined hyperresponders by using an objective measure, sensitivity to methacholine ( $PC_{20} < 0.5$  mg/mL), a clinical feature identified by the unsupervised analysis (Table II). Using this established metric as an objective measure of hyperresponsiveness, 15 of the 67 for which methacholine measurements were performed were identified as hyperresponders and 52 LRs (Fig 4). Pairwise comparison showed that only 1 analyte was significantly different between hyperresponder and LR (IL-2R;  $P < .016$ ; Kruskal-Wallis test). To determine whether combinations of the cytokines could distinguish hyperresponders from LRs, the 18 cytokines were subjected to attribute reduction. This process identified IL-2, IL-4, IL-5, TNF- $\alpha$ , MIG, and RANTES as the most significant attributes. These cytokines were then used in a decision tree–based learning method. The cytokine that best separates the hyperresponder from the LR is selected first, and the process is repeated. Ten-fold cross-validation, a process dividing the data into random training and test sets, is performed to estimate classification error and to prune the tree to prevent overfitting. Performing 10 trials using 10-fold cross-validation resulted in the best model with an average accuracy of 88.1% (Fig 5). The root node was IL-2, which produced a split identifying 5 hyperresponders (class A). Two more hyperresponder classes were identified with low IL-5 and low IL-4 (class B), and the other with high IL-5 and low IL-4 (class C).

The clinical demographics were compared for hyperresponder classes A, B, and C. These groups differed significantly in BAL eosinophils and, interestingly,  $PC_{20}$  methacholine (Fig 6). Hyperresponder class A had low BAL eosinophils and the lowest  $PC_{20}$  methacholine. These results indicated that the decision tree separated 3 distinct hyperresponder subclasses.

### DISCUSSION

A current major challenge in the management of asthma is to identify accurately subtypes that differ in disease pathogenesis



**FIG 6.** Pairwise comparison of methacholine hyperresponder (HR) classes. Demographic variables of methacholine HR classes A, B, and C were compared by ANOVA. Top, BAL eosinophils; bottom,  $PC_{20}$  methacholine.

and response to therapy. Severe asthma is differentiated from mild-moderate disease by age of onset, duration of disease, degree of airflow impairment, cellular inflammation, presence of sinusitis, and history pneumonia.<sup>2,7</sup> However, these clinical features as well as noninvasive measurements of airway inflammation (sputum eosinophils and exhaled nitric oxide) have not resulted in a method for unambiguous separation of clinical phenotype, response to therapy, or disease course.<sup>2,6,18</sup> This experience indicates that a single discriminator that accurately separates different asthma types is not yet available. Recognizing that multiple discriminators may be necessary, we have conducted this study for 2 purposes. First, we sought to determine whether we could use informatics analysis to subgroup subjects on the basis of patterns of BAL cytokine expression. This approach has the potential to identify distinct groups of patients with asthma on the basis of commonalities in inflammatory disease mechanisms. Second, we sought to determine whether clinically objective phenotypes could be predicted on the basis of patterns of cytokine expression. Similar approaches have used mRNA expression patterns to separate different groups of responsive and unresponsive cancers. By using groupings based on mRNA expression profiles, patients can be identified that are indistinguishable by conventional clinical staging criteria.<sup>16,19-21</sup> However, whether this method for molecular phenotyping is generally applicable to other diseases, including asthma, has not been addressed. Although differences in mRNA expression patterns have been identified in patients with asthma, not all mRNAs are translated. Therefore, our focus was on proteins (cytokines), which represent proximal mediators of the inflammatory component of asthma.

Our unsupervised hierarchical clustering analysis indicates that at least 4 phenotypically distinct subgroups of asthma can be identified. Recognizing that a potential confounding variable

could be a result of different treatment regimens, we have sought to determine whether these groupings are significantly affected as a result of chronic glucocorticoid use. Reassuringly, similar groupings are identified in subjects not taking glucocorticoids (Fig 2). We were particularly interested in the phenotype of G1, a group enriched in patients that meet the ATS consensus definition for severe asthma. This group is characterized by increased body mass index, reduced FEV<sub>1</sub>, reduced FVC, and enhanced sensitivity to methacholine. Importantly, our findings indicate that no single cytokine value can be used to separate patients into these groups, but rather, an expression pattern consisting of a minimum of 10 distinct cytokines must be considered.

The biological roles of the chemokines that influence the G1 classification deserve some comment. Relative to the other groups, BAL samples from the subjects in the G1 group are characterized by reduced levels of IL-1 receptor antagonist (IL-1Ra), MIP-1 $\alpha$ , and MIG (and others; Fig. 4), which are known to play various important roles in coordinating cellular trafficking and inflammation in the airways.<sup>22</sup> Reduction of IL-1Ra may be important because this protein competes for IL-1 $\alpha$ / $\beta$  receptor binding blocking IL-1 $\alpha$ / $\beta$ -mediated inflammation. In subjects in G1, a reduction in IL-1Ra may result in enhanced IL-1 signaling, a signaling event that could cause neutrophilic inflammation, features characteristic of severe asthma.<sup>9</sup> Similarly, MIP-1 $\alpha$  (CCL3) is a CC chemokine expressed by macrophages and airway epithelial cells to induce chemotaxis of CD8 T lymphocytes and eosinophils and whose expression is enhanced in the asthmatic lung.<sup>23</sup> Again, reduction in MIP-1 $\alpha$ /CCL3 may result in neutrophil-predominant inflammation. Although MIP-1 is acutely inhibited by glucocorticoids, the level of MIP-1 in BAL from patients on glucocorticoid therapy is not statistically different from patients not taking glucocorticoids (see this article's Table E2 in the Online Repository at [www.jacionline.org](http://www.jacionline.org)). The IFN- $\gamma$  inducible MIG (CXCL9) significantly reduces airway hyperresponsiveness and eosinophil accumulation in animal models of allergen challenge.<sup>24</sup> MIG diminishes IL-4 and enhances IL-12 levels, directing activated T cells toward a T<sub>H</sub>1 phenotype. Therefore, the reduction in these cytokines in the G1 group relative to the patients with less severe asthma is biologically plausible with our understanding of the pathogenesis of severe asthma.

Although most of the molecular classifiers that have been produced have been based on mRNA expression patterns, we suggest that protein expression profiles may be more useful markers of disease than gene expression patterns. This may be particularly true when the proteins themselves play important roles in the underlying disease process, such as cytokines in asthma, and therefore represent bona fide biomarkers.

Previous work examining expression of selected cytokines has associations with asthma severity; however, these associations are not strong and have not been widely replicated. For example, IL-8 was shown to be enhanced in patients with severe asthma during an acute exacerbation and correlates with the number of neutrophils.<sup>25</sup> Our study was conducted on patients with stable asthma, and the processes involved in chronic inflammatory state may be different from those producing acute exacerbations. Other studies have shown increased IL-2 and IL-4 levels in patients with severe asthma.<sup>26</sup> Although we were unable to demonstrate a difference in BAL concentrations for IL-2 or IL-4 between patients with severe and nonsevere asthma in our data set (see this article's Fig E4 in the Online Repository at [www.jacionline.org](http://www.jacionline.org)), these cytokines do contribute to a model predicting methacholine hyperresponders.

Our attempt to classify methacholine hyperresponders has yielded 2 important findings. First, methacholine hyperresponders can be accurately separated from LRs on the basis of cytokine profiles in BAL; and second, the group of methacholine hyperresponders consists of at least 3 phenotypically distinct classes. Hyperresponder class A is a group characterized by low levels of IL-2, IL-4, and IL-5 and, as a group, contains those subjects with the lowest PC<sub>20</sub> methacholine responses in our study. It will be interesting to extend this analysis to larger independent data sets to determine whether the subjects in the hyperresponder classes have distinct clinical outcomes.

The feature reduction analysis has shown that at least 5 cytokines can be used to separate hyperresponders from LRs in our data set. It is also important to note that there are important differences in the groups of cytokines that identify G1 from those that identify hyperresponders. Specifically, IL-5, secreted into the BAL in response to allergen challenge, and a molecule important in eosinophil recruitment and survival, was identified as an important classification variable in hyperresponders, but not for G1. This observation may explain why single cytokine values alone may not be useful discriminators in asthma. Further exploration of these data are underway.

Our studies provide a first proof of concept that informative patterns of cytokines can be detected and interpreted in BAL from patients with asthma and may contribute to more objective classification of disease type. We interpret these findings to suggest that subjects with apparently similar clinical characteristics are in fact composed of heterogeneous subtypes that can be further distinguished on the basis of BAL cytokine profiles. However, before these findings can be applied to patient classification or management, a number of critical questions remain: What is the reproducibility of serial measurement of BAL cytokines? How stable are the individual phenotypes? Do these distinct subtypes differ in clinical outcome or response to therapeutic interventions? Importantly, these findings and models will need to be tested in an independent study population. For these reasons, the findings of this study are not ready for translation into the clinic. Nevertheless, our findings indicate that important new diagnostic and prognostic information is available in airway fluids and indicate that future research in biomarker identification will be informative.

We thank Douglas Curran-Everett, PhD, for assistance with data mining.

**Clinical implications: Definition of asthmatic phenotypes will aid in clinical investigation of the etiology of and intervention for asthma.**

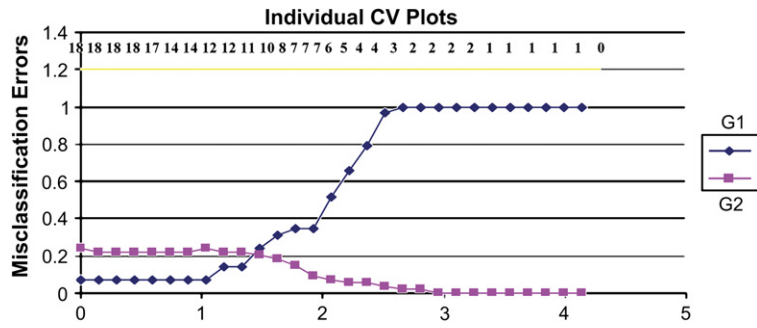
#### REFERENCES

1. Busse WW, Lemanske RF. Asthma. *N Engl J Med* 2001;344:350-62.
2. Moore WC, Peters SP. Severe asthma: an overview. *J Allergy Clin Immunol* 2006; 117:487-94.
3. Godard P, Chanez P, Siraudin L, Nicoloyannis N, Duru G. Costs of asthma are correlated with severity. *Eur Respir J* 2002;19:61-7.
4. Serra-Battles J, Plaza V, Morejon E, Comella A, Bruges J. Costs of asthma according to the degree of severity. *Eur Respir J* 1998;12:1322-6.
5. Wenzel SE, Schwartz LB, Langmack EL, Halliday JL, Trudeau JB, Gibbs RL, et al. Evidence that severe asthma can be divided pathologically into two inflammatory subtypes with distinct physiologic and clinical characteristics. *Am J Respir Crit Care Med* 1999;160:1001-8.
6. Chanez P, Wenzel SE, Anderson GP, Anto JM, Bel EH, Boulet LP, et al. Severe asthma in adults: what are the important questions? *J Allergy Clin Immunol* 2007;119:1337-48.

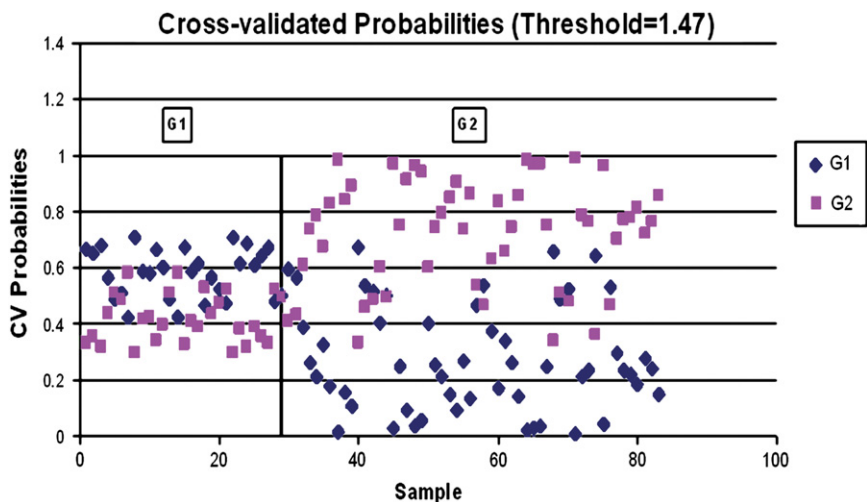
7. Moore WC, Bleecker ER, Curran-Everett D, Erzurum SC, Ameredes BT, Bacharier L, et al. Characterization of the severe asthma phenotype by the National Heart, Lung, and Blood Institute's Severe Asthma Research Program. *J Allergy Clin Immunol* 2007;119:405-13.
8. Wenzel S. Pathology of difficult asthma. *Paediatr Respir Rev* 2003;4:306-11.
9. Jatakanon AN, Uasuf CA, Maziak WA, Lim SA, Chung KF, Barnes PJ. Neutrophilic inflammation in severe persistent asthma. *Am J Respir Crit Care Med* 1999;160:1532-9.
10. Lemiere C, Ernst P, Olivenstein R, Yamauchi Y, Govindaraju K, Ludwig M, et al. Airway inflammation assessed by invasive and noninvasive means in severe asthma: eosinophilic and noneosinophilic phenotypes. *J Allergy Clin Immunol* 2006;118:1033-9.
11. Wenzel SE, Busse WW. Severe asthma: lessons from the Severe Asthma Research Program. *J Allergy Clin Immunol* 2007;119:14-21.
12. Proceedings of the ATS Workshop on Refractory Asthma: Current Understanding, Recommendations, and Unanswered Questions. *Am J Respir Crit Care Med* 2000;162:2341-51.
13. Hankinson J, Odencrantz J, Fedan K. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med* 1999;159:179-87.
14. Bronchoalveolar Lavage Cooperative Steering Committee. BAL constituents in healthy individuals, idiopathic pulmonary fibrosis, and selected comparison groups. *Am Rev Respir Dis* 1990;141:S169-202.
15. Quinlan JR. C4.5: Programs for machine learning. San Mateo (CA): Morgan Kaufmann; 1993.
16. 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.
17. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 1998;95:14863-8.
18. Deykin A. Biomarker-driven care in asthma: are we there? *J Allergy Clin Immunol* 2006;118:565-8.
19. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 1999;286:531-7.
20. Bhattacharjee A, Richards WG, Staunton J, Li C, Monti S, Vasa P, et al. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc Natl Acad Sci U S A* 2001;98:13790-5.
21. Kapp A, Jeffrey S, Langerod A, Borresen-Dale AL, Han W, Noh DY, et al. Discovery and validation of breast cancer subtypes. *BMC Genomics* 2006;7:231.
22. Pease JE, Williams TJ. Chemokines and their receptors in allergic disease. *J Allergy Clin Immunol* 2006;118:305-18.
23. Alam R, York J, Boyers M, Stafford S, Grant JA, Lee J, et al. Increased MCP-1, RANTES, and MIP-1 $\alpha$  in bronchoalveolar lavage fluid of allergic asthmatic patients. *Am J Respir Crit Care Med* 1996;153:1398-404.
24. Thomas MS, Kunkel SL, Lukacs NW. Regulation of cockroach antigen-induced allergic airway hyperreactivity by the CXCR3 ligand CXCL9. *J Immunol* 2004;173:615-23.
25. Fahy JV, Kim KW, Liu J, Boushey HA. Prominent neutrophilic inflammation in sputum from subjects with asthma exacerbation. *J Allergy Clin Immunol* 1995;95:843-52.
26. Leung DY, Martin RJ, Szefer SJ, Sher ER, Ying S, Kay AB, et al. Dysregulation of interleukin 4, interleukin 5, and interferon gamma gene expression in steroid-resistant asthma. *J Exp Med* 1995;181:33-40.

### Availability of Journal back issues

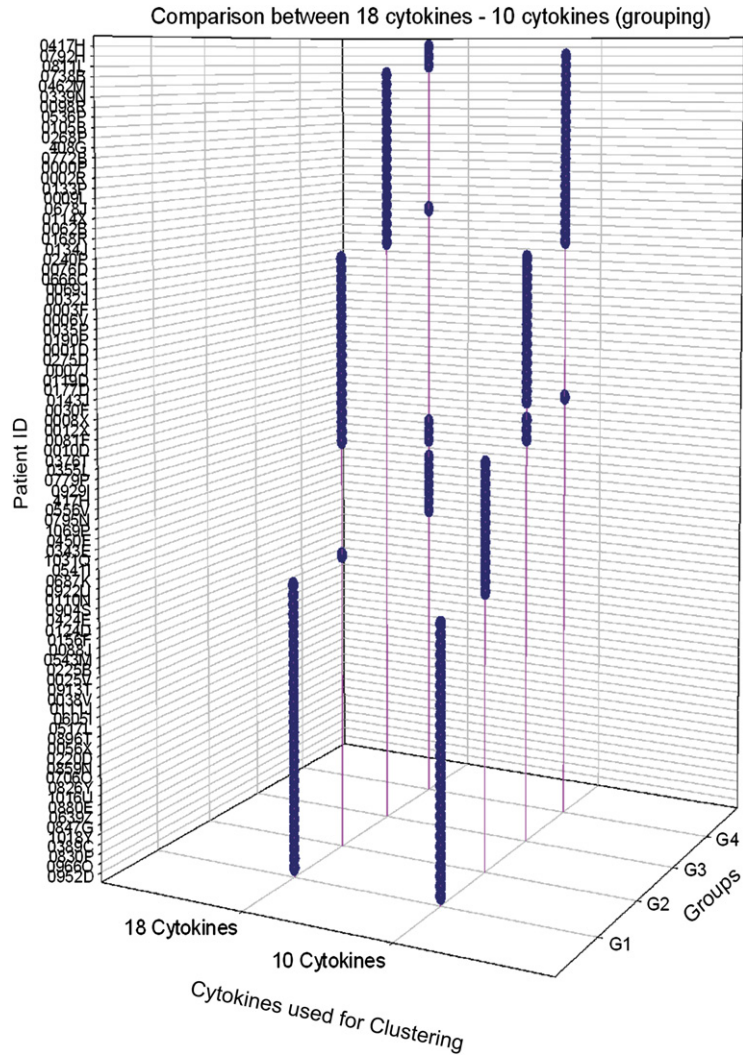
As a service to our subscribers, copies of back issues of *The Journal of Allergy and Clinical Immunology* for the preceding 5 years are maintained and are available for purchase until inventory is depleted. Please write to Elsevier Inc, Subscription Customer Service, 6277 Sea Harbor Dr, Orlando, FL 32887, or call (800) 654-2452 or (407) 345-4000 for information on availability of particular issues and prices.



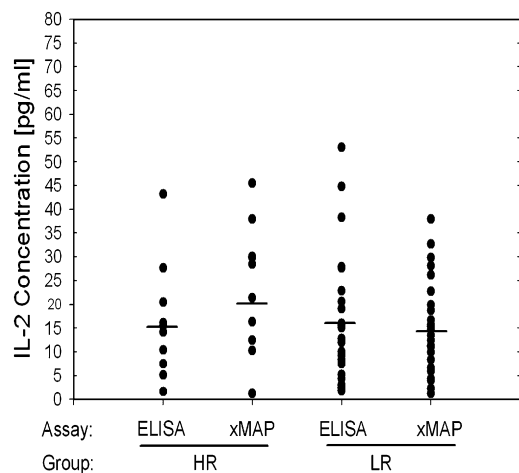
**FIG E1.** Misclassification error for cytokines to group G1 versus G2 (representing G2, G3, and G4 combined). The misclassification error is calculated at each step after reducing the number of cytokines used in the classification with the method of shrunken centroids. Y-axis is the misclassification error and x-axis the threshold. The point where the misclassification error is least is considered the minimal number of cytokines that discriminate G1 versus G2. This occurs at the threshold where the feature reduction from 18 cytokines to 10 cytokines occurs. Above a threshold of value of 1.4, the misclassification error rate rises sharply. The shrunken centroid lists the 10 cytokines identified from the misclassification plot, which is later used for analysis. *CV*, Crossvalidation.



**FIG E2.** Cross-validated probabilities for 2 class prediction. Y-axis is the probability of each subject being classified into group G1 or the combined group G2-G4 (termed G2) on the basis of the linear discriminant classifier. X-axis represents each individual subject. The predicted class is the class having the largest probability (G1 or G2) as indicated in the plot. The correct class is indicated in *white box* for the group.



**FIG E3.** Comparison of 4 groups (G1, G2, G3, and G4) using 18 cytokines and 10 cytokines identified by feature reduction (the identity of the 10 cytokines is shown in Fig 3). For each subject in the y-axis, grouping into each of 4 groups is compared. Note that the subjects in the 4 groups are largely similar, irrespective of whether 18 or 10 cytokines are used for the groupings. This shows that we are still able to identify clearly the same 4 groups with reduced features (10 cytokines).



**FIG E4.** Validation of IL-2 measurements. BAL samples from 10 hyper-responder and 26 LR subjects were measured by using IL-2 ELISA. Shown is the *scattergram* for each group by the type of assay used. The *horizontal bar* is the mean value for each group. The measurements for IL-2 are highly similar regardless of the assay used. *xMAP*, Multiplex bead assay.

**TABLE E1.** Multiple pairwise comparison among the 4 groups identified by unsupervised clustering\*

Feature	Group (means $\pm$ SDs)				Pairwise comparison ( <i>P</i> values)					
	G1	G2	G3	G4	1 vs 2	1 vs 3	1 vs 4	2 vs 3	2 vs 4	3 vs 4
Body mass index	31.95 $\pm$ 8.45	26 $\pm$ 4.16	29.1 $\pm$ 5.8	27.4 $\pm$ 6.7	.02	NS	NS	NS	NS	NS
Macrophages	82.18 $\pm$ 15.4	89.7 $\pm$ -6.6	81.6 $\pm$ 31.33	89.4 $\pm$ 5.4	NS	NS	NS	NS	NS	NS
Lymphocytes	13.07 $\pm$ 12.9	6.55 $\pm$ 3.6	11.9 $\pm$ 9.93	6.7 $\pm$ 3.4	NS	NS	NS	NS	NS	NS
Eosinophils	-0.69 $\pm$ 0.5	-0.78 $\pm$ 0.3	-0.85 $\pm$ 0.36	-0.35 $\pm$ 0.5	NS	NS	NS	NS	NS	NS
FEV <sub>1</sub>	3 $\pm$ 0.6	3.4 $\pm$ 1	2.7 $\pm$ 0.6	2.7 $\pm$ 0.9	.000	NS	NS	.047	NS	NS
FEV <sub>1pp</sub>	72.3 $\pm$ 21.6	89.3 $\pm$ 16.3	83.2 $\pm$ 17.2	80 $\pm$ 19.5	.014	NS	NS	NS	NS	NS
FVC	3.18 $\pm$ 0.84	4.6 $\pm$ 1.2	3.7 $\pm$ 0.8	3.7 $\pm$ 1.1	.000	NS	NS	.049	NS	NS
FVCpp	86 $\pm$ 15.7	100 $\pm$ 17	93.5 $\pm$ 17.7	89.3 $\pm$ 15.8	.024	NS	NS	NS	NS	NS
Maximum FEV <sub>1</sub> /MPV	2.79 $\pm$ 0.6	3.9 $\pm$ 1.2	3 $\pm$ 0.6	3 $\pm$ 1	.000	NS	NS	.024	.024	NS
Maximum FVC/MPV	3.7 $\pm$ 1	5 $\pm$ 1.3	4 $\pm$ 0.8	3.8 $\pm$ 1.1	.001	NS	NS	.017	.013	NS
Maximum FVCpp/MPVLung	100.4 $\pm$ 14.3	109.4 $\pm$ 15	100 $\pm$ 17.3	90.8 $\pm$ 14.6	NS	NS	NS	NS	.024	NS
logPC <sub>20</sub> methacholine	-0.1 $\pm$ 0.5	0.5 $\pm$ 0.6	0.3 $\pm$ 0.65	0.25 $\pm$ 0.5	.004	NS	NS	NS	NS	NS
FEV <sub>1</sub> reversal (%)	15.3 $\pm$ 16	7.4 $\pm$ 5.6	7.2 $\pm$ 6.5	9.9 $\pm$ 7.5	NS	NS	NS	NS	NS	NS
FEV <sub>1</sub> albuterol reversal	2.6 $\pm$ 0.6	3.7 $\pm$ 1	2.9 $\pm$ 0.6	2.9 $\pm$ 1	.000	NS	NS	.034	.035	NS

pp, Percent predicted; MPV, maximum postbronchodilator value.

\*The clinical features for each group (G1, G2, G3, and G4) shown in Fig 1 were compared against each other. For each group, the mean and SD of each feature is calculated. At right, the numeric *P* value of the pairwise comparison is shown for groups G1 vs G2, and so forth.

**TABLE E2.** Comparison of cytokine expression for low vs high steroid use\*

<b>Cytokine</b>	<b>P value</b>
IL-1Ra	.943
IL-2	.202
IL-2R	.058
IL-4	.472
IL-5	.210
IL-6	.589
IL-8	.414
IL-15	.088
IL-17	.282
TNF- $\alpha$	.295
IFN- $\gamma$	.404
MIP1- $\alpha$	.514
MIP-1 $\beta$	.212
MIG	.121
Eotaxin	.696
RANTES	.291
MCP-1	.446
IP-10	.674

\*Tabulated is a group comparison of the BAL concentrations of each cytokine in subjects with asthma taking inhaled corticosteroids versus those not taking corticosteroids. Values were compared by nonparametric Kruskal-Wallis test with a *P* value of .05 considered significant. No cytokine reaches statistical significance as a function of steroid use. These data indicate that no statistical differences are found between patients taking inhaled corticosteroids versus those not taking corticosteroids.