Validation of test performance characteristics and minimal clinically important difference of the 6-minute walk test in patients with idiopathic pulmonary fibrosis

Steven D. Nathan [a],*, Roland M. du Bois [b], Carlo Albera [c], Williamson Z. Bradford [d], Ulrich Costabel [e], Alex Kartashov [f], Paul W. Noble [g], Steven A. Sahn [h], Dominique Valeyre [i], Derek Weycker [f], Talmadge E. King Jr [j]

[a] Inova Fairfax Hospital, 3300 Gallows Rd., Falls Church, VA 22042, USA
[b] Imperial College, South Kensington Campus, London SW7 2AZ, UK
[c] University of Turin, San Luigi Gonzaga Medical School, Regione Gonzole, 1010043 Orbassano, Turin, Italy
[d] InterMune, Inc., 3280 Bayshore Blvd., Brisbane, CA 94005, USA
[e] Ruhrlandklinik, University Hospital, University of Duisburg-Essen, 45117 Essen, Germany
[f] Policy Analysis Inc. (PAI), 4 Davis Ct., Brookline, MA 02445, USA
[g] Duke University School of Medicine, 106 Research Dr., Durham, NC 27710, USA
[h] Medical University of South Carolina, 96 Jonathan Lucas St., Charleston, SC 29425, USA
[i] Assistance Publique-Hôpitaux de Paris, 3, Ave. Victoria, 75004 Paris, France
[j] University of California, San Francisco, 505 Parnassus Ave., Box 0120, San Francisco, CA 94143, USA

Summary
Background: The 6-minute walk test distance (6MWD) has been shown to be a valid and responsive outcome measure in patients with idiopathic pulmonary fibrosis (IPF). The analyses were based, however, on a single phase 3 trial and require validation in an independent cohort.
Objective: To confirm the performance characteristics and estimates of minimal clinically important difference (MCID) of 6MWD in an independent cohort of patients with IPF.
Methods: Patients randomized to placebo in the phase 3 CAPACITY trials who had a baseline 6MWD measurement were included in these analyses. The 6MWD and other functional parameters (lung function, dyspnea, and health-related quality of life) were measured at baseline and 24-week intervals. Validity and responsiveness were examined using Spearman correlation coefficients. The MCID was estimated using distribution- and anchor-based methods.
Results: The analysis comprised 338 patients. Baseline 6MWD was significantly correlated with lung function measures, patient-reported outcomes, and quality-of-life measures (validity). Compared with baseline 6MWD, change in 6MWD (responsiveness) showed stronger correlations with change in lung function parameters and quality-of-life measures. Dyspnea measured by the University of California San Diego Shortness of Breath Questionnaire showed the strongest correlations with 6MWD (baseline: coefficient −0.35; 48-week change: coefficient −0.37; both p < 0.001). The distribution-based analyses of MCID using standard error of measurement yielded an MCID of 37 m, and distribution-based analyses by effect size resulted in 29.2 m. The MCID by anchor-based analysis using criterion referencing (health events of hospitalization or death) was 21.7 m.
Conclusions: The 6MWD is a valid and responsive clinical endpoint, which provides objective and clinically meaningful information regarding functional status and near-term prognosis. These results confirm previous findings in an independent cohort of patients with IPF.

Introduction

Idiopathic pulmonary fibrosis (IPF) is a chronic, progressive, and irreversible parenchymal lung disease of unknown cause, which primarily occurs in older individuals [1]. This disease is characterized by reduced lung volumes and impaired gas exchange, and is associated with symptoms of progressive dyspnea, cough, and declining exercise capacity. Although its natural history is quite variable, IPF is typically associated with a poor prognosis; the estimated median survival after diagnosis is only 2−5 years [2−4].

The 6-minute walk test (6MWT) is a widely used measure of exercise tolerance, which displays favorable performance characteristics in a variety of cardiac and pulmonary diseases [5−9]. Several studies have evaluated the prognostic utility of the 6MWT in IPF [10−16]; however, until recently, studies evaluating the performance characteristics of the 6MWT in IPF were limited by small sample size or enrollment of narrowly defined patient cohorts [10−12,15]. However, in a recent study of 822 patients from a randomized controlled study evaluating interferon-gamma 1b, the 6MWT was shown to be a reliable, valid, and responsive measure of disease status and a predictor of 1-year mortality in patients with IPF [17]. In a subsequent study, the investigators found that both baseline 6MWT distance (6MWD) and 24-week change in 6MWD were independent predictors of near-term mortality in an analysis of 748 patients with IPF [18].

In addition to test performance characteristics, several studies have evaluated the minimal clinically important difference (MCID) for 6MWD in patients with IPF. The MCID is the smallest difference in a measure that can be perceived to be important, whether beneficial or harmful, and that would lead a clinician to consider a change in a patient's therapy. Reported MCID estimates, based on maximal distance walked in 6 min, ranged between 10 m and 58 m [14,16,17]. In the largest study to date in patients with IPF, du Bois et al. reported an estimated MCID of 24−45 m [17]. This finding was consistent with the previously reported estimates of 29−34 m and 28 m (range, 10.8−58.5 m) [14,16].

The objectives of the present study were to confirm the test performance characteristics and MCID of the 6MWT in an independent cohort of patients with IPF and, secondarily, to confirm that exercise tolerance as measured by the 6MWT is a clinically important endpoint in IPF. This work has been presented in part at the 2012 international meeting of the European Respiratory Society [19] and the 2013 international meeting of the American Thoracic Society [20].

Materials and methods

Study population

The study population included patients with IPF randomized to the placebo arms of the two phase 3 CAPACITY studies of pirfenidone (ClinicalTrials.gov, numbers NCT00287729 and NCT00287716) [20]. Patients randomized to the pirfenidone arms were not included in the

present analysis due to the drug's significant treatment benefit. Eligibility criteria for the CAPACITY studies were described previously [21]. Briefly, eligible patients were aged 40–80 years with a diagnosis of IPF within the previous 48 months; had no evidence of improvement in disease severity over the previous year; and had predicted FVC ≥50%, hemoglobin-corrected predicted DL_{CO} ≥35%, either predicted FVC or predicted DL_{CO} ≤90%, and a 6MWD ≥150 m.

Study assessments

At screening, oxygen titration was performed to determine the amount of supplemental oxygen each patient needed to complete the 6MWT without developing oxygen desaturation (blood oxygen saturation [SpO_2] <83%). The oxygen flow rate required to maintain SpO_2 of at least 83% for 10 min at rest was determined at screening (subjects in whom SpO_2 ≥83% was not maintained at rest with oxygen flow of 6 L/min were not eligible for study participation). The titration proceeded in increments of 2 L/min, starting with no supplemental oxygen. This oxygen flow rate was used as the initial testing rate to determine the lowest oxygen flow required to maintain SpO_2 ≥83% during the 6MWT. Again, the titration was done in increments of 2 L/min. Subjects who were unable to walk for 6 min or at least 150 m while receiving 6 L/min of oxygen were ineligible for participation in the study. The flow rate determined from this titration was applied for all subsequent 6MWTs in the study.

The 6MWT was administered at baseline and 24-week intervals until week 72. The test was performed on a flat surface according to a standardized protocol. Patients were instructed to walk as far as they could without jogging or running. They were permitted to slow down or stop to rest, if needed, and were encouraged to resume walking as soon as they were able. The 6MWT was stopped if the patient experienced chest pain, intolerable dyspnea, leg cramps, diaphoresis, or desaturation <83% [19].

At regular intervals, assessments were made of physiologic function (FVC, hemoglobin-corrected DL_{CO}, and resting alveolar–arterial gradient of partial pressure of oxygen [A–a PO_2]), dyspnea (University of California San Diego Shortness of Breath Questionnaire [UCSD SOBQ]), and health-related quality of life (HRQOL; St George's Respiratory Questionnaire [SGRQ]). The UCSD SOBQ is a 24-item questionnaire that evaluates the severity of dyspnea during 21 different activities of daily living associated with varying levels of exertion, and includes 3 additional questions about limitations due to shortness of breath, fear of harm due to overexertion, and fear of shortness of breath [22]. Each question is scored from 0 to 5; the total score ranges from 0 to 120, with higher scores indicating greater shortness of breath. The SGRQ is a 50-item questionnaire that evaluates the frequency and severity of respiratory symptoms, activities that are limited by breathlessness or cause breathlessness, and the impact of these disturbances on social and psychological functioning [23]. The total SGRQ score ranges from 0 to 100, with higher scores indicating worsening HRQOL.

Analyses

The validity of the 6MWT was assessed from the association between baseline 6MWD and measures of pulmonary function, dyspnea, and HRQOL [20]. Test responsiveness was assessed by the association between the change from baseline to week 48 in 6MWD and the changes from baseline to week 48 in FVC, hemoglobin-corrected DL_{CO}, UCSD SOBQ score, and SGRQ score; the association with resting A–a PO_2 was analyzed on the basis of change to week 72. In addition, the association between the change in 6MWD from baseline to week 24 and 1-year risk of mortality was evaluated. Associations were determined using Spearman correlation coefficients; the strength of the correlation was determined based on Cohen's criteria, under which an absolute value of a coefficient >0.5 is indicative of a large correlation, 0.5–0.3 of a moderate correlation, 0.3–0.1 of a weak correlation, and <0.1 of a trivial correlation [24]. Associations were also evaluated through a one-way analysis of covariance by comparing mean 6MWD values across subgroups presumed to have different capacities for physical endurance based on each functional parameter (FVC, hemoglobin-corrected DL_{CO}, resting A–a PO_2, UCSD SOBQ score, and SGRQ score). The relation between 24-week change in 6MWD and mortality was evaluated using a Cox proportional hazards model with baseline 6MWD and change in 6MWD as covariates. The proportional hazard assumption for change in 6MWD was evaluated using published methods and was satisfied [25].

Both distribution- and anchor-based methods were used to estimate the MCID for the 6MWD. Distribution-based analyses included standard error of measurement (SEM) and effect size. To calculate the SEM for the 6MWD, the estimated standard deviation (SD) at baseline was multiplied by the square root of 1 minus the estimated reliability coefficient, with 1 SEM defined to be the MCID [17]. To calculate the effect size, the difference in 6MWD mean values at baseline and week 48 was divided by the estimated SD at baseline [19]. Anchor-based analysis used the criterion-referencing approach to estimate the mean difference in baseline 6MWD between patients who did and did not meet the composite endpoint of hospitalization or death [26,27].

Results

Patient characteristics

In all, 347 patients were randomized to the placebo arms in the CAPACITY studies, of whom 338 (97.4%) were included in the present analysis (Table 1). The remaining 9 patients were excluded owing to missing data for the baseline 6MWT. The study cohort had a mean age of 66.5 years, and the majority were men (72.5%). The mean baseline 6MWD was 404.6 m (standard deviation 90.4 m). Mean FVC was 74.7% of predicted, and mean UCSD SOBQ and SGRQ scores were 33.5 and 36.8, respectively. There was, however, substantial variation in these measures across the study cohort. In all, 71 patients (21.0%) were using supplemental oxygen at baseline [19].

Table 1 Baseline characteristics.

Characteristic	Value
Age, y (n = 338)	
Mean (SD)	66.5 (7.6)
Median (IQR)	68.0 (61.0−72.0)
Gender, n (%) (n = 338)	
Men	245 (72.5)
Women	93 (27.5)
6MWD, m (n = 338)	
Mean (SD)	404.6 (90.4)
Median (IQR)	404.5 (353.0−457.0)
FVC, % predicted (n = 338)	
Mean (SD)	74.7 (14.9)
Median (IQR)	72.2 (63.2−84.7)
DL_{CO}, % predicted[a] (n = 336)	
Mean (SD)	46.8 (9.8)
Median (IQR)	45.2 (39.2−52.0)
Resting A−a PO_2, mm Hg (n = 333)	
Mean (SD)	17.8 (12.6)
Median (IQR)	17.4 (10.6−24.1)
UCSD SOBQ score (n = 331)	
Mean (SD)	33.5 (21.5)
Median (IQR)	30.0 (16.0−49.0)
SGRQ score (n = 326)	
Mean (SD)	36.8 (16.7)
Median (IQR)	36.2 (24.3−46.9)
Smoking status, n (%) (n = 338)	
Never	111 (32.8)
Former	210 (62.1)
Current	17 (5.0)
Supplemental oxygen use, n (%)	71 (21.0)

A−a PO2, alveolar–arterial gradient of partial pressure of oxygen; DL_{CO}, carbon monoxide diffusion capacity; IQR, interquartile range; SD, standard deviation; SGRQ, St. George's Respiratory Questionnaire; UCSD SOBQ, University of California San Diego Shortness of Breath Questionnaire; 6MWD, 6-minute walk test distance.

[a] Average hemoglobin-corrected DL_{CO}.

Table 2 Correlation between 6MWD and other measures: baseline and 48-week change.

Variable	Patients, n	Coefficient	P value
Baseline			
FVC, % predicted	338	0.111	0.041
DL_{CO}, % predicted[a]	336	0.212	<0.001
Resting A−a PO_2, mm Hg	333	−0.171	0.002
UCSD SOBQ score	331	−0.346	<0.001
SGRQ score	326	−0.290	<0.001
48-Week change			
FVC, % predicted	295	0.289	<0.001
DL_{CO}, % predicted[a]	291	0.194	<0.001
Resting A−a PO_2, mm Hg[b]	260	−0.191	0.002
UCSD SOBQ score	274	−0.366	<0.001
SGRQ score	271	−0.303	<0.001

DL_{CO}, carbon monoxide diffusion capacity; FVC, forced vital capacity; SGRQ, St. George's Respiratory Questionnaire; UCSD SOBQ, University of California San Diego Shortness of Breath Questionnaire; 6MWD, 6-minute walk test distance.

[a] Average hemoglobin-corrected DL_{CO}.

[b] 72-Week change for resting alveolar–arterial gradient of partial pressure of oxygen (A−a PO_2).

Validity

Correlations between 6MWD and selected measures of functional status were statistically significant (Table 2). The strongest correlations with the 6MWD were observed for dyspnea as measured by UCSD SOBQ score and HRQOL as measured by SGRQ score (both p < 0.001). Patients in the highest quintile (ie, best values) of hemoglobin-corrected DL_{CO}, dyspnea, and HRQOL had a significantly better 6MWD than those in the lowest quintile (ie, poorest values; all p < 0.001; Fig. 1). There was no significant relationship between FVC quintiles and 6MWD.

Responsiveness

In comparison with the baseline values, stronger correlations were generally observed between 48-week change in 6MWD and 48-week changes in functional parameters, including % predicted FVC (p < 0.001), A−a PO_2 (p = 0.002), UCSD SOBQ score (p < 0.001), and SGRQ score

(p < 0.001) (Table 2). As with baseline values, the strongest correlation with 6MWD was observed for the USCD SOBQ measure of dyspnea. The decrease in 6MWD over 48 weeks was significantly greater in quintiles with the largest 48-week decline in functional status than in quintiles with the smallest decline for every functional measure (FVC, hemoglobin-corrected DL_{CO}, resting A−a PO_2, UCSD SOBQ score, and SGRQ score) (Fig. 2). Significant decreases in 6MWD were also evident in the quintiles with the second largest functional declines in several parameters including FVC % predicted, UCSD SOBQ score, and SGRQ score.

In the Cox proportional hazards model, patients with a shorter baseline 6MWD tended to have higher mortality risk (Table 3). Similarly, patients with a >50-m 6MWD decline over the first 24 weeks tended to have greater 1-year mortality risk than patients with a ≤50-m decline (hazard ratio 2.53; 95% confidence interval 0.94−6.79; p = 0.066).

Minimal clinically important difference

Distribution- and anchor-based analyses of 6MWD yielded MCID estimates ranging between 21.7 and 37 m (Table 4). In the distribution-based analyses, the MCID derived from the estimated standard error of measurement was 37 m (95% confidence interval 34−40), calculated by multiplying the estimated standard deviation at baseline by the square root of 1 minus the estimated reliability coefficient (0.83) [17]. The estimated effect size was 0.32 based on a mean change from baseline to week 48 of 29.2 m. This effect size is considered moderate according to Cohen's criteria [24]. Using the criterion-referencing approach, mean baseline 6MWD was significantly different between patients who met the composite endpoint of hospitalization or death and those who did not (p = 0.047). The corresponding MCID,
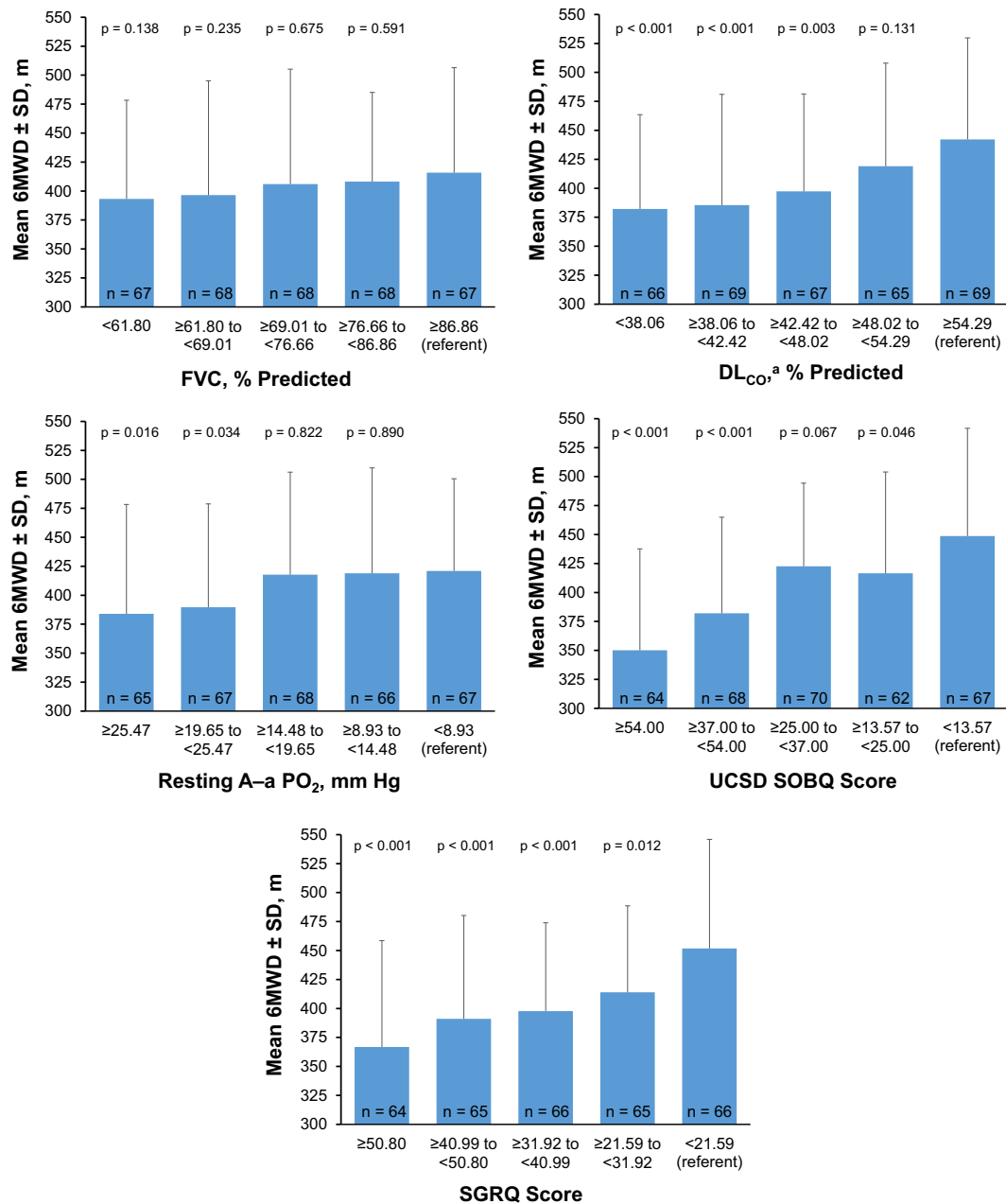
Figure 1 Mean baseline 6-minute walk test distance (6MWD) by quintiles of other measures. Error bars represent standard deviation (SD). P values shown are nominal; no multiplicity adjustments have been performed. [a]Average hemoglobin-corrected $DL_{CO.}$ A−a PO_2, alveolar–arterial gradient of partial pressure of oxygen; DL_{CO}, carbon monoxide diffusion capacity; FVC, forced vital capacity; SGRQ, St. George's Respiratory Questionnaire; UCSD SOBQ, University of California San Diego Shortness of Breath Questionnaire.

based on the difference in mean 6MWD between these patient subsets, was 21.7 m.

Discussion

The present analysis, conducted in an independent cohort of patients with IPF, confirms that 6MWD is a valid and responsive measure of disease status in patients with mild–moderate physical impairment due to IPF [17]. Compared

with baseline 6MWD, change in 6MWD (responsiveness) showed stronger correlations with change in lung function parameters and quality-of-life measures. The estimated MCID for 6MWD in the present study was 21.7−37 m.

The patients evaluated in the present study had been allocated to the placebo arms of the CAPACITY trials and were similar in baseline characteristics to the cohort of the INSPIRE trial evaluated by du Bois and colleagues [17]. Notably, coefficients characterizing the validity and responsiveness of the 6MWD were largely similar between
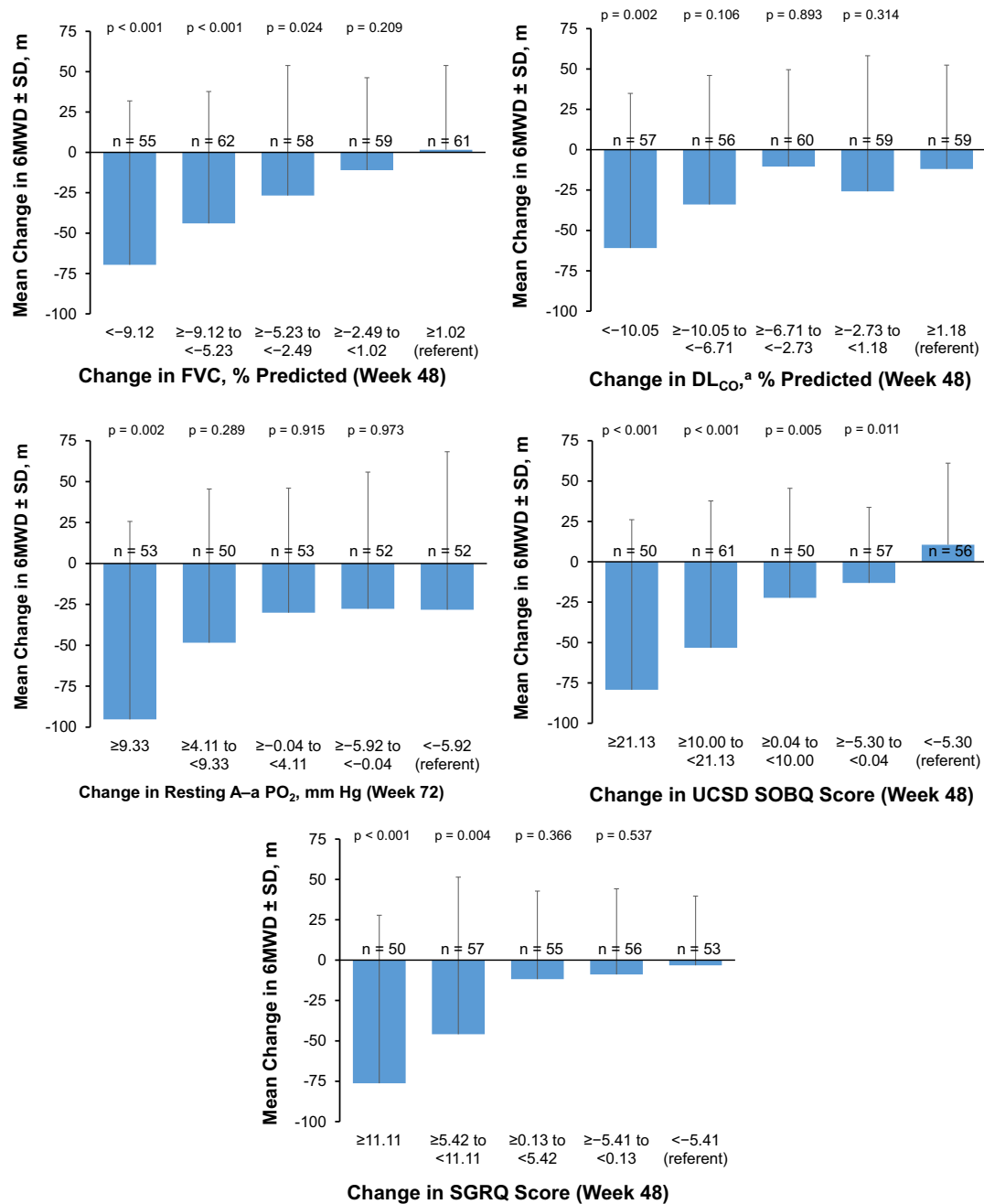
Figure 2 Mean change in 6-minute walk test distance (6MWD) by quintiles of changes in other measures. Error bars represent standard deviation (SD). P values shown are nominal; no multiplicity adjustments have been performed. [a]Average hemoglobin-corrected DL_{CO}. A–a PO_2, alveolar–arterial gradient of partial pressure of oxygen; DL_{CO}, carbon monoxide diffusion capacity; FVC, forced vital capacity; SGRQ, St. George's Respiratory Questionnaire; UCSD SOBQ, University of California San Diego Shortness of Breath Questionnaire.

the 2 independent cohorts (see Supplementary Data). The absolute correlation coefficients reported by du Bois et al. were larger than those in the present study for FVC % predicted and A–a PO_2. In the 48-week comparison (72 weeks for A–a PO_2 in CAPACITY), the present study found larger correlation coefficients for FVC % predicted, hemoglobin-corrected DL_{CO}, UCSD SOBQ score, and SGRQ score; the changes in the latter 2 measures increased the correlation strength to the next level. The associations

between 6MWD and 1-year mortality were also similar between the 2 independent cohorts: the hazard ratios reported here are similar in magnitude to those observed in the larger cohort by du Bois and colleagues. Statistical significance was not achieved for mortality in the present study, which was likely due to a lack of power.

The estimated MCID for 6MWD in the present study was 21.7–37 m, which is consistent with the range of 24–45 m estimated in the study by du Bois and colleagues [17].

Table 3 Cox proportional hazards model of 1-year mortality.

6MWD, m	Patient visits, n	Deaths, n	HR (95% CI)	P value
Baseline				
<250	32	2	2.51 (0.55—11.46)	0.235
250—349	88	4	1.90 (0.60—6.05)	0.279
≥350 (referent)	404	10	—	
24-week change				
<−50	125	7	2.53 (0.94—6.79)	0.066
≥−50 (referent)	399	9	—	

CI, confidence interval; HR, hazard ratio; 6MWD, 6-minute walk test distance.

Because measurement error may limit the utility of small MCID values in the assessment of individual patients, it has been suggested that the high end of the range should be used to assess clinical relevance of changes in individual patients, whereas the lower limit may be used to assess changes at the population level [28]. On the basis of this recommendation, the MCID for 6MWD in individual patients with IPF is 37—45 m and the MCID for IPF populations is 21.7—24 m.

From a clinical perspective, the 6MWT offers a practical and safe method for evaluating disease status in patients with IPF, and serial measurements allow tracking of disease progression over time. The 6MWT requires no special equipment or advanced training and can be performed in all patients with IPF, except the most severely impaired. Because the 6MWT is self-paced, allowing patients to slow down or rest when needed, it is reflective of submaximal exercise capacity. The 6MWT therefore provides a measure of one's ability to complete normal activities of daily living. In the absence of validated and practical clinical instruments for measuring dyspnea and HRQOL in patients with IPF, the 6MWT may serve as a useful benchmark for assessing the degree to which one's ability to engage in activities that affect one's quality of life may be impaired.

Further, the responsiveness of 6MWD to the effect of treatment in IPF was indicated by the findings of a recent phase 3 trial of 555 patients [2]. In that study, pirfenidone therapy resulted in a significant difference compared with placebo for the prespecified secondary efficacy endpoints of change from baseline to week 52 in 6MWD and progression-free survival, a composite endpoint that included a ≥50-m decrease in 6MWD as one indicator of disease progression. These results coupled with those in the present study support the use of 6MWD as a clinical trial endpoint either alone or in the context of a composite.

In addition to its potential utility as a measure of responsiveness to treatment, the 6MWT has been shown to be a reliable, valid, and responsive measure of disease status and a predictor of 1-year mortality in patients with IPF. Although the present study may have been

underpowered for examination of relationships between 6MWD and mortality outcomes, the 24-week change in 6MWD did show a predictive trend consistent with the results of the analyses by du Bois and colleagues [17]. These investigators demonstrated the predictive value of the 6MWT for mortality risk, including the role of both baseline and 24-week change in 6MWD as independent predictors of 1-year all-cause mortality in a multivariate model [18].

Several study limitations are noteworthy. First, the CAPACITY trials enrolled patients with mild—moderate impairment in lung function caused by IPF and, consequently, the present results may not be generalizable to patients with more severe IPF. Second, the analysis included 338 patients randomized to the placebo arms of the CAPACITY trials, but formal statistical analyses of the adequate sample size were not conducted. Although the

Table 4 Estimation of minimal clinically important difference in 6MWD in patients with IPF.

Standard error of measurement			
	Baseline[a]	ICC	SEM[b] (95% CI)
6MWD at baseline (n = 338)	404.6 (90.4)	0.83[c]	37 (34—40)

Effect size				
	Baseline[a]	Week 48[a]	Difference[a]	Effect size[d]
6MWD at baseline and week 48 (n = 296)	410.5 (90.0)	381.3 (120.9)	−29.2 (79.6)	0.32

Criterion referencing				
Criterion	Patients, n	Baseline[a]	Difference[e]	P value[f]
Hospitalization				
No	246	409.2 (91.0)	16.9 (11.0)	0.126
Yes	92	392.3 (88.1)		
Death				
No	318	406.7 (89.7)	36.0 (20.8)	0.084
Yes	20	370.7 (96.0)		
Hospitalization or death				
No	244	410.6 (89.7)	21.7 (10.9)	0.047
Yes	94	388.9 (90.6)		

CI, confidence interval; ICC, intraclass correlation coefficient; MCID, minimal clinically important difference; SEM, standard error of measurement.

[a] Data presented in meters as mean (standard deviation).

[b] Calculated for 6MWD by multiplying the estimated standard deviation at baseline by the square root of 1 minus the estimated reliability coefficient. One SEM was defined to be the MCID; because the SEM is sample-independent, MCID estimates based on the SEM are considered bidirectional in nature.

[c] Based on data from du Bois et al., 2011.

[d] Calculated as difference in 6-minute walk test distance (6MWD) from baseline to week 48 divided by estimated standard deviation at baseline.

[e] Calculated as difference (standard error) in 6MWD in patients by hospitalization, death, or a composite of hospitalization or death.

[f] P value from independent samples t-test.

sample size appeared adequate for confirming the validity and responsiveness of the 6MWT, it may have been insufficient for relating the 6MWT and change in 6MWD to 1-year outcomes. Third, several distribution- and anchor-based methods were used to estimate MCID. Hospitalization and death were included for purposes of comparison with prior studies. It was anticipated that the criterion-referencing approach based on hospitalization and death might overestimate MCID. Given that death and hospitalization are major events, one would expect the MCID for minor events to be smaller than those values reported (and thus those reported could be interpreted as "upper bounds"). However, it appeared that the result was generally consistent with—and perhaps somewhat lower than—the value determined by the distribution-based approach. Lastly, patients receiving pirfenidone were not included in this analysis as this pharmacologic intervention has been shown to slow the rate of decline in the 6MWT. The 6MWT changes are not necessarily concordant with other potential effects of the drug, which would therefore have affected our analyses of serial change.

Finally, the 2 independent cohorts from the INSPIRE and CAPACITY trials were very similar but not identical. For example, the INSPIRE patients were enrolled at 81 centers in 7 European countries (Belgium, France, Germany, Ireland, Italy, Spain, and the United Kingdom), the United States, and Canada, whereas the CAPACITY patients came from those countries plus Poland, Switzerland, Mexico, and Australia. There were also minor differences between the INSPIRE and CAPACITY trials in their inclusion and exclusion criteria (for example, a lower inclusion threshold for FVC % predicted of 55% in INSPIRE and 50% in CAPACITY). However, we perceive these subtle differences as a strength of our analysis, as they demonstrate that a particular trial population need not be replicated to achieve consistently valid and reliable results with the 6MWT.

In summary, the present analysis of results from the placebo arms of 2 large multinational, randomized, phase 3 trials in IPF with similar study designs and populations provides confirmatory evidence pertaining to the performance characteristics of the 6MWT. Specifically, the 6MWD is a valid and responsive clinical endpoint that provides objective and clinically meaningful information regarding the functional status and near-term prognosis of patients with IPF. Our study further helps to establish this easy-to-perform test as an important assessment tool in the practical management of IPF.

Role of the funding source

Author disclosures

Acknowledgments

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.rmed.2015.04.008.

References

[1] Raghu G, Collard HR, Egan JJ, et al. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. Am J Respir Crit Care Med 2011;183:788—824.

[2] King Jr TE, Bradford WZ, Castro-Bernardini S, et al. A phase 3 trial of pirfenidone in patients with idiopathic pulmonary fibrosis. N Engl J Med 2014;370:2083—92.

[3] Nathan SD, du Bois RM, Albera C, et al. 6-minute walk test (6MWT) in patients with idiopathic pulmonary fibrosis (IPF): confirmation of test performance characteristics. Am J Respir Crit Care Med 2013;187:A2360.

[4] Rudd RM, Prescott RJ, Chalmers JC, et al. British Thoracic Society study on cryptogenic fibrosing alveolitis: response to treatment and survival. Thorax 2007;62:62—6.

[5] Baughman RP, Sparkman BK, Lower EE. Six minute walk test and health status assessment in sarcoidosis. Chest 2007;132: 207—13.

[6] Bernstein ML, Despars JA, Singh NP, et al. Reanalysis of the 12-minute walk in patients with chronic obstructive pulmonary disease. Chest 1994;105:163—7.

[7] Cahalin LP, Mathier MA, Semigran MJ, et al. The six-minute walk test predicts peak oxygen uptake and survival in patients with advanced heart failure. Chest 1996;110:325—32.

[8] Gomberg-Maitland M, Huo D, Benza RL, et al. Creation of a model comparing 6-minute walk test to metabolic equivalent in evaluating treatment effects in pulmonary arterial hypertension. J Heart Lung Transpl 2007;26:732—8.

[9] Lipkin DP, Scriven AJ, Crake T, et al. Six minute walking test for assessing exercise capacity in chronic heart failure. Br Med J 1986;292:653—5.

[10] Caminati A, Bianchi A, Cassandro R, et al. Walking distance on 6-MWT is a prognostic factor in idiopathic pulmonary fibrosis. Respir Med 2009;103:117—23.

[11] Chetta A, Aiello M, Foresi A, et al. Relationship between outcome measures of six-minute walk test and baseline lung function in patients with interstitial lung disease. Sarcoidosis Vasc Diffuse Lung Dis 2001;18:170—5.

[12] Eaton T, Young P, Milne D, et al. Six-minute walk, maximal exercise tests: reproducibility in fibrotic interstitial pneumonia. Am J Respir Crit Care Med 2005;171:1150—7.

[13] Hallstrand TS, Boitano LJ, Johnson WC, et al. The timed walk test as a measure of severity and survival in idiopathic pulmonary fibrosis. Eur Respir J 2005;25:96—103.

[14] Holland AE, Hill CJ, Conron M, et al. Small changes in six-minute walk distance are important in diffuse parenchymal lung disease. Respir Med 2009;103:1430—5.

[15] Lederer DJ, Arcasoy SM, Wilt JS, et al. Six-minute-walk distance predicts waiting list survival in idiopathic pulmonary fibrosis. Am J Respir Crit Care Med 2006;174:659—64.

[16] Swigris JJ, Wamboldt FS, Behr J, et al. The 6 minute walk in idiopathic pulmonary fibrosis: longitudinal changes and minimum important difference. Thorax 2010;65:173—7.

[17] du Bois RM, Weycker D, Albera C, et al. Six-minute walk test in idiopathic pulmonary fibrosis. Test validation and minimal clinically important difference. Am J Respir Crit Care Med 2011;183:1231—7.

[18] du Bois RM, Albera C, Bradford WZ, et al. 6-Minute walk test distance is an independent predictor of mortality in patients with idiopathic pulmonary fibrosis. Eur Respir J 2014;43:1421—9.

[19] Nathan SD, Albera C, du Bois RM, et al. 6-Minute walk test in patients with idiopathic pulmonary fibrosis: confirmation of the minimal clinically important difference. Poster presented at: European Respiratory Society. September 1-5, 2012 [Vienna, Austria].

[20] Nathan SD, Shlobin OA, Weir N, et al. Long-term course and prognosis of idiopathic pulmonary fibrosis in the new millennium. Chest 2011;140:221—9.

[21] Noble PW, Albera C, Bradford WZ, et al. Pirfenidone in patients with idiopathic pulmonary fibrosis (CAPACITY): two randomised trials. Lancet 2011;377:1760—9.

[22] Eakin EG, Resnikoff PM, Prewitt LM, et al. Validation of a new dyspnea measure: the UCSD Shortness of Breath Questionnaire. Chest 1998;113:619—24.

[23] Jones PW, Quirk FH, Baveystock CM. The St. George's Respiratory Questionnaire. Respir Med 1991;85(suppl B):25—31.

[24] Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. New Jersey: Lawrence Erlbaum Associates; 1988.

[25] Allison PD. Survival analysis using the SAS system: a practical guide. Cary, NC: SAS Institute, Inc.; 1995. p. 292.

[26] Wyrich KW, Nienaber NA, Tierney WM, et al. Linking clinical relevance and statistical significance in evaluating intra-individual changes in health-related quality of life. Med Care 1999;37:469—78.

[27] Wyrich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. J Clin Epidemiol 1999;52:861—73.

[28] Yost KJ, Eton DT. Combining distribution- and anchor-based approaches to determine minimally important differences: the FACIT experience. Eval Health Prof 2005;28:172—91.