Analyses of Efficacy End Points in a Controlled Trial of Interferon-γ1b for Idiopathic Pulmonary Fibrosis*

Talmadge E. King, Jr, MD, FCCP; Sharon Safrin, MD; Karen M. Starko, MD; Kevin K. Brown, MD, FCCP; Paul W. Noble, MD; Ganesh Raghu, MD, FCCP; and David A. Schwartz, MD, FCCP

Background: Idiopathic pulmonary fibrosis (IPF) is a devastating disease, yet validated, reliable criteria for evaluating patient response to therapies in clinical trials are lacking.
Methods: To optimize selection of end point criteria for the study of interferon (IFN)-γ1b in patients with IPF, we retrospectively analyzed the components of the primary efficacy end point used in a large, controlled study of 330 patients for reliability, validity, and sensitivity to treatment effect. The primary end point components were death, disease progression defined as a ≥ 5 mm Hg increase in resting alveolar-arterial oxygen pressure gradient ($P[A\text{-}a]O_2$), and disease progression defined as a $\geq 10\%$ decrease in percentage of predicted FVC.
Results: We found that the $P(A\text{-}a)O_2$ criterion was not reliable and was not associated with mortality. In contrast, the FVC criterion was reliable and was associated with a 2.4-fold increase in the risk of death. Of the three measures, only mortality was sensitive to a treatment effect of IFN-γ1b. Additionally, the tendency for mortality benefit was observed in nearly all patient subgroups defined by baseline physiology. The effect of IFN-γ1b on mortality was strongest in patients with baseline percentage of predicted FVC $\geq 55\%$ (p = 0.004) or percentage of predicted diffusing capacity of the lung for carbon monoxide $\geq 30\%$ (p = 0.008).
Conclusion: We conclude that mortality is the most inclusive end point for future trials of IFN-γ1b in patients with IPF, and that a $> 10\%$ decrement in the percentage of predicted FVC represents a valid measure of disease progression.   *(CHEST 2005; 127:171–177)*

Key words: arterial-alveolar pressure gradient; efficacy end points; idiopathic pulmonary fibrosis; mortality; percentage of predicted FVC

Abbreviations: DLCO = diffusing capacity of the lung for carbon monoxide; IFN = interferon; IPF = idiopathic pulmonary fibrosis; $P(A\text{-}a)O_2$ = arterial-alveolar oxygen pressure gradient

Idiopathic pulmonary fibrosis (IPF) is a discrete clinical and histopathologic entity with a uniformly poor prognosis; however, the pace of the disease may vary substantially in individuals. Although a number of therapeutic agents are in development for treatment of IPF, reliable and objective parameters to assess disease progression and re-sponse to therapy remain poorly defined and not well validated.[1] Consequently, standard parameters to define disease progression or distinguish between responders and nonresponders are lacking. The joint consensus statement on IPF by the American Thoracic Society and the European Respiratory Society identified thresholds for change in several lung function parameters as representative of a favorable

or improved response to therapy in patients with IPF, including total lung capacity, vital capacity, single-breath diffusing capacity of the lung for carbon monoxide (Dlco), arterial oxygen pressure, alveolar-arterial oxygen pressure gradient (P[A-a]O$_2$), and oxygen saturation.[1] However, inadequate data exist to determine if these parameters or thresholds are the most reliable and valid measurements of therapeutic efficacy in patients with IPF.

In 2000, we embarked on a 330-patient phase III clinical trial to assess the impact of interferon (IFN)-γ1b on the progression of IPF. Since no therapy had been shown to be effective for this disease, no model for the assessment of therapeutic efficacy existed. We selected a primary efficacy end point that we hoped had the following characteristics: reliability, validity as a measure of clinical relevance, and sensitivity to the treatment under study. We assumed that P(A-a)O$_2$ and FVC would be sensitive to the effects of IFN-γ1b based on the findings of a small controlled study[2] indicating that patients with IPF treated with IFN-γ1b for 1 year demonstrated significant improvement in total lung capacity and arterial oxygen pressure at rest. Additionally, we estimated the expected response rates in both the treated and the placebo group to plan for the study sample size and power. The results of the trial have been previously published.[3] We now undertake a retrospective analysis of the components in our selected primary efficacy end point (*ie*, change in P[A-a]O$_2$, change in FVC, and death) to gain insights that will inform decisions regarding end point selection and facilitate the design of subsequent clinical trials of IFN-γ1b in patients with IPF.

Materials and Methods

Patients and Study Design

We used data from a recently reported randomized study comparing subcutaneous IFN-γ1b (200 μg; n = 162) with placebo (n = 168), administered three times weekly, in 330 patients who met the diagnostic criteria for IPF according to the American Thoracic Society.[1,3] Eligible patients were aged 20 to 79 years, had mild-to-moderate IPF (*eg*, FVC of 50 to 90% of predicted and Dlco ≥ 25% of predicted), had definite or probable IPF on high-resolution CT scan based on protocol-specified criteria, and had worsening IPF during the preceding year despite therapy with corticosteroids. FVC, Dlco, and resting arterial blood gases at ambient temperature were measured at 3-month intervals. The median duration of follow-up was 58 weeks (range, 2 to 92 weeks).

Statistical Analysis

The primary efficacy end point was progression-free survival time, defined as time from randomization to the first occurrence of either death or disease progression. Disease progression was defined as either an increase of at least 5 mm Hg in P(A-a)O$_2$ or a decrease of at least 10% in percentage of predicted FVC compared to baseline. Threshold changes in P(A-a)O$_2$ and percentage of predicted FVC required confirmation at a subsequent visit within 4 to 14 weeks. Vital status was ascertained in all randomized patients at the completion of the trial.

We conducted a series of analyses that examined the rates of primary end point events in all randomized patients. We first examined the association of single components or combinations of components with the rate of end point events. We next assessed the impact of different thresholds in the components on the rate of primary end point events by progressively increasing the level of change from baseline (P[A-a]O$_2$ in 5-mm increments, percentage of predicted FVC in 5% increments) while holding the other measurements constant or by testing the parameter alone. These analyses assessed only placebo patients to avoid obfuscation by a treatment effect. We assessed end point reliability (*ie*, reproducible and stable [not subject to fluctuation over a brief period of time]) by examining serial measurements of P(A-a)O$_2$ and percentage of predicted FVC between the screening and baseline visits (*ie*, prior to therapeutic intervention). The relationships between change in P(A-a)O$_2$, change in percentage of predicted FVC, and change in percentage of predicted Dlco and death were examined. The risk of death according to different thresholds of change in P(A-a)O$_2$, percentage of predicted FVC, and percentage of predicted Dlco was calculated as a ratio relative to the reference placebo group (*ie*, no change or improvement) in patients receiving placebo. Finally, the sensitivity of treatment effect was evaluated by comparing disease progression and mortality rates according to treatment group in various subgroups defined by baseline physiologic parameters.

Results

Components of the Primary Efficacy End Point

A primary end point event (*ie*, either disease progression according to change in P(A-a)O$_2$ or percentage of predicted FVC criteria, or death; see "Methods and Methods") occurred in 75 IFN-γ1b patients (46.3%) and 87 placebo patients (51.8%) [p = 0.53, likelihood score test from the Cox proportional hazards model; Table 1]. The majority of study patients reaching the primary end point did so on the basis of disease progression rather than death (IFN-γ1b, 90.7%; placebo, 86.2%), and the majority of these events were due to a protocol-specified increase in P(A-a)O$_2$ (43 of 75 events in IFN-γ1b patients and 46 of 87 events in placebo patients). The frequency of FVC-dependent end points was considerably lower in both treatment groups: 8 of 75 events in the IFN-γ1b group and 12 of 87 events in the placebo group. Concurrent changes in both P(A-a)O$_2$ and percentage of predicted FVC occurred in 17 subjects in each treatment group, while death prior to documentation of disease progression occurred in 7 IFN-γ1b patients and 12 placebo patients.

Exploration of the Components of the Primary End Point

Effects of Varying the Threshold Levels of Physiologic Parameters: When the definition of the primary end point included only the P(A-a)O$_2$ criterion and death, 78 placebo patients met the end point, a 10% reduction compared with the number of patients meeting the end point using the original protocol definition (Table 2). In contrast, when the definition of the end point included only the FVC-based criterion and death, 48 placebo patients met the end point, a 45% reduction compared with the original end point definition. Thus, the proportion of patients with the protocol-specified change in P(A-a)O$_2$ was substantially higher than those with the prespecified change in percentage of predicted FVC.

As the definition for disease progression was varied by progressively increasing the threshold for change in P(A-a)O$_2$ from baseline in 5-mm increments (while holding other measures constant), the number of end point events decreased substantially in the placebo group initially (87 events at the ≥ 5 mm Hg threshold vs 65 events at ≥ 10 mm Hg increase) but remained relatively constant at thresholds ≥ 15 mm Hg (Table 2). In contrast, incremental increases of 5% in the threshold for change in percentage of predicted FVC resulted in only a small reduction in the number of events initially (87 events at ≥ 10% decrease vs 79 events at ≥ 15%) and had minimal effect at higher threshold levels (*ie*, 78 events at thresholds of ≥ 20%, ≥ 25%, and ≥ 30% decrease).

When P(A-a)O$_2$ and percentage of predicted FVC were analyzed as sole measures of the efficacy end point, we found that 37.5% of placebo patients experienced an increase of ≥ 5 mm Hg in P(A-a)O$_2$ during the study, with progressively fewer patients manifesting greater changes in P(A-a)O$_2$ (Table 2).

For example, only 7.1% of placebo patients had a ≥ 15 mm Hg increase in P(A-a)O$_2$. In contrast, 17.3% of placebo patients had a decrease of ≥ 10% in percentage of predicted FVC during the study, and < 1% had decreases of ≥ 20%.

Reliability of P(A-a)O$_2$ and Percentage of Predicted FVC: Changes in P(A-a)O$_2$ and percentage of

predicted FVC between the screening visit and the baseline visit were assessed, with a median duration of this interval of 20 to 21 days (range, 3 to 57 days). Since repeating these tests was not required by the protocol, only a subset of patients was available for analysis. We found that 31 of 73 subjects (42.5%) tested had a change of ≥ 5 mm Hg in $P(A\text{-}a)O_2$ between these two pretreatment time points: 18 subjects (24.7%) had an increase in $P(A\text{-}a)O_2 \geq 5$ mm Hg, while 13 subjects (17.8%) had a decrease ≥ 5 mm Hg (Table 3). In contrast, none of 81 tested patients had a $\geq 10\%$ decrease in percentage of predicted FVC, and only 1 patient (1.2%) had a $\geq 10\%$ increase in percentage of predicted FVC.

Effect of Mortality on Nonmortality End Points:
Mortality tends to result in an underestimation of the changes in our physiologic measures of disease progression. Overall, 12 of the 168 patients (7%) receiving placebo died prior to one of the specified progression end points, while 75 patients (45%) demonstrated the specified end point changes in either gas exchange or FVC. If a 10% change in percentage of predicted FVC had been the sole progression end point, 19 patients (11%) would have died prior to reaching that end point, whereas 29 patients (17%) would have reached the progression end point. Thus, mortality clearly decreases the number of study subjects at risk for a nonmortality end point, and probably leads to an underestimate of the number of patients who would have experienced a loss of gas exchange or lung volume.

Relationship Between Change in Physiologic Parameters and Mortality: The relationships between the greatest change in $P(A\text{-}a)O_2$ and mortality, and between the percentage of predicted FVC and mortality were assessed in all placebo patients (Table 4). Change in $P(A\text{-}a)O_2$ was not associated with an increased risk of death in patients at increases of 1 to 14 mm Hg, but mortality increased 2.4-fold in those who had increases of ≥ 15 mm Hg. In contrast, the protocol-defined threshold of $\geq 10\%$ decrease in

Table 3—*Change From the Screening Visit to the Baseline Visit in $P(A\text{-}a)O_2$ and Percentage of Predicted FVC**

Variables	IFN-γ1b	Placebo
Change in $P(A\text{-}a)O_2$, mm Hg	n = 36	n = 37
≥ 5 increase	11 (30.6)	7 (18.9)
≥ 5 decrease	9 (25.0)	4 (10.8)
Change in % predicted FVC, %	n = 41	n = 40
≥ 10 decrease	0 (0)	0 (0)
≥ 10 increase	0 (0)	1 (2.5)

*Data are presented as No. (%).

Table 4—*Mortality According to Greatest Change in Physiologic Parameters During the Study Period in Patients Receiving Placebo**

Variables	Total No. of Deaths/Total No. of Patients (%)	Relative Risk of Death
Change from baseline in $P(A\text{-}a)O_2$, mm Hg		
No change or improvement†	3/22 (14)	1.0
1–4 increase	4/38 (11)	0.8
5–9 increase	0/27 (0)	NA
10–14 increase	5/37 (14)	1.0
≥ 15 increase	13/39 (33)	2.4
Missing‡	3/5 (60)	4.3
Change from baseline in % predicted FVC, %		
No change or improvement†	3/24 (13)	1.0
1–4 decrease	1/41 (2)	0.2
5–9 decrease	6/49 (12)	0.9
≥ 10 decrease	15/49 (31)	2.4
Missing‡	3/5 (60)	4.6
Change from baseline in % predicted DLCO, %		
No change or improvement†	5/26 (19)	1.0
1–4 decrease	2/46 (4)	0.2
5–9 decrease	9/44 (20)	1.1
10–14 decrease	3/23 (13)	0.7
≥ 15 decrease	5/23 (22)	1.2
Missing‡	4/6 (67)	3.5

*NA = not applicable.
†Reference group.
‡Second measurement was not available for comparison to baseline in these patients.

percentage of predicted FVC was associated with a 2.4-fold increase in the relative risk of death.

We also examined the change from baseline in percentage of predicted DLCO. There was no obvious threshold that was associated with a substantially increased risk of death (Table 4).

Analyses To Assess the Impact of IFN-γ1b Treatment on Disease Progression and Mortality

Although analysis of the protocol-specified definition for the primary efficacy end point did not appear to reflect a treatment effect of IFN-γ1b (p = 0.53), a smaller proportion of IFN-γ1b patients reached the end point than did placebo patients in every analysis that used any combination of components of the end point definition (Table 2). Similarly, patients receiving IFN-γ1b had lower frequencies of every outcome in which the threshold for change in $P(A\text{-}a)O_2$ or percentage of predicted FVC was varied, either in combination or as isolated parameters (Table 2). An intent-to-treat analysis identified a trend toward enhanced survival in patients receiving IFN-γ1b, with death in 16 IFN-γ1b patients (9.9%) vs 28 placebo patients (16.7%) [p = 0.08, log-rank test]. In

every subcategory of baseline $P(A-a)O_2$, percentage of predicted FVC, and percentage of predicted DLCO composed of five or more deaths, the proportion of patients dying was similar to or lower in the IFN-γ1b group than in the placebo group (Table 5). Similarly, an analysis in which dichotomized subgroups of these three baseline variables were assessed in association with mortality showed a lower

Variables	IFN-γ1b	Placebo
Baseline $P(A-a)O_2$, mm Hg		
Categorical subgroups		
< 10	1/16 (6.3)	1/21 (4.8)
10–19	2/31 (6.5)	1/32 (3.1)
20–29	5/56 (8.9)	9/59 (15.3)
30–39	3/43 (7.0)	10/42 (23.8)
40–49	4/13 (30.8)	7/14 (50.0)
50–59	1/2 (50)	0/0
Dichotomous subgroups		
< 10	1/16 (6.3)	1/21 (4.8)
< 20	3/47 (6.4)	2/53 (3.8)
< 30	8/103 (7.8)	11/112 (9.8)
< 40	11/146 (7.5)	21/154 (13.6)
< 50	15/159 (9.4)	28/168 (16.7)
< 60	16/161 (9.9)†	28/168 (16.7)
Baseline % predicted FVC		
Categorical subgroups		
≥ 90	0/1 (0)	1/3 (33.3)
80–89	0/14 (0)	2/16 (12.5)
70–79	3/31 (9.7)	6/33 (18.2)
60–69	1/52 (1.9)	5/44 (11.4)
50–59	12/59 (20.3)	12/65 (18.5)
40–49	0/5 (0)	2/7 (28.6)
Dichotomous subgroups		
≥ 70	3/46 (6.5)	9/52 (17.3)
≥ 65	3/67 (4.5)	10/77 (13.0)
≥ 60	4/98 (4.1)	14/96 (14.6)
≥ 55	6/126 (4.8)	21/128 (16.4)
≥ 50	16/157 (10.2)	26/161 (16.1)
≥ 45	16/160 (10.0)	28/168 (16.7)
≥ 40	16/162 (9.9)	28/168 (16.7)
Baseline % predicted DLCO		
Categorical subgroups		
≥ 50	0/15 (0)	1/23 (4.3)
40–49	3/42 (7.1)	3/31 (9.7)
30–39	1/60 (1.7)	12/67 (17.9)
20–29	11/44 (25.0)	11/44 (25.0)
10–20	1/1 (100)	1/3 (33.3)
Dichotomous subgroups		
≥ 50	0/15 (0)	1/23 (4.3)
≥ 45	2/33 (6.1)	1/37 (2.7)
≥ 40	3/57 (5.3)	4/54 (7.4)
≥ 35	4/87 (4.6)	11/85 (12.9)
≥ 30	4/117 (3.4)	16/121 (13.2)
≥ 25	12/152 (7.9)	25/158 (15.8)
≥ 20	15/161 (9.3)	27/165 (16.4)
≥ 15	16/162 (9.9)	28/168 (16.7)

*Data are presented as total No. of deaths/total No. of patients (%).
†Measurement of $P(A-a)O_2$ at baseline was not performed in one study patient.

proportion of deaths in the IFN-γ1b group than in the placebo group in every analysis composed of five or more events.

The mortality end point was more sensitive to a treatment effect of IFN-γ1b than physiologic markers of disease progression. In the mortality analysis, there were two subgroups of patients in which the evidence for a treatment effect was strongest: those with baseline percentage of predicted FVC ≥ 55% (death in 4.8% IFN-γ1b patients vs 16.4% placebo patients; 71% relative reduction in the risk of death; p = 0.004, log-rank test) and those with baseline percentage of predicted DLCO ≥ 30% (death in 3.4% IFN-γ1b patients vs 13.2% placebo patients; 74% relative reduction in the risk of death; p = 0.008). Of note is that each of these two subgroups included a majority of study patients (254 patients [77.0%] had baseline percentage of predicted FVC ≥ 55%, and 238 patients [72.1%] had baseline percentage of predicted DLCO ≥ 30%).

DISCUSSION

Using data from a parallel-group, randomized, double-blind study comparing treatment with IFN-γ1b vs placebo in patients with IPF,[3] we sought to explore the components of the protocol-specified primary efficacy end point to guide selection of end points for future clinical trials of IFN-γ 1b.

Percentage of Predicted FVC Is a More Reliable Parameter Than $P(A-a)O_2$

The primary efficacy end point, based on the best available information at the time of study design, was defined as the time to first occurrence of either disease progression (ie, ≥ 5-mm Hg increase in $P(A-a)O_2$ or ≥ 10% decrease in percentage of predicted FVC) or death. Our data, derived by analyzing several different thresholds of change in $P(A-a)O_2$ and in percentage of predicted FVC in relation to disease progression and mortality, suggest that the protocol-specified threshold for $P(A-a)O_2$ was unreliable for use as a primary end point, because 42.5% of tested patients had at least a 5-mm Hg change in either direction in the absence of intervention, and during the brief interval of time between the screening and baseline visits, and 24.7% experienced disease progression based on this criterion during this interval. Given that this criterion alone accounted for 62% of disease progression events overall during the trial, this is an important finding. In contrast, only 1% of patients had an absolute change of at least 10% in percentage of predicted FVC during the interval between the screening and baseline visits, and no patient met the protocol-specified criterion

for worsening in percentage of predicted FVC (*ie*, $\geq 10\%$) during this interval. Thus, we conclude that the threshold of $\geq 10\%$ increase for percentage of predicted FVC constitutes a more reliable parameter than a ≥ 5 mm Hg increase in P(A-a)O$_2$.

Change in Percentage of Predicted FVC, But Not P(A-a)O$_2$, Is Predictive of Death

We next examined changes in P(A-a)O$_2$ and percentage of predicted FVC in relation to mortality, with mortality representing the most compelling of all end points in a progressive and fatal illness such as IPF. In patients assigned to receive placebo, we found that change in P(A-a)O$_2$ was not associated with an increased risk of death until an increase of ≥ 15 mm Hg occurred, whereas a $\geq 10\%$ decrease in percentage of predicted FVC carried a 2.4-fold risk of increased mortality. An analysis of change in percentage of predicted DLCO failed to identify a clinically useful threshold that predicted mortality. Therefore, we conclude that the chosen threshold for percentage of predicted FVC is a more clinically relevant parameter than that for P(A-a)O$_2$, and this threshold is recommended for use in future trials of patients with IPF if a surrogate measure for mortality is required.

Mortality Is the Most Sensitive Measure of Treatment Effect

Given the trend toward enhanced overall survival in patients receiving IFN-γ1b in this study (p = 0.08), we explored the impact of treatment on mortality in relation to baseline physiologic characteristics. The proportion of patients dying was similar or lower in every category of baseline P(A-a)O$_2$, percentage of predicted FVC, and percentage of predicted DLCO that was composed of five or more deaths. However, a beneficial effect of IFN-γ1b on mortality was strongest in patients with percentage predicted FVC $\geq 55\%$ (p = 0.004) and percentage of predicted DLCO $\geq 30\%$ (p = 0.008).

It remains intriguing that although differences in primary efficacy end point rates did not show a statistically significant difference between treatment groups (p = 0.53), all analyses of varied thresholds for this composite end point showed a lesser proportion of IFN-γ1b than placebo patients meeting the definition for disease progression. Several potential but untested explanations may address this discrepancy. First, loss of subjects to death during the trial may have obscured the effect of IFN-γ1b on physiologic measures of disease progression, particularly if patients with more severe disease died before confirmation of physiologic worsening. It is of note that of the 44 patients who died during the study, 19

patients (43%) did so before documentation of disease progression according to the protocol-specified changes in physiologic parameters; conceivably this was due to the relative infrequency of protocol-specified assessments (*ie*, every 3 months). Thus, mortality decreases the number of study subjects at risk for a nonmortality end point and appears to lead to an underestimate of the number of patients that would have experienced a loss of gas exchange or lung volume. Second, it is possible that the conventional physiologic markers of disease progression selected for our study, particularly P(A-a)O$_2$, are not optimal for monitoring the course of IPF. Third, it may not be feasible to standardize the performance of testing of physiologic markers such as P(A-a)O$_2$ and DLCO in multicenter clinical trials. In particular, given the pathophysiology of the disease, P(A-a)O$_2$ may vary considerably in this population. Thus, variations in barometric pressure, effective alveolar ventilation, and random measurement error may all contribute to the lack of reliability noted in the P(A-a)O$_2$ in this study.[4] Fourth, IFN-γ1b may act on biological pathways other than fibrosis in patients with IPF. For instance, IFN-γ1b may decrease the frequency of episodic acute respiratory decompensation without modulating the underlying fibroproliferative process. Finally, the trial may simply have been underpowered to detect a reduction in disease progression. For example, using the optimized thresholds identified for definition of disease progression (*ie*, ≥ 15-mm Hg change in P(A-a)O$_2$ or $\geq 10\%$ decrease in percentage of predicted FVC) in combination with death as a primary efficacy end point, detection of a difference as small as that observed in our trial (*ie*, from 30% in the placebo group to 24% in the IFN-γ1b group) would require a total of 1,714 patients, using a two-tailed α of 0.05% and 80% power.

This trial was designed to monitor physiologic parameters that would serve as surrogates for mortality, as the duration of 1 year was assumed to be insufficient to detect a difference in survival between the two treatment groups. Our findings regarding the relationship between P(A-a)O$_2$ and percentage of predicted FVC and mortality generally mirror, and corroborate those of recent investigations into predictors of survival in patients with IPF. Hanson and colleagues[5] did not find P(A-a)O$_2$ to be a statistically significant predictor of mortality, while the studies by Flaherty and colleagues[6] and Latsi and colleagues[7] did not monitor P(A-a)O$_2$ as a predictive parameter. Collard and colleagues[8] did find that an increase of ≥ 5 mm Hg from baseline at 6 months was predictive of survival (p = 0.001) in a sample of 68 patients. In contrast, many authors have identified change in percentage of predicted FVC from base-

line as a significant predictor of mortality.[6–8] Specifically, both Collard and colleagues[8] and Flaherty and colleagues[6] identified a decrease of $\geq 10\%$ in percentage of predicted FVC at 6 months as predictive of mortality in recent reports.

In conclusion, we have explored the components of the primary efficacy end point selected for a large, prospective, multicenter randomized trial evaluating IFN-γ1b in patients with IPF. Although IPF is acknowledged to be a progressive interstitial pneumonia without proven effective medical therapy,[1,9] there are no validated and standardized measurements that represent either improvement or progression of disease in individual patients. We found that the occurrence of changes ≥ 5 mm Hg in P(A-a)O$_2$ was not reliable over short periods of time and was not predictive of mortality, whereas a $\geq 10\%$ decrease in percentage of predicted FVC was both reliable and predictive of mortality. While the limitations of subgroup analyses have been well documented,[10] such analyses may offer insights into the effects of an intervention that may prove to be important for enhanced understanding of the natural history of IPF as well as the design of future therapeutic clinical trials of IFN-γ1b in patients with IPF. These data may prove useful in the design of other potential therapeutic agents for patients with IPF as well. Moreover, the suggestion of benefit of IFN-γ1b on both disease progression and mortality in almost all subgroups of lung function, while not reaching statistical significance in all, is promising and requires further exploration in larger and longer clinical trials of IFN-γ1b.

REFERENCES

1 American Thoracic Society. Idiopathic pulmonary fibrosis: diagnosis and treatment; international consensus statement. Am J Respir Crit Care Med 2000; 161:646–664

2 Ziesche R, Hofbauer E, Wittmann K, et al. A preliminary study of long-term treatment with interferon γ-1b and low-dose prednisolone in patients with idiopathic pulmonary fibrosis. N Engl J Med 1999; 341:1264–1269

3 Raghu G, Brown KK, Bradford WZ, et al. A placebo-controlled trial of interferon γ-1b in patients with idiopathic pulmonary fibrosis. N Engl J Med 2004; 350:125–133

4 Hill AR. Interferon-γ1b for pulmonary fibrosis. N Engl J Med 2004; 350:1794–1797; author reply 1794–1797

5 Hanson D, Winterbauer RH, Kirtland SH, et al. Changes in pulmonary function test results after 1 year of therapy as predictors of survival in patients with idiopathic pulmonary fibrosis. Chest 1995; 108:305–310

6 Flaherty KR, Mumford JA, Murray S, et al. Prognostic implications of physiologic and radiographic changes in idiopathic interstitial pneumonia. Am J Respir Crit Care Med 2003; 168:543–548

7 Latsi PI, du Bois RM, Nicholson AG, et al. Fibrotic idiopathic interstitial pneumonia: the prognostic value of longitudinal functional trends. Am J Respir Crit Care Med 2003; 168:531–537

8 Collard HR, King TE Jr, Bartelson BB, et al. Changes in clinical and physiologic variables predict survival in idiopathic pulmonary fibrosis. Am J Respir Crit Care Med 2003; 168: 538–542

9 Selman M, Thannickal VJ, Pardo A, et al. Idiopathic pulmonary fibrosis: pathogenesis and therapeutic approaches. Drugs 2004; 64:405–430

10 Pocock SJ, Hughes MD, Lee RJ et al. Statistical problems in the reporting of clinical trials. N Engl J Med 1987; 317:426–432