Experience using centralized spirometry in the phase 2 randomized, placebo-controlled, double-blind trial of denufosol in patients with mild to moderate cystic fibrosis☆

C.H. Gossa,a,c,⁎, E.F. McKonea, D. Mathews b, D. Kerr c, J.S. Wanger d, S.P. Millarde and the Cystic Fibrosis Therapeutics Development Network

a Department of Medicine, University of Washington, Seattle WA, United States
b Inspire Pharmaceuticals, Inc., NC, United States
c Axio Research Corp., Seattle WA, United States
d Pharmaceutical Research Associates, Inc., Lenexa, KS, United States
e Cystic Fibrosis Therapeutics Development Network, Children's Hospital and Regional Medical Center, Seattle, WA, United States

Received 14 November 2006; received in revised form 14 May 2007; accepted 13 July 2007
Available online 28 August 2007

Abstract

Background: Centralized spirometry may significantly improve quality of spirometry and reduce variability of this outcome measure in clinical trials in cystic fibrosis (CF).
Methods: Spirometry was performed during the phase 2 randomized, placebo-controlled, double-blind clinical trial of denufosol in patients with mild to moderate CF using American Thoracic Society guidelines. Uniform spirometers were used with electronic data transmission of all the data to a reading center. Spirometry was evaluated for quality by a central reader based on start of test, cough during the test, and evidence of a plateau.
Results: A total of 1418 spirometry values were assessed in 89 subjects during the trial. In only 5 instances did the central reading center need to give feedback to sites regarding the quality of spirometry. The study site data matched the central reading center’s data for all but 78 (6%) spirometry values in 33 patients. Many of these differences were small with only 35 (3%) values differing by more than 50 mL in 26 patients.
Conclusion: Spirometry in this clinical trial was of high quality with low rate of significant centralized over-read.

© 2007 European Cystic Fibrosis Society. Published by Elsevier B.V. All rights reserved.

Keywords: Spirometry; Clinical trials; Outcome measures; Cystic fibrosis; Centralized

1. Introduction

Cystic fibrosis (CF) is one of the most common inherited fatal diseases in Caucasians with a reported incidence of CF from 1 in 2000 to 1 in 3200 live births [1]. Over the past 30 years, the median age of survival has improved from 14 years in 1969 to over 35 years in 2004 in the United States [2]. Advancing the care of patients with CF requires the continued search for more efficacious treatments that can modify the natural history of the disease and improve symptom control and prognosis. Improving treatment options for patients with CF requires outcome measures that are clinically relevant and readily responsive to change with treatments. One of the most commonly used outcome measures in CF has been lung function. Lung function is primarily assessed via spirometry, principally using the measure of forced vital capacity (FVC) and forced expiratory volume in the first
second (FEV\textsubscript{1}). Both FEV\textsubscript{1} and FVC continue to be critical outcome measures in both randomized clinical trials and epidemiologic studies related to lung disease in CF; FEV\textsubscript{1} continues to be one of the most important surrogate endpoints to document response to therapy. Improvement in lung function has been used to gain approval from the Food and Drug Administration (FDA) for two important therapies in the treatment of CF [recombinant human deoxyribonuclease (DNase), and inhaled tobramycin] \cite{3,4}.

As with any clinical endpoint used in clinical trials, improving accuracy and precision of spirometry is integral to decreasing both biased results and excessive variability of results. Irwin and colleagues have clearly shown that variability can bias the measurement of lung function due to regression to the mean \cite{5}. High quality spirometry relies on the actual spirometer, the technician performing the study, the patient’s participation during the study, and lastly the patient’s true lung function. Centralized spirometry is one approach to potentially reduce variability of spirometry. Centralized spirometry incorporates standardized instrumentation, site/technician training, centralized review by an expert over-reader, and immediate feedback to sites on poor-tation, site/technician training, centralized review by an expert over-reader, and immediate feedback to sites on poor-quality test sessions. This program is put in place to ensure that spirometry is performed according to American Thoracic Society (ATS) standards, merely stating that these standards are employed in a clinical trial does not mean that they are indeed rigorously applied \cite{6}.

Only limited data exist regarding the impact of centralized spirometry on spirometry performed in clinical trials \cite{7–12}. These studies suggest that centralized spirometry can reduce variability in spirometry in clinical trials in studies involving patients with both asthma and chronic obstructive lung disease. Given these data, we prospectively assessed the impact of using centralized spirometry in a phase 2 multi-center randomized double-blind placebo-controlled trial in patients with CF.

2. Methods

2.1. Participants

Subjects in this study were all participants in a multi-center, double-blind, randomized, placebo-controlled trial (DBPCRCT) comparing the safety and efficacy of three dose levels of inhaled denufosol, a novel P2Y2 receptor agonist, to placebo in subjects with mild to moderate CF lung disease; the goal of this novel agent is to improve lung function in patients with CF. Results from this study are reported by Deterding et al. \cite{13}. This DBPCRCT trial was conducted within the CF Therapeutics Development Network (CF TDN) with six additional sites from outside the network. This clinical trial network is funded by the National Institutes of Health and Cystic Fibrosis Therapeutics, Inc. and currently has 18 major CF clinical sites situated around the United States representing the largest CF centers caring for both pediatric and adult CF patients in the nation. A full description of this clinical trial network has been previously reported \cite{14}.

To be included in this clinical trial, patients had to have a confirmed diagnosis of CF, be 8 years of age or greater and less than or equal to 50 years of age, have an oxygen saturation > 90% on room air, have a negative pregnancy test at the screening visit (all females of childbearing potential), be clinically stable with no evidence of acute upper or lower respiratory tract infection or current pulmonary exacerbation, have an FEV\textsubscript{1} percent of predicted greater than 75%, and be able to reproducibly perform spirometry maneuvers. Reproducibility was defined as repeated FEV\textsubscript{1} within 12% on two separate occasions separated by 1–7 days. Key exclusion criteria included abnormal renal or liver function; HRCT scan at screening with clinically significant findings atypical for moderate CF; changes in physiotherapy technique, bronchodilator, anti-inflammatory, or corticosteroid medications during the week prior to screening; and use of certain concomitant medications that could have affected study measurements, including the use of TOBI® or other inhaled antibiotics within 30 days prior to screening or anti-pseudomonal antibiotics, oral macrolide antibiotics or intravenous (IV) antibiotics within 14 days prior to screening. The data for this analysis were anonymous and are included in the Therapeutics Development Network Data Bank, which is approved by the Institutional Review Board at Children’s Hospital and Regional Medical Center in Seattle, WA.

2.2. Design and procedures

The Vitalograph® Spirotrac Centralized Spirometry System was used for this study. The centralized spirometry program was overseen by Pharmaceutical Research Associates (PRA), Inc., Lenexa, Kansas; PRA acted as the central reading center for the study. The centralized spirometry program included initial site education when the spirometers were delivered to each site, an educational session at the investigator meeting, and feedback regarding the site performance of spirometry during the conduct of the study.

Each site was provided with the same spirometer; a Vitalograph 6800 spirometer attached to a PC running the software “Spirotrac for Over-Read.” The Spirotrac spirometer provides two different reports for a session. The first report shows results only from the three “best” tests (i.e., the tests with the largest values of (FVC+FEV\textsubscript{1})), along with a “best” column that shows the largest value of FEV\textsubscript{1} among the best tests, the largest value of FVC among the best tests, and the value of FEF\textsubscript{25–75} associated with the largest value of (FVC+FEV\textsubscript{1}) among the best tests. The second report presents results for all tests that were performed, both acceptable and unacceptable. Unacceptable tests are noted with error flags that indicate the problem with the test. Spirometry was conducted using ATS Standards \cite{6} using a standard operating procedure developed at the CF Therapeutics Development Network. The sites were instructed to pick the best FEV\textsubscript{1} (i.e., largest FEV\textsubscript{1}) and best FVC (i.e., largest FVC) from all acceptable curves as per the ATS
Standard. FEF_{25-75} values were also recorded electronically and reported on the case report forms.

The actual spirometry data, including the spirometry values (volume–time and flow–volume), were then transmitted electronically to a central reading center for over-reading. The central over-reader assessed the loops for: quality based on start of test; cough in the first second during the test; and evidence of a plateau at the end of the maneuver. The over-reader may or may not have agreed with: 1) what the computer said was acceptable or unacceptable, and 2) what the site research coordinator (RC) said was acceptable. The over-reader included comments if there was disagreement (e.g., over-reader does not agree with the site RC about a particular test the RC said is acceptable).

To evaluate the potential impact of centralized spirometry on variability in lung function assessed during a clinical trial, we compared spirometry data from the denufosol study gathered during the first two clinic visits, which occurred prior to initiation of active drug or placebo, with pooled pre-dose data from two previous studies [15,16] conducted within the CF TDN which used similar sites and inclusion criteria and which did not use centralized spirometry. Spirometry values used in the comparison included only those values from screening and baseline visits prior to initiation of any drug therapy. To ensure that subjects had similar lung function, only those subjects with an FEV_{1%} predicted > 75% were included in the analysis, matching the lung function entry criteria for the denufosol study. The percent predicted values for FEV_{1} and FVC were calculated using the reference equations of Knudson et al. [17].

2.3. Statistical analysis

The results of both the central over-read values and site values were compared. Several results are possible from this comparison: 1) the site chooses the “correct” value, enters it correctly; 2) the site chooses the “correct” value but enters it incorrectly; and 3) the site chooses an “incorrect” value. Spirometry was measured at all clinic visits. Two clinic visits occurred prior to dosing (Visits 1 and 2), followed by four post-treatment visits (Visits 3–6). Because denufosol is an investigational drug, in this analysis we were not interested in modeling efficacy for the post-treatment visits (efficacy results are presented in Deiterding et al. [13]). Instead, we were interested in looking at agreement between spirometry measures reported at a site and reported by the over-reader. We were also interested in comparing variability within subjects prior to dosing in this study versus variability within subjects prior to dosing in previous studies.

Site and central reading center parameter values (FEV_{1} and FVC) were compared by computing differences in the values at each visit. Linear mixed-effects models were used to account for repeated measures within subjects and estimate variance components [18]. Mean absolute difference between center and site values within subject was compared between sites using the Kruskal–Wallis test [19]. Within-subject variability was assessed using spirometry values obtained at screening and baseline (pre-1st dose) spirometry. Absolute difference in lung function from screening to baseline was compared between studies using the two-sample t-test, and compared within the denufosol study for values with and without over-read using the paired t-test. We used SAS 8.2 for Windows (SAS Institute Inc., Cary, NC) and S-PLUS 6.2 and 7.0 (Insightful Corporation, Seattle, WA).

3. Results

A total of 107 subjects were screened for enrollment in the denufosol versus placebo study, 90 (84%) were randomized, with a final sample size of 89 in the intent-to-treat (ITT) population. The reasons subjects failed screening included: FEV_{1} % predicted < 75 (n = 7), not clinically stable (n = 4), and unable or not willing to comply with protocol, be present for required visits, or complete study procedures.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Placebo (N=21)</th>
<th>Denufosol 20 mg (N=23)</th>
<th>Denufosol 40 mg (N=22)</th>
<th>Denufosol 60 mg (N=23)</th>
<th>All groups combined (N=89)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years, mean, (SD)</td>
<td>15.8 (7.78)</td>
<td>14.3 (7.43)</td>
<td>18.0 (9.12)</td>
<td>16.2 (8.92)</td>
<td>16.1 (8.31)</td>
</tr>
<tr>
<td>Male, n (%)</td>
<td>15 (71)</td>
<td>16 (70)</td>
<td>11 (50)</td>
<td>8 (35)</td>
<td>50 (56)</td>
</tr>
<tr>
<td>Caucasian, n (%)</td>
<td>21 (100)</td>
<td>23 (100)</td>
<td>22 (100)</td>
<td>22 (96)</td>
<td>88 (99)</td>
</tr>
<tr>
<td>CF genotype, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔF508 homozygous</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔF508 heterozygous</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Height, cm, mean (SD)</td>
<td>156.00 (16.313)</td>
<td>148.29 (18.968)</td>
<td>157.82 (15.027)</td>
<td>152.48 (14.706)</td>
<td>153.55 (16.490)</td>
</tr>
<tr>
<td>Weight, kg, mean (SD)</td>
<td>59.83 (14.620)</td>
<td>44.61 (18.811)</td>
<td>53.76 (18.216)</td>
<td>49.46 (17.064)</td>
<td>49.59 (17.327)</td>
</tr>
<tr>
<td>FEV_{1} (L)</td>
<td>2.690 (0.7967)</td>
<td>2.389 (0.9094)</td>
<td>2.623 (0.7775)</td>
<td>2.547 (0.8104)</td>
<td>2.559 (0.8198)</td>
</tr>
<tr>
<td>FEV_{1} % predicted, mean (SD)</td>
<td>92.678 (10.6774)</td>
<td>93.936 (9.4483)</td>
<td>88.731 (8.9992)</td>
<td>96.353 (13.7590)</td>
<td>92.977 (11.0738)</td>
</tr>
<tr>
<td>FEF % predicted, mean (SD)</td>
<td>78.924 (27.8006)</td>
<td>84.719 (23.3562)</td>
<td>66.498 (21.8684)</td>
<td>82.039 (24.0664)</td>
<td>78.155 (24.9051)</td>
</tr>
<tr>
<td><em>P. aeruginosa</em> positive, n (%)</td>
<td>6 (29)</td>
<td>11 (48)</td>
<td>11 (50)</td>
<td>4 (17)</td>
<td>32 (36)</td>
</tr>
</tbody>
</table>

* A patient is considered positive for Pa if the patient’s culture at screening was positive, or if the patient could not expectorate sputum and the patient had a history of two known positive bronchoalveolar lavage or sputum cultures within the past two years.
(n=9) (note that some subjects had multiple reasons for failing screening). A full description of the subjects involved in this study is shown in Table 1. This population of subjects has mild lung disease with an average FEV\textsubscript{1} % predicted of 93\% (SD=11). Thirty-nine (44\%) were females, and the mean age was 16.1 years (SD=8.3).

A total of 1418 spirometry values (709 FVC values and 709 FEV\textsubscript{1} values) were obtained in the 89 subjects included in the intent-to-treat population. Despite this very large number of spirometry test sessions, the central over-reading center needed to give feedback to specific sites regarding the quality of the spirometry in only 5 instances. When specific site data was compared to the final data picked by the central over-reader, site data matched the central reading center’s data for all but 78 (5.5\%) of the 1418 spirometry values (40 FEV\textsubscript{1} and 38 FVC values differed). These differences occurred when the over-reader did not agree with the site with regards to the most appropriate spirometry data to evaluate for the best FEV\textsubscript{1} and FVC. The majority of these differences were small. Only 35 (2.5\%) values differed by more than 50 mL for FEV\textsubscript{1} and FVC combined. The mean differences between the central over-read and site values for FEV\textsubscript{1} and FVC were 0.5 mL (SD=20 mL) and −4.2 mL (SD=55 mL) respectively. Given that almost 95\% of the differences were 0, it is not surprising that linear mixed-effects models to account for repeated measures within subject showed that within-subject variability for the differences was much greater than between-subject variability: 2 mL versus 19 mL for between- and within-subject SDs of FEV\textsubscript{1} differences, and 15 mL versus 54 mL for between- and within-subject SDs for FVC differences. Not surprisingly, there were a few “outliers”: the minimum and maximum values of the differences were: −130 mL to +350 mL for FEV\textsubscript{1}; and −840 mL to +320 mL for FVC.

Fig. 1 shows for each subject the mean of the absolute values of the center minus site differences for FEV\textsubscript{1} (mL), plotted by site. There was no clear evidence that the distribution of mean absolute differences differed by site.

Fig. 2. Variability of spirometry as expressed by change in FEV\textsubscript{1} (L) from the screening visit to the baseline visit was reduced compared to historical controls but not statistically significant. The difference between individual subjects’ FEV\textsubscript{1} at screening and at visit 1 prior to dosing was calculated for the denufosol Phase 2 study with and without central over-read, and these differences are compared to differences seen in historical controls. Filled circles represent the mean values and bars represent the 95\% confidence intervals.

3.1. Comparison with historical controls

One of the important questions that we posed was whether the use of centralized spirometry with central over-read resulted in a reduction of the variability of spirometry used in CF clinical trials. The only definitive way to answer this question would be to randomize subjects to having centralized spirometry versus no centralized spirometry and assess the impact of the two programs, but this study design would be extremely challenging to fund. Comparing pre-dose data with pre-dose data from historical controls from CF clinical trials with similar inclusion criteria and lung function may represent the best alternative.

Pre-dose data at two separate visits from the denufosol study was available for 92 subjects: the 89 subjects in the ITT population, as well as two subjects who were dropped from the study because they did not meet the “FEV\textsubscript{1} (L) within 12\% on two separate occasions” inclusion criterion, and one subject who was dropped from the study after Visit 2 due to increased cough. Pre-dose data at two separate visits from two previous studies conducted with the CF TDN was available for 52 subjects. When compared with historical controls with similar lung function, the absolute difference in lung function (FEV\textsubscript{1}) from screening to baseline was lower in the denufosol study (Fig. 2), but not statistically significant (p=0.2 for denufosol with over-read versus historical control; p=0.2 for denufosol without over-read versus historical control; two-sample \textit{t}-test). The mean
absolute difference was larger in the historical control group (124 mL versus 102 mL) with a larger standard deviation of the difference (130 mL versus 77 mL). We also looked at the difference between screening and baseline lung function (FEV₁) with and without the central over-read. There was no evidence that the central over-read in and of itself reduced variability compared merely to centralized spirometry with uniform spirometers at each site (p=0.5, paired t-test). Mean absolute difference in FEV₁ between screening and baseline was 103 mL without over-read versus 102 mL with over-read, and the standard deviation of the absolute difference was 79 mL without over-read versus 77 mL with over-read. Similar results were obtained for FVC.

Table 2 shows variance components estimates for raw FEV₁ (L) values using just the screening and baseline pre-dose data based on linear mixed-effects models. Although the within-subject variability as a percentage of total variability is slightly greater for the historical controls (1.8% versus 1.2%), the confidence intervals for these variance components makes it clear that there is no firm evidence that within-subject variability has been reduced in the denufosol study compared to the historical controls, with or without over-reading. The fact that within-subject variability as a percentage of total variability is less than 2% for the historical controls is a testament to the high-level of quality control in these historical studies.

One concern with this historical comparison is that the duration of time in days from the screening visit to the baseline visit would be related either linearly or non-linearly to the variation in lung function. Thus, if the historical control group had a much greater period of time pass between screening and baseline, this alone could explain the reduction in variation observed in our study. When we examined the relationship between the number of days between screening and baseline spirometry values, we found that in fact a much higher proportion of visits for the historical controls were two weeks apart compared to the denufosol study, for which most visits were at most one week apart (Fig. 3). Variability for the historical controls increased with days between visits (slope=0.006, 95% CI=[−0.011, 0.012]), while variability for the denufosol study decreased with days between visits (slope=−0.008, 95% CI=[−0.016, 0.001]), but neither slope was significantly different from 0. Thus, we found no clear effect of days between visits and variability.

**4. Discussion**

Site spirometry in the denufosol versus placebo CF clinical trial was of high quality with low rate of over-read from the central reading center. In comparison to other CF clinical trials conducted within the CF TDN, a program of centralized reading reduced within-subject variability that did not reach statistical significance. The central over-read appeared to play a minimal role in reducing variability. In the current study, high quality of spirometry diminished the role of centralized spirometry. Overall, decreasing variability in lung function measurements will be necessary if we are to be able to move forward with more therapies in rare lung diseases like cystic fibrosis. As clinical research in CF expands to less experienced research sites, the quality and variability of spirometry may worsen.

Most of the evidence regarding the variability of spirometry and the effect of quality assurance programs (like centralized spirometry) come from large epidemiologic and randomized clinical trials [8,9,11,12]. The effect of these programs has been assessed primarily using before/after designs and by comparing their results to those of other large cohorts, much like our design. The primary evidence to support the use of quality assurance programs has come form the Multiple Risk Factor Intervention Trial (MRFIT) [7], Lung Health Study (LHS) [8–10], the National Health and Nutrition Examination Survey (NHANES) III study [20,21], the Children’s Health Study [11], and most recently asthma clinical trials [12].

The earliest report of quality oversight of spirometry using a 10-liter Stead–Wells water-filled spirometer in a
major clinical trial came from the MRFIT study [7]. The conclusion from this study was that centralized training on testing techniques and spirometer maintenance along with periodic retraining during long trials is strongly recommended. In the LHS, 73,000 current cigarette smokers ages 45 to 60 were screened at 10 sites on three occasions prior to enrollment in a prospective randomized clinical trial to assess the effect of smoking cessation and an inhaled bronchodilator on annual lung function decline over 5 years [8]. Specialized software with real-time displays helped coach both the subject and the technician if tests were inadequate, giving precise instructions to perform an adequate test. Eighteen months after study initiation, central reading was started because of excessive variability noted early in the study. The coefficient of variation found in this study was 5.8% after repeating the spirometry at two time points separated by 30–90 days. This was much less than that noted in earlier cohort studies from a single time point (range 9.9% to 20.2% in smokers) [22–24]. Sustained improved quality scores of spirometry occurred only after central reading/feedback and scoring was initiated. Three other studies have reported results and experience with centralized spirometry [11,12,20,21]. These studies noted that centralized spirometry was associated with minimal rates of unacceptable spirometry loops [20,21], low rates spirometry loops that failed to meet ATS standards [11], and low rates of failing spirometry using a spirometry grading system (A to F with the best studies achieving a grade of A) [12]. Of all of these reported studies, only the LHS convincingly showed the impact on centralized spirometry.

Experience from the LHS, NHANES, the Child’s Health Study, and asthma clinical trials does not necessarily apply to the CF clinical research setting. The study sites in these examples used both community-based and academic sites that may or may not have had prior experience conducting clinical research in lung disease. In comparison, the majority of CF care is delivered at specialized centers with a multi-disciplinary approach, with a significant percentage of the centers located at academic medical centers. The centers in our study were all tertiary referral CF centers with experience doing clinical research in CF with established standard operating procedures for the conduct of spirometry. Thus, it was unclear whether we could show any benefit of centralized spirometry in this specialized setting.

The reason for the reduced variability in CF clinical trials is likely multi-factorial. Certainly, intrinsic subject characteristics may be important in explaining variability in spirometry, but these factors cannot be routinely controlled for in a clinical trial other than excluding specific subject populations. Reasons that our sample showed reduced though not statistically significant variability compared to historic controls include: different subject populations, protocol exclusions, site personnel who are very familiar with the spirometry procedure, site training, the use of uniform spirometer, and experience with spirometry in clinical trial. Additional information from the literature has noted that subject experience with spirometry in and of itself is an important predictor of non-intrinsic variability in spirometry [12,25–28]. Given that the historical data came from these same sites within the prior 3 years and included similar subject populations based on FEV1, we believe most of the reduction in variability we noted was due to the use of uniform spirometers and detailed site training at the start of the study. Our study was also of short duration in comparison to some of those mentioned above. Our sites had also already gone through two internal reviews of spirometry (169 and 160 spirograms respectively) conducted by Dr. Paul Enright. In the first review, only 1.2% of 169 spirometry studies did not fulfill ATS standards — this all prior to the introduction of centralized spirometry. In our current study, we found no clear evidence that the act of central reading had much impact on the variability of spirometry conducted in this study. Clearly, quality assurance programs in spirometry utilized in clinical trials are important. It remains to be proven which component of a program that includes centralized spirometry is critical to reducing variability in spirometry values. Our data may not apply to CF subjects with FEV1 below 75% of predicted; further assessment of centralized spirometry is needed in this population of CF patients.

A component of test assessment not formally addressed in this analysis is accuracy; an ideal test has both high precision (low variability) and high accuracy (low bias). We focused on variability because accuracy cannot be formally assessed in spirometry without the availability of a gold standard. We did not have data regarding the actual experience of the technicians performing the spirometry nor the volume at each given site. Thus, issues related to the size of a site are much more complex and relate to training and experience.

In conclusion, we believe that centralized spirometry (instrument standardization; site/technician training, centralized review of data by an expert over-reader, and immediate feedback to sites on poor quality sessions) has the potential to decrease test variability. However, we found that for CF clinical trials within the established CF TDN, having central over-read did not significantly enhance spirometry quality. We did see a non-statistically significant reduction in variability with centralized spirometry which may justify it use in further clinical trials. Careful training of the sites, uniform spirometers, centralized transfer of data may reduce variability and improve our capacity to assess change due to treatments particularly as less experienced research sites are used in the setting of larger and more numerous CF clinical trials.

Acknowledgments

We are indebted to Robert Beall, PhD, President, CF Foundation, Preston Campbell, III, MD, Vice President for Medical Affairs, CF Foundation, Investigators and Research staff at each of the clinical sites of the CF Therapeutics Development Network, and the CF subjects who participated in these studies. We are also indebted to Inspire Pharmacau-
ticals for access to the data for this analysis. The following is a list of all the participating sites and principal investigators:

<table>
<thead>
<tr>
<th>City</th>
<th>Lead Study PI</th>
<th>Principal Investigators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denver (Lead Study PI)</td>
<td>Robin Deterding, MD</td>
<td>George Retsch-Bogart, MD</td>
</tr>
<tr>
<td>Columbus</td>
<td>Karen McCoy, MD</td>
<td>Ronald Gibson, MD, PhD</td>
</tr>
<tr>
<td>Baltimore</td>
<td>Pamela Zeitlin, MD, PhD</td>
<td>Boston Children's, David Waltz, MD</td>
</tr>
<tr>
<td>Houston</td>
<td>Peter Hiatt, MD</td>
<td>Robert Wilmott, MD</td>
</tr>
<tr>
<td>Boston Mass Gen</td>
<td>Hank Dorkin, MD</td>
<td>Cori Daines, MD</td>
</tr>
<tr>
<td>Palo Alto</td>
<td>Noreen Henig, MD</td>
<td>Carlos Milla, MD</td>
</tr>
<tr>
<td>Chicago</td>
<td>Susanna McColley, MD</td>
<td>Christopher Harris, MD</td>
</tr>
</tbody>
</table>

References


