Association between Pulmonary Function and Sputum Biomarkers in Cystic Fibrosis

Nicole Mayer-Hamblett^{1,2}, Moira L. Aitken³, Frank J. Accurso⁴, Richard A. Kronmal^{2,5}, Michael W. Konstan⁶, Jane L. Burns¹, Scott D. Sagel⁴, and Bonnie W. Ramsey^{1,2}

¹Department of Pediatrics, University of Washington, Seattle, Washington; ²Cystic Fibrosis Foundation Therapeutics Development Network Coordinating Center, Children's Hospital and Regional Medical Center, Seattle, Washington; ³Department of Medicine, University of Washington, Seattle, Washington; ⁴Department of Pediatrics, University of Colorado Health Sciences Center, Denver, Colorado; ⁵Department of Biostatistics, University of Washington, Seattle, Washington; ⁶Rainbow Babies and Children's Hospital and Case Western Reserve University School of Medicine, Cleveland, Ohio

Rationale: Sputum biomarkers of infection and inflammation are noninvasive measures that enable quantification of the complex pathophysiology of cystic fibrosis (CF) lung disease. Validation of these biomarkers as correlates of disease severity is a key step for their application.

Objectives: We constructed a large database from four multicenter studies to quantify the strength of association between expectorated sputum biomarkers and FEV₁.

Methods: FEV₁ (range, 25–120% predicted) and quantitative data on expectorated sputum biomarkers including free neutrophil elastase, IL-8, neutrophils, *Pseudomonas aeruginosa*, and *Staphylococcus aureus* were obtained from 269 participants (ages, 9–54 years) from 33 centers. Cross-sectional and longitudinal statistical analyses were performed to estimate associations between the markers and FEV₁, including the use of multivariable analyses.

Results: Elastase was negatively correlated with FEV₁ (correlation [*r*] = −0.35; 95% confidence interval [CI]: −0.46, −0.22). On average, patients with CF who differed in their elastase measurements by 0.5 log differed in their FEV₁ values by −7.3% (95% CI: −9.7, −4.6). Neutrophil counts and IL-8 were also each negatively correlated. In a multivariable regression, elastase and neutrophil counts were able to explain the majority of variation in FEV₁. Elastase was further shown to have a significant longitudinal association with FEV₁, specifically a −2.9% decline in FEV₁ (95% CI: −5.0, −0.9) per 1-log increase in elastase. Although correlated with FEV₁, bacterial densities were unable to explain clinically meaningful differences in FEV₁ within and across patients.

Conclusions: These data support the role of sputum biomarkers as correlates of disease severity in a diverse CF population.

Keywords: cystic fibrosis; pulmonary function; sputum; infection; inflammation

The interrelated pulmonary manifestations of cystic fibrosis (CF) involve both neutrophil-dominated airway inflammation and chronic airway infection by pathogens such as *Pseudomonas aeruginosa* (1, 2). Although the temporal relationship between infection and inflammation remains unclear in early disease (3), later in disease there is a self-perpetuating cycle of airway ob-

AT A GLANCE COMMENTARY

Scientific Knowledge on the Subject

There are limited data from single-center studies supporting the association between sputum markers and clinical parameters in CF. Associations have yet to be described in a more diverse CF population from multiple centers.

What This Study Adds to the Field

Sputum biomarkers, including elastase, neutrophil counts, and IL-8, are correlates of disease severity in diverse populations of patients with CF.

struction, chronic bacterial infection, and vigorous inflammation that results in structural damage to the airway (4).

Given the importance of infection and inflammation in the pathophysiology of CF lung disease, noninvasive sputum-derived measures of these processes may play a key role in monitoring disease progression and predicting future clinical course (5). These measures may also serve a critical role as outcome measures in clinical trials assessing early biological efficacy of new antimicrobial and antiinflammatory agents (6). They are often denoted as biomarkers, which suggests they are sensitive for measuring biological activity and response to therapy. An important step in the validation process of these biomarkers is to establish that they are correlates of established clinical outcome measures.

There are limited but compelling data from small, single-center studies that support the association between sputum biomarkers and clinical parameters (7, 8). Sagel and coworkers (7) reported results from a single-center study that demonstrated significant correlations between pulmonary function and induced sputum measurements of IL-8, neutrophil elastase, total cell counts, and neutrophil counts but nonsignificant associations between pulmonary function and bacterial colony counts. However, this study was based on only 19 participants with CF who, with the exception of 1 participant, had FEV₁ predicted values greater than 70%. Similarly, Kim and coworkers (8) demonstrated significant correlations between pulmonary function and both IL-8 and myeloperoxidase in a small, single-center study of 16 patients with CF. A critical step in the validation of these sputum biomarkers as correlates of disease severity is to demonstrate significant associations between these markers and clinical parameters such as FEV₁ among a diverse CF population.

(Received in original form September 21, 2006; accepted in final form January 18, 2007)

Supported by the Cystic Fibrosis Foundation and by the National Center for Research Resources (NCRR grants MO1-RR00037 and MO1-RR00069).

Correspondence and requests for reprints should be addressed to Nicole Mayer-Hamblett, Ph.D., Department of Pediatrics, University of Washington, 4800 Sand Point Way N.E., Box 5371, Seattle, WA 98105-0371. E-mail: nicole.hamblett@seattlechildrens.org

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

Am J Respir Crit Care Med Vol 175. pp 822–828, 2007

Originally Published in Press as DOI: 10.1164/rccm.200609-1354OC on January 18, 2007
Internet address: www.atsjournals.org

Through the Cystic Fibrosis Foundation Therapeutics Development Network (CFF TDN), data on both clinical end points and expectorated sputum markers of infection and inflammation have been obtained from multiple clinical studies. These studies used standardized sputum processing procedures and centralized laboratories, which provides a unique opportunity to combine the data from these studies to investigate associations between sputum biomarkers and pulmonary function across multiple study centers. These data provide a more diverse and larger CF population with which to explore these associations than has been examined in previously published small, single-center studies. The main objectives of this study were to quantify the cross-sectional association between FEV₁ and each expectorated sputum marker, and to determine which sputum biomarkers can best explain the interpatient variation in pulmonary function. A secondary objective was to describe longitudinal associations between changes in FEV₁ and corresponding changes in each marker, using a subset of patients for whom repeated measurements over time were available.

METHODS

Study Cohort

The cohort was composed of 269 participants from 33 centers in 4 CF studies. These studies consisted of a randomized, placebo-controlled trial of azithromycin (9), a randomized, placebo-controlled trial of IFN- γ -1b (10), a randomized, placebo-controlled trial of tgAAVCF (11), and a study measuring sputum biomarkers before and after administration of intravenous antibiotics (12). Select data are anonymously included in the TDN Data Bank, approved by the Institutional Review Board at Children's Hospital and Regional Medical Center (Seattle, WA).

Data Collection

Sputum collection and processing were performed according to standard procedures developed by the CFF TDN and at core TDN laboratories (13) (see the online supplement). The following data from preintervention baseline visits were obtained: FEV₁% predicted (14) and expectorated sputum markers IL-8 (picograms per milliliter), free neutrophil elastase (micrograms per milliliter), neutrophil counts (cells per milliliter and percent), and *P. aeruginosa* and *Staphylococcus aureus* densities (colony-forming units per gram of sputum). These data were also available in the TDN Data Bank at postbaseline visits for randomized placebo patients (n = 78 of 269). These repeated measures were used to improve estimation of cross-sectional correlations and enable longitudinal analyses.

Statistical Methods

Sputum markers (except percent neutrophils) were logarithmically transformed. For neutrophil counts and percents, associations with FEV₁ were summarized using Pearson correlations (15). To obtain correlations for markers with some values either at or below a limit of detection (LOD) (elastase and IL-8) or nondetectable (*P. aeruginosa* and *S. aureus*), the following linear regression model was fit for each marker:

$$\text{FEV}_1 = \beta_0 + \beta_1 \times I (\text{marker not detected}) + \beta_2 \times I (\text{marker detected}) \times (\text{marker value})$$

where I (marker not detected) is equal to 1 if the marker value is below an LOD (applicable for IL-8 and elastase) or is not present (applicable for *P. aeruginosa* and *S. aureus* densities when the density was 0), and is equal to 0 otherwise. In this model, $\beta_0 + \beta_1$ is a point estimate of the average FEV₁ among all observations for which the marker is not detected or below an LOD, and $\beta_0 + \beta_2 \times (\text{marker value})$ is the regression line that describes the association when the marker value is detected. This approach allows for the possibility that participants with nondetectable values represent a clinically distinct group whose FEV₁ does not fall on the same continuum as the participants with detectable values. The coefficient of determination from this regression model,

R^2 is the proportion of variance in FEV₁ explained by this model and R is an overall estimate of the correlation between the marker and FEV₁. Note that when the marker has all detectable values, the correlation estimated from this model equals the Pearson correlation.

Multivariable regression was used to investigate which set of markers could explain the majority of the interpatient variation in FEV₁, and linear regression was used to estimate longitudinal associations between each marker and FEV₁. For all analyses, bootstrap methods were used to derive 95% confidence intervals (16), which accounts for the repeated measurements from placebo patients and provides accurate estimates of variability. Further details on the statistical methods can be found in the online supplement.

RESULTS

Description of the Study Cohort

The baseline characteristics of the study cohort are summarized in Table 1. The cohort represents a diverse population of patients with CF, ranging in age from 9 to 54 years and with FEV₁% predicted values ranging from 25 to 120%. Of the 269 participants, 78 were randomized to placebo and had more than one expectorated sputum measurement available for a particular marker. Note that some markers were not available from all studies, and Table 2 provides an overview of the availability of marker data by participating study and overall.

Cross-sectional Association between FEV₁ and Each Sputum Biomarker

Figure 1 displays scatterplots of each sputum marker versus FEV₁% predicted. For IL-8, free elastase, *P. aeruginosa*, and *S. aureus*, the regression approach we implemented calculates the average FEV₁% predicted among all the nondetectable (or zero density) values of the marker, which is plotted in Figure 1. Importantly, and unique to our approach, the average FEV₁% predicted among the nondetectable values is allowed to differ from the extrapolated value of a regression line derived from the detectable values. For both IL-8 and free elastase, the average FEV₁% predicted among the values at the LOD are below the average value that would have been predicted from the regression model based on the values above the LOD. For the purposes of graphical display, we note that nondetectable values for IL-8 and elastase were plotted at the LOD whereas in truth these values could lie in any interval between zero and the LOD.

TABLE 1. CHARACTERISTICS OF STUDY COHORT BY PARTICIPATING STUDY AND OVERALL

	Expectorated Sputum Cohort				
	Study 1*	Study 2*	Study 3*	Study 4*	Total
No. of participants	33	64	148	24	269
No. of sites	8	12	22	6	33
Female, %	61	63	45	54	52
Age, %					
6–12 yr	9	0	5	0	4
> 12–18 yr	48	30	40	25	37
> 18 yr	42	70	55	75	59
Age (yr), range	9–43	12–45	9–51	14–54	9–54
FEV ₁ % predicted, %					
25–50	15	11	28	0	20
50–65	36	33	22	21	26
65–80	33	16	20	29	22
> 80	15	41	30	50	32
FEV ₁ % predicted, range	40–115	38–120	25–120	52–120	25–120
Body mass index, mean (SD)	19 (2.8)	21 (3.3)	20 (3.1)	22 (2.9)	21 (3.1)

* Study 1: Pre- and postintravenous antibiotic study (12); study 2: IFN- γ -1b study (10); study 3: azithromycin study (9); study 4: tgAAVCF study (11).

TABLE 2. AVAILABILITY OF SPUTUM MARKER DATA BY PARTICIPATING STUDY AND OVERALL

	Study 1*	Study 2*	Study 3*	Study 4*	Total
No. of unique participants	33	64	148	24	269
IL-8					
No. of observations	22	109	190	0	322
No. of participants	22	53	141	0	217
Free elastase					
No. of observations	18	107	190	0	315
No. of participants	18	51	141	0	210
Neutrophil count					
No. of observations	20	89	0	0	109
No. of participants	20	45	0	0	65
Percent neutrophils					
No. of observations	20	92	0	0	112
No. of participants	20	47	0	0	67
<i>P. aeruginosa</i> density					
No. of observations	27	97	165	36	325
No. of participants	27	54	124	24	229
<i>S. aureus</i> density					
No. of observations	27	97	165	0	289
No. of participants	27	54	124	0	205

* Study 1: Pre- and postintravenous antibiotic study (12); study 2: IFN- γ -1b study (10); study 3: azithromycin study (9); study 4: tgAAVCF study (11).

Thus, what is shown is the minimal departure between the set of nondetectable values and the regression line derived from the detectable values. In contrast, for *P. aeruginosa* density and *S. aureus* density, the average FEV₁% predicted among the ob-

servations with zero density falls much closer to the extrapolated regression line in that range of the data.

Figure 2 displays the associated correlation estimates and 95% confidence intervals between each sputum marker and FEV₁% predicted. As seen in Figure 2, IL-8, free elastase, neutrophil counts, percent neutrophils, and *P. aeruginosa* density were all significantly negatively associated with FEV₁% predicted. Note that when the one outlying observation in the neutrophil count data was removed (see Figure 1), the correlation estimate was slightly lower at -0.26 , but still statistically significant (95% confidence interval [CI], -0.42 to -0.08). *S. aureus* density was positively associated with FEV₁% predicted.

Although significant correlations were identified in our analysis, it is important to compare the relative impact of each biomarker in explaining the interpatient variation in FEV₁% predicted. Figure 3 displays the average difference in FEV₁% predicted associated with an average difference in each of the biomarkers, as derived from the regression models used to estimate the correlations. Figure 3 presents these associations among the patients with detectable values for each marker. Free elastase had the largest associated difference, such that patients with CF who differed in their free elastase measurements by 0.5 log also differed in their FEV₁% predicted values by an average of -7.3% (95% CI, -9.7 to -4.6). There were no clinically meaningful differences in terms of FEV₁ among patients who had different nonzero bacterial densities of *P. aeruginosa* and/or *S. aureus*. Figure 3 displays the average differences in FEV₁ associated with 1-log differences in sputum densities between patients, and

Figure 1. Scatterplots of FEV₁% predicted versus each sputum marker. For IL-8, free elastase, *Pseudomonas aeruginosa* density, and *Staphylococcus aureus* density, the plots display both the average FEV₁% predicted among the nondetectable values (solid diamonds) and the regression lines corresponding to the detectable values. Both the point estimate and regression line contribute to the estimate of the overall correlation of FEV₁ with each marker. For neutrophil counts and percent neutrophils, the regression lines pertaining to the Pearson correlation are displayed.

Figure 2. Correlation estimates and 95% confidence intervals for the association between FEV₁% predicted and each sputum marker. CI = confidence interval.

we note that even smaller differences in FEV₁ are present with a 0.5-log difference in sputum densities, as contrasted to the inflammatory marker results.

Multivariable Regression Analyses of the Cross-sectional Relationship between FEV₁ and the Sputum Biomarkers

As seen in Figure 2, free elastase had the highest correlation with FEV₁% predicted as estimated by the regression model approach and explained 12% of the total variation. After accounting for the variation in FEV₁% predicted explained by free elastase, neutrophil counts were able to explain the greatest additional proportion of variation (10% additional). IL-8 and free elastase were in fact highly correlated (correlation, 0.72), which explains why the addition of IL-8 to the model after

Figure 3. Estimates of the average difference in FEV₁% predicted associated with an average difference in each sputum marker. Differences in FEV₁% predicted are derived from the regression lines displayed in Figure 1. Because neutrophil percent is on a different scale from the other markers, only results for neutrophil counts are presented.

adjusting for free elastase did not explain any additional variation. Free elastase and neutrophil counts were therefore able to explain 22% of the total variation in FEV₁% predicted and, after adjusting for these two markers, the others could not explain more than 2% additional variation and were therefore excluded from the model. The corresponding correlation coefficient from this model was -0.46 (95% CI, -0.64 to -0.30), indicating that these two markers combined provide a stronger correlation than just one alone. The multivariate model including free elastase and neutrophil counts is displayed in Table 3.

Longitudinal Associations between FEV₁ and Each Sputum Biomarker

Longitudinal associations were explored, using the subset of the cohort that was randomized to placebo and for which measurements were available at baseline as well as at least one postbaseline time point. The data for this analysis comprise a much smaller set of measurements available for each marker as compared with the previous cross-sectional analysis. Despite this, there was a significant association identified between an increase in free elastase and decrease in FEV₁ (Figure 4). Specifically, a -2.9% decline in FEV₁ (95% CI, -5.0 to -0.9) was associated with a 1-log increase in free elastase; a -1.5% decline (95% CI, -2.5 to -0.5) was associated with a 0.5-log increase in free elastase. There is notable variability in the data but, importantly, these data are only from patients receiving placebo and thus are not meant to reflect a more systematic change across patients, as would be expected in an intervention trial. No other markers were found to have a significant longitudinal relationship with FEV₁ (see Figures E1–E4 in the online supplement), but these exploratory analyses are hindered for some markers by the limited number of measurements available.

Further longitudinal analysis of free elastase provided estimates of variability for the change in this marker over time, one of the key components for deriving sample size calculations. Time points for the longitudinal analyses ranged from approximately 1 week to 6 months, and on average the variability in the change in free elastase was estimated as a standard deviation of 0.5 for studies of 3 months or less, and a standard deviation of 0.7 was calculated for studies of approximately 6 months (Figure E5). Figure 5 displays the sample size per group required for a two-sample comparison of the change in free elastase for a range of effect sizes and two standard deviation estimates. The previous cross-sectional analysis demonstrated that an effect size

TABLE 3. COEFFICIENTS AND CORRELATION ESTIMATES FOR MULTIVARIABLE MODEL, INCLUDING FREE ELASTASE AND NEUTROPHIL COUNTS

	Coefficient Estimate (SE)
Intercept	126.4 (19.40)
I (Elastase $\leq$ LOD)	-13.2 (3.92)
I (Elastase $>$ LOD) $\times$ elastase value	-9.5 (7.60)
Neutrophil count	-4.3 (3.07)
Percentage of explained variance	21.6% (95% CI: 8.7%, 40.4%)
Overall correlation	-0.46 (95% CI: -0.64 , -0.30)

Definition of abbreviations: 95% CI = 95% confidence interval; LOD = limit of detection.

The model is: $\text{FEV}_1\% \text{ predicted} = 126.4 - 13.2 \times I(\text{elastase} \leq \text{LOD}) - 9.5 \times I(\text{elastase} > \text{LOD}) \times \text{elastase value} - 4.3 \times \text{neutrophil count}$, where $I(\text{elastase} \leq \text{LOD})$ is an indicator equal to 1 when elastase is at or below the LOD and 0 otherwise. Although data from all studies were used to select the most informative model, the final model was derived on the basis of data from study 1 and study 2 because these were the only two studies with both neutrophil counts and elastase measured for each participant.

Figure 4. Longitudinal association between change in FEV₁% predicted and change in free elastase. Changes in elastase that represent an increase from the limit of detection (LOD) (positive changes) and a decrease to the LOD (negative changes) are denoted by solid circles.

of 0.5 log may represent a target effect size associated with clinically meaningful acute improvements in FEV₁. However, a smaller effect size may be just as important for antiinflammatory therapies meant to slow the rate of FEV₁ decline. In particular, 44 patients would be required per group to detect a 0.3-log effect size assuming a standard deviation of 0.5 and 80% power, whereas 59 patients per group would be needed assuming a standard deviation of 0.7.

DISCUSSION

We present the largest published study to investigate associations between pulmonary function and sputum markers of infection and inflammation in CF. Our diverse population of patients with CF was retrospectively obtained from 33 centers participating

in 4 studies conducted through the CF TDN. This patient population represents a diverse range of ages and disease severities for which our results can be generalized.

We found significant negative correlations between FEV₁% predicted and spontaneously expectorated sputum inflammatory markers including free elastase, IL-8, neutrophil counts, and percent neutrophils. We report a correlation estimate of -0.28 for IL-8 in our study, which is comparable but smaller than the correlation estimate of -0.4 reported in a study of 16 patients with CF by Kim and coworkers (8) using expectorated sputum samples and an estimate of -0.72 reported in a study of 19 patients with CF reported by Sagel and coworkers (7) using induced sputum. Similarly, we report correlation estimates of -0.35 for free elastase and -0.29 for neutrophil counts in our study as compared with estimates of -0.75 and -0.57 reported

Figure 5. Power curves for a two-sample comparison of the change in free elastase.

in the study by Sagel and coworkers. However, our study differs greatly from the studies by Kim and coworkers and Sagel and coworkers, which may explain these dissimilarities. Most notably, a greater degree of variability in a multiple center study, as compared with one at a single center, should be expected and may be attributed to either diversity in the patient population or to minor differences in site collection and processing that may not have been standardized in the standard operating procedures used across the sites. Whereas nearly all the patients in the study by Sagel and coworkers had normal or mild obstructive lung disease, our study included patients with various degrees of lung disease. In addition, the study by Sagel and coworkers reported data from induced rather than expectorated sputum and, although other studies have demonstrated comparability between the two (17, 18), this may also be a source of discrepancy between our two findings as induced sputum may result in less variability and stronger correlations.

A significant negative correlation between pulmonary function and *P. aeruginosa* bacterial density was also found in our study, but was of a smaller magnitude than correlations with the inflammatory markers. This may be because quantitative sputum culture is a bioassay with inherent biological variability, whereas cytokines such as IL-8 are biochemical assays and cell counts are microscopic assays. *S. aureus* density was positively associated with FEV₁% predicted, and we note that previous studies have found significant positive associations between *S. aureus* and mortality (19, 20).

There are two novel methodologic aspects to our investigation of the associations between pulmonary function and markers of infection and inflammation. First, the results for IL-8 and free elastase support the use of our flexible approach for estimating correlations in the presence of values below a limit of detection. Importantly, commonly used methods for estimating correlation coefficients such as the Pearson correlation or Spearman rank correlation estimates (15) are undesirable in the presence of nondetectable values, as they do not adequately account for biological mechanisms that may suggest the data do not arise from a single distribution. Our approach therefore allows the possibility that patients with values below a lower limit of detection or nondetectable in a marker may represent a distinct group of patients from those with measurable amounts of the marker. There is evidence from our analysis to suggest that for IL-8 and free elastase, it may not be appropriate in future studies to assume that the group of participants with nondetectable values shares the same relationship with FEV₁ as the group with detectable values.

The second unique aspect of our study is that besides reporting simply the correlation estimates, we also compare the average differences in FEV₁% predicted associated with average differences in each of the markers. This is a critical step for any study exploring associations, because the correlation estimates are limited in their interpretation. Our investigation revealed that there were significant correlations between bacterial densities and pulmonary function in our study (Figure 2). However, among patients with these organisms, differences in FEV₁% predicted associated with differences in the bacterial densities were not clinically meaningful compared with analogous results for the inflammatory markers (Figure 3). Because the presence of *P. aeruginosa* is an established correlate of disease severity (21, 22), our results support the hypothesis that significant correlations between *P. aeruginosa* density and FEV₁% predicted are driven more by the presence or absence of this organism than by the measured density. The ratio between planktonic and biofilm-growing organisms may also account for variation in FEV₁, as evidenced by the association between mucoidy in *P. aeruginosa* and disease severity (21).

Unlike prior studies, our study included sufficient data with which to pursue a multivariate model between pulmonary function and markers of infection and inflammation to identify the set of markers best able to explain the observed variation in FEV₁% predicted. Free elastase and neutrophil counts were able to explain the greatest amount of variation in FEV₁% predicted across patients. We recognize, however, that there was a great deal of correlation between many of the markers, such as elastase and IL-8.

Another unique aspect of our study was our ability to examine longitudinal relationships between pulmonary function and sputum markers. Cross-sectional analyses are most commonly reported in the literature, but longitudinal analyses are essential for the validation of these markers as correlates for disease severity. Our results support the hypothesis that there is a correlation between change in elastase and change in FEV₁ among patients receiving placebo, and this correlation may be stronger and more clinically meaningful among patients receiving a successful intervention. There is a possibility that the specific placebo formulations used in the studies could have altered the natural change in elastase over time, but we were unable to address this hypothesis within the context of this study. The use of elastase as a primary end point may be of most pertinent consideration for clinical trials in which acute reductions in inflammation are expected before increases in or stabilization of clinical efficacy measures such as FEV₁. Our analyses further indicate that the sample sizes required for this sputum marker end point are feasible and comparable to those needed for more traditional study end points.

There are several limitations to this exploratory study. Our primary aim was to estimate associations rather than to perform formal hypothesis testing. Thus, our results should be viewed as hypothesis generating and motivation for future studies. Unfortunately, data availability limited this study to a subset of inflammatory markers that may play a role in the pathophysiology of CF. These did not include, for instance, myeloperoxidase, DNA, and metalloproteinase-9 (8, 23). We did not have sufficient data to adequately look for important interactions of the relationship between FEV₁ and the sputum markers by factors such as age group and sex, and these analyses will be important for a future study to further characterize these relationships. Longitudinal analyses were also limited by the smaller number of patients for whom data were available. Although our cohort represents a much larger group of patients than represented in prior studies, there was admittedly only 4% of patients 12 years of age or younger, and thus these results may not be generalizable to a patient population with more mild disease. Despite these limitations, we were still able to demonstrate significant and meaningful associations between pulmonary function and expectorated sputum biomarkers among a diverse patient population and beyond a single center. Further, it is important to note that there is reason to believe that the true magnitude of associations described in our study is even greater than reported, because measurement error can considerably alter the correlation estimates (24). We chose to focus on expectorated sputum biomarkers given their availability across multiple studies in the CFF TDN Data Bank, but it would be of interest to extend these results to the same biomarkers obtained from induced sputum.

This study represents a critical step in the validation of sputum biomarkers as correlates of disease severity. However, while FEV₁ may be considered a good indicator of clinical status, alone it does not capture all aspects of disease severity. FEV₁ is only one of several known predictors of mortality (19, 20, 25, 26). Therefore strong correlations between FEV₁ and markers of inflammation and infection should not be expected, because these end points represent unique components of the disease

process. In addition, variability in FEV₁ is determined by several factors not necessarily directly related to infection and inflammation, including genetic determinations of lung size and airway caliber, mucus plugging, airway edema, and bronchoconstriction. Strong correlations with FEV₁ should not then be a necessary criterion for sputum biomarkers to be considered correlates of disease severity.

The next defining step in the validation process for sputum biomarkers will be to quantify the level of change in the biomarker needed to produce a change in clinical status. Our study provides proof of concept that there are differences in sputum inflammatory marker levels between patients with different disease severity levels as captured by FEV₁% predicted. There is also the suggestion that changes in markers of inflammation, in particular free elastase, correlate with changes in pulmonary function. Future studies will need to further investigate longitudinal associations between changes in the markers and clinically meaningful changes in clinical efficacy parameters. The best data with which to conduct these studies will be from studies of interventions that are able to elicit a systematic response from both clinical outcomes and sputum biomarkers. Thus, it is critical in the validation of these biomarkers that they be included not only in studies for which biological efficacy is the primary aim, but also as secondary end points in larger studies in which investigating clinical efficacy is the primary aim.

Conflict of Interest Statement: N.M.-H. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. M.L.A. entered into a contractual agreement under the auspices of the University of Washington between the period 1/1/04–present as a principal investigator of a grant sponsored by Sucampo Pharmaceuticals with total award of \$23,393. F.J.A. received \$2,000 for consulting work for Novartis in 2006. R.A.K. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. M.W.K. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. J.L.B. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. S.D.S. received \$2,600 in 2006 for consulting work for Conatus Pharmaceuticals. His site received \$27,000 from 2004 to 2006 for conducting a clinical trial partially sponsored by INO Therapeutics. B.W.R. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript.

Acknowledgment: The authors gratefully acknowledge the following principal investigators of the included studies: Bruce Marshall, M.D., Richard Moss, M.D., Claudia Ordoñez, M.D., and Lisa Saiman, M.D., M.P.H. Statistical review was provided by Mary Emond, Ph.D. Heidi Sucharew, M.S., and Umer Khan, M.S., prepared the data for analysis. The authors acknowledge Bob Beall, Ph.D., and Preston Campbell, M.D., from the Cystic Fibrosis Foundation and Intermune Pharmaceuticals, Pfizer, Inc., and Targeted Genetics for use of data from their sponsored studies. Most importantly, the authors thank the patients with CF for participating in these studies.

References

- Sagel SD, Kapsner R, Osberg I, Sontag MK, Accurso FJ. Airway inflammation in children with cystic fibrosis and healthy children assessed by sputum induction. *Am J Respir Crit Care Med* 2001;164:1425–1431.
- Konstan MW, Hilliard KA, Norvell TM, Berger M. Bronchoalveolar lavage findings in cystic fibrosis patients with stable, clinically mild lung disease suggest ongoing infection and inflammation. *Am J Respir Crit Care Med* 1994;150:448–454.
- Gibson RL, Burns JL, Ramsey BW. Pathophysiology and management of pulmonary infections in cystic fibrosis. *Am J Respir Crit Care Med* 2003;168:918–951.
- Chmiel JF, Berger M, Konstan MW. The role of inflammation in the pathophysiology of CF lung disease. *Clin Rev Allergy Immunol* 2002;1:5–27.
- Sagel SD. Noninvasive biomarkers of airway inflammation in cystic fibrosis. *Curr Opin Pulm Med* 2003;9:516–521.
- Ramsey BW, Boat TF; Consensus Group. Outcome measures for clinical trials in cystic fibrosis: summary of a Cystic Fibrosis Foundation Consensus Conference. *J Pediatr* 1994;124:177–192.
- Sagel SD, Sontag MK, Wagener JS, Kapsner RK, Osberg I, Accurso FJ. Induced sputum inflammatory measures correlate with lung function in children with cystic fibrosis. *J Pediatr* 2002;141:811–817.
- Kim JS, Okamoto K, Rubin BK. Pulmonary function is negatively correlated with sputum inflammatory markers and cough clearability in subjects with cystic fibrosis but not those with chronic bronchitis. *Chest* 2006;129:1148–1154.
- Saiman L, Marshall BC, Mayer-Hamblett N, Burns JL, Quittner AL, Cibene DA, Coquillet S, Fieberg AY, Accurso FJ, Campbell PW III; Macrolide Study Group. Azithromycin in patients with cystic fibrosis chronically infected with *Pseudomonas aeruginosa*: a randomized controlled trial. *JAMA* 2003;290:1749–1756.
- Moss RB, Mayer-Hamblett N, Wagener J, Daines C, Hale K, Ahrens R, Gibson R, Anderson P, Retsch-Bogart G, Nasr S, et al. A randomized, double-blind, placebo-controlled, dose-escalating study of aerosolized interferon γ -1b in patients with mild to moderate cystic fibrosis lung disease. *Pediatr Pulmonol* 2005;39:209–218.
- Moss RB, Rodman D, Spencer TL, Aitken M, Zeitlin P, Waltz D, Milla C, Brody A, Clancy JP, Ramsey B, et al. Repeated AAV2 aerosol-mediated CFTR gene transfer to the lungs of patients with cystic fibrosis: a multi-center, double-blind, placebo-controlled trial. *Chest* 2004;125:509–521.
- Ordoñez CL, Henig NR, Mayer-Hamblett N, Accurso FJ, Burns JL, Chmiel JF, Daines CL, Gibson RL, McNamara S, Retsch-Bogart GZ, et al. Inflammatory and microbiologic markers in induced sputum after intravenous antibiotics in cystic fibrosis. *Am J Respir Crit Care Med* 2003;168:1471–1475.
- Ordoñez CL, Stulberg M, Grundland H, Liu JT, Boushey HA. Effect of clarithromycin on airway obstruction and inflammatory markers in induced sputum in cystic fibrosis: A pilot study. *Pediatr Pulmonol* 2001;32:29–37.
- Knudson RJ, Lebowitz MD, Holberg CJ, Burrows B. Changes in normal maximal expiratory flow-volume curve with growth and aging. *Am Rev Respir Dis* 1983;127:725–734.
- Belle GV, Fisher LD, Heagerty PJ, Lumley TS. Biostatistics: a methodology for the health sciences. London: John Wiley & Sons; 2004.
- Efron B, Tibshirani RJ. An introduction to the bootstrap. Boca Raton, FL: CRC Press; 1998.
- Sagel SD, Kapsner R, Osberg I, Sontag MK, Accurso FJ. Airway inflammation in children with cystic fibrosis and healthy children assessed by sputum induction. *Am J Respir Crit Care Med* 2001;164:1425–1431.
- Henig NR, Tonelli MR, Pier MV, Burns JL, Aitken ML. Sputum induction as a research tool for sampling the airways of subjects with cystic fibrosis. *Thorax* 2001;56:306–311.
- Liou TG, Adler FR, Fitzsimmons SC, Cahill BC, Hibbs JR, Marshall BC. Predictive 5-year survivorship model of cystic fibrosis. *Am J Epidemiol* 2001;153:345–352.
- Belkin RA, Henig NR, Singer LG, Chaparro C, Rubenstein RC, Xie SX, Yee JY, Kotloff RM, Lipson DA, Bunin GR. Risk factors for death of patients with cystic fibrosis awaiting lung transplantation. *Am J Respir Crit Care Med* 2006;173:659–666.
- Demko CA, Byard PF, Davis PB. Gender differences in cystic fibrosis: *Pseudomonas aeruginosa* infection. *J Clin Epidemiol* 1995;48:1041–1049.
- Emerson J, Rosenfeld M, McNamara S, Ramsey B, Gibson RL. *Pseudomonas aeruginosa* and other predictors of mortality and morbidity in young children with cystic fibrosis. *Pediatr Pulmonol* 2002;34:91–100.
- Sagel SD, Kapsner RK, Osberg I. Induced sputum matrix metalloproteinase-9 correlates with lung function and airway inflammation in children with cystic fibrosis. *Pediatr Pulmonol* 2005;39:224–232.
- Yanez ND, Kronmal RA, Shemanski LR. The effects of measurement error in response variables and tests of association of explanatory variables in change models. *Stat Med* 1998;17:2597–2606.
- Mayer-Hamblett N, Rosenfeld M, Emerson J, Goss C, Aitken ML. Developing cystic fibrosis lung transplant referral criteria using predictors of two year mortality. *Am J Respir Crit Care Med* 2002;166:1550–1555.
- Nixon PA, Orenstein DM, Kelsey SF, Doershuk CF. The prognostic value of exercise testing in patients with cystic fibrosis. *N Engl J Med* 1992;327:1785–1788.