Six-Minute-Walk Test in Idiopathic Pulmonary Fibrosis

Test Validation and Minimal Clinically Important Difference

Roland M. du Bois[1], Derek Weycker[2], Carlo Albera[3], Williamson Z. Bradford[4], Ulrich Costabel[5], Alex Kartashov[2], Lisa Lancaster[6], Paul W. Noble[7], Steven A. Sahn[8], Javier Szwarcberg[4], Michiel Thomeer[9], Dominique Valeyre[10], and Talmadge E. King, Jr.[11]

[1]National Heart and Lung Institute, Imperial College, London, United Kingdom; [2]Policy Analysis, Inc., Brookline, Massachusetts; [3]Department of Clinical and Biological Sciences, University of Turin, Turin, Italy; [4]InterMune, Inc., Brisbane, California; [5]Ruhrlandklinik and Medical Faculty, University of Duisburg/Essen, Essen, Germany; [6]Vanderbilt University Medical Center, Nashville, Tennessee; [7]Duke University School of Medicine, Durham, North Carolina; [8]Medical University of South Carolina, Charleston, South Carolina; [9]University Hospitals Leuven, Leuven, Belgium; [10]Assistance Publique-Hôpitaux de Paris, Hospital Avicenne, Bobigny, France; [11]University of California, San Francisco, California

***Rationale*: The 6-minute-walk test (6MWT) is a practical and clinically meaningful measure of exercise tolerance with favorable performance characteristics in various cardiac and pulmonary diseases. Performance characteristics in patients with idiopathic pulmonary fibrosis (IPF) have not been systematically evaluated.**

***Objectives*: To assess the reliability, validity, and responsiveness of the 6MWT and estimate the minimal clinically important difference (MCID) in patients with IPF.**

***Methods*: The study population included all subjects completing a 6MWT in a clinical trial evaluating interferon gamma-1b (n = 822). Six-minute walk distance (6MWD) and other parameters were measured at baseline and at 24-week intervals using a standardized protocol. Parametric and distribution-independent correlation coefficients were used to assess the strength of the relationships between 6MWD and measures of pulmonary function, dyspnea, and health-related quality of life. Both distribution-based and anchor-based methods were used to estimate the MCID.**

***Measurements and Main Results*: Comparison of two proximal measures of 6MWD (mean interval, 24 d) demonstrated good reliability (coefficient = 0.83; $P < 0.001$). 6MWD was weakly correlated with measures of physiologic function and health-related quality of life; however, values were consistently and significantly lower for patients with the poorest functional status, suggesting good construct validity. Importantly, change in 6MWD was highly predictive of mortality; a 24-week decline of greater than 50 m was associated with a fourfold increase in risk of death at 1 year (hazard ratio, 4.27; 95% confidence interval, 2.57–7.10; $P < 0.001$). The estimated MCID was 24–45 m.**

***Conclusions*: The 6MWT is a reliable, valid, and responsive measure of disease status and a valid endpoint for clinical trials in IPF.**

Keywords: interstitial lung disease; idiopathic pulmonary fibrosis; exercise test

AT A GLANCE COMMENTARY

Scientific Knowledge on the Subject

The 6-minute-walk test is a widely used measure of exercise tolerance that has been validated in a variety of cardiac and pulmonary diseases. Performance characteristics of the 6-minute-walk test have not been systematically evaluated in patients with idiopathic pulmonary fibrosis.

What This Study Adds to the Field

Our findings suggest that the 6-minute-walk test is a reliable, valid, and responsive measure of exercise tolerance in patients with idiopathic pulmonary fibrosis and that the minimal clinically important difference in 6-minute-walk distance is between 24 and 45 meters.

Idiopathic pulmonary fibrosis (IPF) is a progressive, life-threatening, interstitial lung disease of unknown etiology. It is characterized by a progressive decline in lung function that limits and eventually precludes routine physical activity. IPF is the most common and the most lethal idiopathic interstitial pneumonia, with a median survival of only 2–5 years after diagnosis (1–3).

The 6-minute-walk test (6MWT) is a practical and reliable measure of exercise tolerance that is widely used to assess the functional status of patients with a variety of cardiac and pulmonary diseases, including heart failure, pulmonary hypertension, and chronic obstructive pulmonary disease (COPD) (4–9). However, studies evaluating the measurement properties of the 6MWT in patients with IPF have been limited by small sample sizes or narrowly defined patient subsets and, presumably because of these limitations, have generally yielded inconsistent results (10–18). Moreover, the minimal clinically important difference (MCID) for the 6MWT in patients with IPF has been estimated in only two studies, each with limitations as noted previously (16, 18). The MCID is the smallest difference in a measure that may be perceived to be important, either beneficial or harmful, and that would lead a clinician to consider a change in a patient's therapy (19). MCID is a clinically important concept, because it may assist with the interpretation of observed changes in a measure and may influence the perceived success of an intervention. In addition, the MCID may have implications for the design of clinical trials in terms of the selection of primary and secondary endpoints and the determination of sample size (20, 21).

In the present study, we used data from the largest clinical trial to date in patients with IPF to assess the reliability, validity, and responsiveness of the 6MWT and estimate the MCID in patients with this disease. This work has been presented in part at the 2010 international meeting of the American Thoracic Society (22).

(*Received in original form July 29, 2010; accepted in final form December 3, 2010*)

Supported by InterMune, Inc., Brisbane, California.

Correspondence and requests for reprints should be addressed to Roland M. du Bois, M.D., National Heart & Lung Institute, Imperial College, London, UK. E-mail: r.dubois@imperial.ac.uk

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

Am J Respir Crit Care Med Vol 183. pp 1231–1237, 2011
Originally Published in Press as DOI: 10.1164/rccm.201007-1179OC on December 3, 2010
Internet address: www.atsjournals.org

METHODS

Study Population

The study population consisted of all randomized subjects in a placebo-controlled Phase 3 clinical trial of interferon gamma-1b who completed the 6MWT at the baseline study visit (n = 822). Analysis of outcomes in the original study revealed no evidence of a treatment effect; therefore, the present analysis included data from both treatment arms to maximize study power (23). Criteria for enrollment in the original trial have been described elsewhere (23). Briefly, eligible patients had a confident IPF diagnosis according to the criteria of the American Thoracic Society (24, 25), FVC greater than or equal to 55% of predicted, diffusing capacity of carbon monoxide (D_{LCO}) greater than or equal to 35% of predicted, either FVC or D_{LCO} less than or equal to 90% of predicted, and a 6MWT distance greater than or equal to 150 m.

Study Protocol

The 6MWT was performed at the screening and baseline visit and at 24-week intervals thereafter using a standardized protocol (*see* online supplement). The test was performed indoors on a flat, straight corridor with a hard surface. An oxygen titration procedure was performed at the screening visit to establish a baseline flow rate for patients who required supplemental oxygen. Before each 6MWT, patients were required to have resting oxygen saturation as measured by pulse oximetry of at least 83% after 10 minutes of rest breathing room air or at the baseline O_2 flow rate. Patients were instructed to walk as far as they could without jogging or running; if they needed to slow down or stop to rest they were permitted to do so and encouraged to resume walking as soon as they were able. The test was stopped if the patient experienced chest pain, intolerable dyspnea, leg cramps, diaphoresis, or desaturation below 83%.

Assessments of physiologic function (FVC, D_{LCO}, resting A-a gradient), dyspnea (University of California San Diego Shortness of Breath Questionnaire), and health-related quality of life (HRQL; St. George's Respiratory Questionnaire) were also performed at the screening or baseline visit and at 24-week intervals thereafter (23). The total score of the University of California San Diego Shortness of Breath Questionnaire ranges from 0–120, and the score increases with extent of dyspnea (26). The St. George's respiratory questionnaire is comprised of three respiratory-specific domains: (*1*) symptoms, (*2*) activity, and (*3*) impacts. Each domain of the questionnaire ranges from 0–100, with an increasing score indicating worsening HRQL (27).

Statistical Analyses

Reliability was assessed based on the stability of 6MWT distance (6MWD) between the screening and baseline visits. The intraclass correlation coefficient was used to assess the strength of the relationship; a value greater than or equal to 0.80 was assumed to represent "good" reliability. Analyses were conducted using observed data for the subgroup of patients who did not receive supplemental oxygen during the 6MWT; for purposes of comparison, analyses were also conducted using data for all randomized patients.

Criterion validity was assessed based on relationships between 6MWD and measures of physiologic status, dyspnea, and HRQL; distribution-independent (Spearman) correlation coefficients were used to assess the strength of these relationships. Strength of correlation was designated as follows: greater than 0.5, large; 0.5–0.3, moderate; 0.3–0.1, small; and less than 0.1, trivial (28). Construct validity was assessed by comparing the mean 6MWD across subgroups of patients presumed to have different capacities for physical endurance, defined on the basis of physiologic function, dyspnea, and HRQL. Patients were stratified into subgroups based on quintiles of the corresponding distributions. One-way analysis of variance was used for statistical comparisons.

Responsiveness was assessed using Spearman correlation coefficients between the 48-week change in 6MWD (i.e., between baseline and the Week 48 study visit) and changes over the same period in measures of physiologic function, dyspnea, and HRQL. The relationship between mean changes in 6MWD and changes in other measures (stratified into quintiles) was examined using analysis of variance. Responsiveness was also evaluated by examining the relationship between 24-week change in 6MWD and 1-year risk of death using a Cox proportional hazards model. Change in 6MWD was evaluated over the 24-week periods immediately preceding the Week 24 and Week 72 trial visits, respectively, and was defined initially based on quintiles of the distribution and subsequently modified based on clinical and statistical considerations. All deaths occurring over the 48-week period after the Week 24 and Week 72 visits, respectively, were included in the analysis.

Both distribution-based and anchor-based methods were used to estimate the MCID. Distribution-based methods included the standard error of measurement (SEM) and effect size. SEM was calculated for 6MWD by multiplying the estimated standard deviation at baseline by the square root of one 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 to be bidirectional in nature (29, 30).

Effect size was calculated by dividing the difference in 6MWD values at baseline and Week 48 by the estimated standard deviation at baseline. A change in value corresponding to a "small" effect size was considered to approximate the MCID (21, 31, 32). One-third of the estimated standard deviation has also been suggested as an approximation of MCID (33).

The anchor-based method of estimating the MCID used the criterion-referencing approach, which involved estimation of differences in 6MWD at baseline between patients who did and did not experience selected health events during the subsequent 48-week period. An independent samples *t* test was used for statistical comparisons.

RESULTS

Patient Characteristics

A total of 826 patients were randomized to interferon gamma-1b (n = 551) or placebo (n = 275) in the Phase III trial; among these, two patients in each group were missing baseline data for the 6MWT and were excluded from the study population (n = 822) (Table 1). Mean (± SD) age at study entry was 66 (± 8) years and 71% of study subjects were male. The mean (± SD) value for 6MWD at baseline was 392 (± 109) m; the interquartile

TABLE 1. BASELINE CHARACTERISTICS

Characteristic	Value*
Age, yr	
Mean (SD)	66 (7.8)
Median (IQR)	67 (61–72)
Sex, n (%)	
Male	582 (70.8)
Female	240 (29.2)
6MWT Distance, m	
Mean (SD)	392.4 (108.5)
Median (IQR)	395 (328–462)
FVC, % predicted	
Mean (SD)	72.5 (12.7)
Median (IQR)	70.5 (62.2–80.7)
D_{Lco}, % predicted	
Mean (SD)	47.4 (9.2)
Median (IQR)	46.1 (40.5–52.7)
Resting $AaPo_2$, mm Hg	
Mean (SD)	19.1 (10.7)
Median (IQR)	18.8 (11.9–25.9)
UCSD SOBQ score	
Mean (SD)	34.9 (22.8)
Median (IQR)	30.4 (17–51)
SGRQ score	
Mean (SD)	41.8 (18)
Median (IQR)	40.8 (28.6–54.3)

Definition of abbreviations: 6MWT = 6-minute-walk test; $AaPo_2$ = alveolar-arterial oxygen gradient; D_{Lco} = carbon monoxide diffusion capacity; IQR = interquartile range; SGRQ = St. George's Respiratory Questionnaire; UCSD SOBQ = University of California San Diego Shortness of Breath Questionnaire.

* n = 822 except for D_{Lco} (820), $AaPo_2$ (807), UCSD SOBQ (788), and SGRQ (742).

TABLE 2. CORRELATION BETWEEN 6MWT DISTANCE AND OTHER MEASURES

Variable	N	Coefficient*	*P* Value†
FVC, % predicted	822	0.121	<0.001
$D_{L_{CO}}$, % predicted	820	0.135	<0.001
Resting $AaPo_2$, mm Hg	807	−0.188	<0.001
UCSD SOBQ score	788	−0.290	<0.001
SGRQ score	742	−0.255	<0.001

Definition of abbreviations: 6MWT = 6-minute-walk test; $AaPo_2$ = alveolar-arterial oxygen gradient; $D_{L_{CO}}$ = carbon monoxide diffusion capacity; SGRQ = St. George's Respiratory Questionnaire; UCSD SOBQ = University of California San Diego Shortness of Breath Questionnaire.

* Spearman correlation coefficient.

† One-way analysis of variance.

range was 328–462 m. There was substantial variation across study subjects in physiologic measures, dyspnea, and HRQL. Baseline characteristics of patients randomized to placebo versus interferon gamma-1b, and the 722 patients who did not require supplemental oxygen to complete the 6MWT, were similar (*see* online supplement).

Reliability

The 6MWT seemed to have good reliability in patients with IPF. The intraclass correlation coefficient for 6MWD between the screening and baseline visits (mean interval, 24 d) was 0.83 ($P < 0.001$) using data for subjects who did not receive supplemental oxygen (n = 718), and 0.82 ($P < 0.001$) using data for all randomized subjects (n = 821). The correlation coefficient was slightly lower (0.72; $P < 0.001$) for the small subset of patients who required supplemental oxygen (n = 103); however, the degree to which this reflects test reliability in this subgroup is uncertain, because these patients may have been more likely to experience real change during the interval between measurements. Because reliability is based on the presumption of stable disease between proximal measurements, these patients are not optimal candidates for tests of reliability.

Validity

Correlations between 6MWD and measures of physiologic function, dyspnea, and HRQL were in the expected direction, but generally weak (absolute values of all coefficients, <0.30; $P < 0.001$) (Table 2). Patients in the fifth quintile (i.e., those with the best scores) of physiologic function, dyspnea, and HRQL generally had significantly better 6MWT values than those in the other quintiles (i.e., those with poorer scores for physiologic function, dyspnea, and HRQL) (Table 3).

Responsiveness

Correlations between changes in 6MWD and changes in measures of physiologic function, dyspnea, and HRQL were in the expected direction, but generally weak (absolute values of all coefficients, ≤ 0.27; $P < 0.001$) (Table 4). Decline in 6MWD was generally significantly greater for patients in the first and second quintiles of change in physiologic function, dyspnea, and HRQL versus those in the fifth quintile (Table 5). Both baseline 6MWD and the 24-week change in 6MWD were highly predictive of death over the subsequent 1-year period (Table 6). Importantly, the risk of death was more than fourfold higher (hazard ratio [HR], 4.27; 95% confidence interval [CI], 2.57– 7.10; $P < 0.001$) for patients with a decline in 6MWD greater than 50 m, and more than threefold higher (HR, 3.59; 95% CI, 1.95–6.63; $P < 0.001$) for those with a decline of 26–50 m, compared with subjects for whom the decline in 6MWD was less than or equal to 25 m. Treatment assignment and the interaction term for treatment assignment and change in 6MWD were not important predictors of death, and the proportional hazards assumption for change in 6MWD was satisfied.

To confirm the validity of including patients from both the experimental and control arms in the analyses, data were analyzed separately for the subgroups of patients who were randomized to treatment with interferon gamma-1b and placebo, respectively, in the original clinical trial. There were no meaningful differences between groups in 6MWD reliability, validity, or responsiveness. Additionally, responsiveness was similar when change in 6MWD was assessed over either 24 or 48 weeks.

Minimal Clinically Important Difference

The estimated SEM for the 6MWT, and the corresponding MCID, was 45 m (95% CI, 42–47) (Table 7). The estimated effect size for 6MWD was 0.28, based on a mean change in value of 31 m between the baseline and Week 48 visits; according to Cohen's criteria (28), such an effect should be considered "small." One-third of the estimated standard deviation at baseline yielded a figure of 36.

Baseline 6MWT values were significantly different for patients who experienced the composite endpoint of hospitalization or death versus those who did not; the corresponding estimated MCID was 24 m ($P = 0.009$). Differences in 6MWD between patients who were hospitalized and those who were not, and patients who died and those who did not, were 18 m ($P = 0.086$) and 27 m ($P = 0.059$), respectively.

DISCUSSION

The 6MWT is a widely used measure of exercise tolerance in patients with various cardiac and pulmonary diseases. From a clinical perspective, it has the advantages of practicality and safety; it requires no special equipment or advanced training, and unlike maximal cardiopulmonary exercise testing, it can be performed by all but the most severely impaired patients (34). Moreover, because the 6MWT is self-paced, it is better tolerated and more reflective of daily activities than other maximal exercise tests (10).

Based on these characteristics, the 6MWT represents a clinically meaningful tool that may be particularly well-suited to the assessment of functional status in patients with IPF. To date, however, studies evaluating the performance characteristics of the 6MWT in patients with IPF have been limited by small sample size and have generally yielded inconsistent results (12, 13, 15–17). In a study of 29 patients with fibrotic idiopathic interstitial pneumonia, Eaton and coworkers (13) reported excellent intrasubject reproducibility for the 6MWT (coefficient of variation = 4.2%) and high correlations between 6MWD and V_{O_2}max on maximal exercise testing ($r^2 = 0.78$) and percent predicted $D_{L_{CO}}$ ($r^2 = 0.61$). However, there was virtually no correlation between 6MWD and percent predicted FVC ($r^2 = 0.06$). Additionally, analysis of outcomes during a median follow-up of 28 months revealed no significant relationship between 6MWD and mortality. In a subsequent retrospective analysis of 44 patients with IPF, Caminati and coworkers (17) reported moderate correlations between 6MWD and both percent predicted FVC and percent predicted $D_{L_{CO}}$ ($r = 0.40$ and 0.42, respectively). In contrast to the findings of Eaton and coworkers (13), both 6MWD and change in 6MWD at 1 year were predictors of mortality, although the magnitude of the HR was small (HR = 0.995 and HR = 0.994 for 6MWD and change in 6MWD, respectively).

TABLE 3. MEAN (SD) 6MWT DISTANCE BY QUINTILES OF OTHER MEASURES

Variable*	N	6MWT Distance*	*P* Value†
FVC, % predicted			
<60.45	165	368.2 (103.7)	0.003
⩾60.45 to <66.63	164	384.7 (114.1)	0.110
⩾66.63 to <74.39	163	400 (106)	0.756
⩾74.39 to <83.01	165	405.5 (108)	0.882
⩾83.01	165	403.7 (107.1)	—
$D_{L_{CO}}$,% predicted			
<39.37	165	368.2 (104.5)	<0.001
⩾39.37 to <43.93	162	379.8 (110.2)	0.014
⩾43.93 to <48.57	164	398 (101)	0.341
⩾48.57 to <54.45	165	405.9 (107.2)	0.771
⩾54.45	164	409.3 (115.1)	—
Resting $AaPo_2$, mm Hg			
⩾27.75	163	353.5 (109.1)	<0.001
⩾21.36 to <27.75	160	386.1 (115.8)	<0.001
⩾15.97 to <21.36	161	400.6 (101)	0.028
⩾10.12 to <15.97	163	393.6 (102.2)	0.005
<10.12	160	426.8 (103.8)	—
UCSD SOBQ score			
⩾56.00	157	344.1 (109.2)	<0.001
⩾38.18 to <56.00	156	383.9 (107.7)	<0.001
⩾25.00 to <38.18	172	390.9 (105.6)	<0.001
⩾13.09 to <25.00	143	400.9 (100.9)	<0.001
<13.09	160	443.3 (96.8)	—
SGRQ score			
⩾58.22	148	351.8 (106.1)	<0.001
⩾46.24 to <58.22	148	369.6 (107.4)	<0.001
⩾36.03 to <46.24	148	399.8 (111.6)	<0.001
⩾25.17 to <36.03	149	398.7 (101.8)	0.005
<25.17	149	433.5 (98.8)	—

Definition of abbreviations: 6MWT = 6-minute-walk test; $AaPo_2$ = alveolar-arterial oxygen gradient; $D_{L_{CO}}$ = carbon monoxide diffusion capacity; SGRQ = St. George's Respiratory Questionnaire; UCSD SOBQ = University of California San Diego Shortness of Breath Questionnaire.

* Data are presented as mean (SD).

† One-way analysis of variance.

In the present study, we assessed the reliability, validity, and responsiveness of the 6MWT and estimated the MCID in a large, well-defined cohort of patients with IPF who performed the test according to a standardized protocol. Consistent with previous studies, the distance walked during the 6MWT was highly reproducible, demonstrating good reliability. Concurrent validity, based on correlations with various measures of pulmonary function, was found to be weak; however, construct validity was found to be much better, as 6MWT values were lowest for patients with the poorest levels of physiologic function and the worst scores for dyspnea and HRQL. Responsiveness, as assessed by the relationship between 24-week changes in 6MWD and the risk of 1-year mortality, was also found to be good. Patients experiencing a decline as small as 26 m over 24 week were found to be at significantly higher risk of death than those in the referent group, whereas those experiencing a decline of greater than 50 m over 24 weeks had a fourfold increase in the risk of death at 1 year. This finding is especially important, as it suggests that the 6MWT is an important measure of prognosis, and thus might be used as a key physiologic outcome parameter in clinical trials in patients with IPF. Indeed, the sensitivity of the progression-free survival endpoint, currently defined in most clinical trials in IPF on the basis of a 10% decline in percent predicted FVC, might be enhanced by the addition of a 6MWD criterion defined on the basis a 50-m decrement. Such an endpoint has the potential to increase the expected event rate and thereby decrease the necessary size and duration of subsequent clinical trials.

A novel and important finding of our study was the observation that the 24-week change in 6MWD was highly predictive of 1-year mortality despite relatively weak correlations between 6MWD and various measures of pulmonary function known to be independent predictors of mortality. This suggests that the

TABLE 4. CORRELATION BETWEEN CHANGE IN 6MWT DISTANCE AND CHANGE IN OTHER MEASURES

Variable	N	Coefficient*	*P* Value†
FVC, % predicted	697	0.270	<0.001
$D_{L_{CO}}$,% predicted	667	0.175	<0.001
Resting $AaPo_2$, mm Hg	644	−0.194	<0.001
UCSD SOBQ score	628	−0.203	<0.001
SGRQ score	571	−0.231	<0.001

Definition of abbreviations: 6MWT = 6-minute-walk test; $AaPo_2$ = alveolar-arterial oxygen gradient; $D_{L_{CO}}$ = carbon monoxide diffusion capacity; IQR = interquartile range; SGRQ = St. George's Respiratory Questionnaire; UCSD SOBQ = University of California San Diego Shortness of Breath Questionnaire.

* Spearman correlation coefficient.

† Analysis of variance.

TABLE 5. MEAN (SD) CHANGE IN 6MWT DISTANCE BY QUINTILES OF CHANGE IN OTHER MEASURES

Variable*	N	6MWT Distance, m*	*P* Value†
FVC, % predicted			
<−9.96	132	−98 (147.6)	<0.001
⩾ −9.96 to <−4.58	139	−48 (110.6)	<0.001
⩾ −4.58 to <−1.30	144	−6.2 (95.6)	0.437
⩾ −1.30 to <2.13	142	−8.9 (108.5)	0.328
⩾2.13	140	4.3 (100.5)	—
$D_{L_{CO}}$,% predicted			
<−11.30	131	−68.6 (141.1)	<0.001
⩾ −11.30 to <−6.67	135	−30.7 (122.2)	0.037
⩾ −6.67 to <−2.92	132	−17.6 (99.9)	0.253
⩾ −2.92 to <1.94	132	−18.5 (102.6)	0.227
⩾1.94	137	−1.6 (103.6)	—
UCSD SOBQ score			
⩾18.00	126	−84.2 (162)	<0.001
⩾8.00 to <18.00	133	−28.9 (114)	0.345
⩾1.00 to <8.00	132	−13.2 (100.2)	0.915
⩾ −8.00 to <1.00	130	−8.6 (84.8)	0.681
<−8.00	107	−14.8 (96.1)	—
SGRQ score			
⩾12.32	110	−74.1(126.9)	<0.001
⩾5.53 to <12.32	117	−38.5 (132.1)	0.016
⩾0.35 to <5.53	115	−19.2 (118.4)	0.250
⩾ −8.37 to <0.35	113	−23.2 (103.5)	0.160
<−8.37	116	−1.60 (98)	—

Definition of abbreviations: 6MWT = 6-minute-walk test; $AaPo_2$ = alveolar-arterial oxygen gradient; $D_{L_{CO}}$ = carbon monoxide diffusion capacity; SGRQ = St. George's Respiratory Questionnaire; UCSD SOBQ = University of California San Diego Shortness of Breath Questionnaire.

* Data are presented as mean (SD).

† One-way analysis of variance.

TABLE 6. COX PROPORTIONAL HAZARDS MODEL

	Patient Visits (*n*)	Deaths (*n*)	HR (*95% CI*)	*P* Value
Δ6MWT distance, m				
<−50	317	40	4.27 (2.57–7.10)	<0.001
−50 to −26	117	18	3.59 (1.95–6.63)	<0.001
⩾ −25	720	24		
6MWT distance, m				
<250	130	15	2.65 (1.48–4.74)	0.001
250 to 349	255	20	1.54 (0.91–2.60)	0.106
⩾350	823	47		

Definition of abbreviations: Δ6MWT distance = 24-week change in 6-minute walk test distance; CI = confidence interval; HR = hazard ratio.

TABLE 7. ESTIMATION OF THE MINIMAL CLINICALLY IMPORTANT DIFFERENCE IN 6MWT DISTANCE IN PATIENTS WITH IPF

Standard Error of Measurement				
	Mean	SD	Correlation	SEM (*95% CI*)
6MWT distance, m	392.4	108.3	0.83	45 (42-47)
Effect Size				
	Baseline	Follow-up*	Difference	Effect Size
6MWT distance, m mean (SD)	397.9 (107.3)	367.4 (132.1)	−30.5 (119)	0.28
Criterion Referencing†				
	N	6MWT Distance, m‡		*P* value§
Hospitalization				
Yes	669	397.1 (107.2)		0.086
No	128	379.3 (111.1)		
Difference		17.9		
Death				
Yes	735	396.4 (107.3)		0.059
No	62	369.5 (112.4)		
Difference		26.9		
Hospitalization or death				
Yes	621	399.6 (106)		0.009
No	176	375.4 (112.8)		
Difference		24.2		

Definition of abbreviations: 6MWT = 6-minute-walk test; CI = confidence interval; IPF = idiopathic pulmonary fibrosis.

* Assessed at the Week 48 study visit.

† Comparison of baseline 6MWT distance between patients who did and did not experience selected health events during the subsequent 48-week period.

‡ Data are presented as mean (SD).

§ Independent samples *t* test.

6MWT measures a clinically important domain of the disease that is not captured by these other measures. Although several studies have identified change in percent predicted FVC as a strong independent predictor of mortality in patients with IPF (3, 35–37), no previous study assessing independent predictors of mortality in patients with IPF has included change in 6MWD as a covariate. Our data suggest that such information might provide important incremental prognostic information. Future studies using multivariate analysis are necessary to assess the contribution of changes in 6MWD to overall prognostic accuracy in patients with IPF.

Consistent with the recommendations of Yost and Eton (33), we used a number of alternative methods to estimate the MCID for the 6MWT, including distribution-based and anchor-based approaches. Based on these methods, our results suggest that the MCID for the 6MWT is between 24 and 45 m. This observation is consistent with the findings of two previous studies in patients with IPF (16, 18). In a study of 48 patients with diffuse parenchymal lung disease, half of whom had IPF, Holland and coworkers (16) reported a MCID ranging from 29 to 34 m. In a subsequent analysis of 6MWD among 123 patients with IPF randomized to the placebo arm of a clinical trial evaluating treatment with bosentan, Swigris and coworkers (18) estimated that the minimum important difference was 28 m (range, 10.8–58.5). Similar values have been reported in patients with COPD. Based on a single methodologic approach, Redelmeier and coworkers (38) reported an estimated minimum important difference of 54 m. A more recent analysis used data from nine trials among patients with COPD and used a variety of methodologic approaches; the MCID ranged from 29–42 m, with the estimate based on the SEM (35 m) favored by the authors (39). Thus, our finding that the MCID for the 6MWT is between 24 and 45 m is consistent with prior research focusing on the use of this instrument in patients with COPD, IPF, and other diffuse parenchymal lung diseases.

Several study limitations should be noted. First, although our analyses of responsiveness would have ideally been limited to patients randomized to placebo in the clinical trial, we concluded based on the absence of evidence for any treatment effect in the clinical trial that the enhanced power of the study to characterize the relationship between 6MWT and mortality justified the inclusion of all randomized patients. Analyses of responsiveness were robust when focusing on different populations (e.g., all subjects, subjects not requiring supplemental oxygen, placebo subjects) and when assessing changes over different periods of time (e.g., 24 wk vs. 48 wk). Second, we did not have access to data on patients' perception of their general well-being, and thus could not use the patient-referencing approach, which is another anchor-based method often used to assess MCID. Additionally, in our criterion-referencing approach, we used hospitalization and death as health events to estimate the MCID. These events are not "minimally important" in nature; therefore, the resulting MCID might overestimate the true smallest differences that would be clinically important to patients with IPF. Estimates of the MCID based on this approach may thus be viewed as conservative, although they were largely consistent with estimates yielded by the other approaches. Finally, patients with severe physiologic impairment were excluded from enrollment in the original clinical trial; therefore, the extent to which our findings are generalizable to patients with severe functional impairment is uncertain.

In conclusion, our results demonstrate that the 6MWT is a reliable, valid, and responsive measure of exercise tolerance in patients with IPF, and that a decline in 6MWD of 24–45 m represents a small but clinically important difference. These findings demonstrate that the 6MWT is a clinically useful measure of disease status and risk of mortality, and represents an attractive endpoint for clinical trials in patients with IPF. Future studies using multivariable analysis are necessary to assess the contribu-

tion of changes in 6WMT distance to overall prognostic accuracy in patients with IPF.

Author Disclosure: R.M.d.B. received consultancy fees from Cambridge Antibody Technology and Bayer; serves on the advisory board for Boehringer-Ingelheim, InterMune, Actelion, and Genzyme; and received lecture fees from InterMune, Actelion, GlaxoSmithKline, and AstraZeneca. D.W. received a sponsored grant from InterMune. C.A. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. W.Z.B. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. U.C. received consultancy fees from Centocor, Boehringer Ingelheim, Intermune, Actelion, and Bayer; served on the advisory board for Serono; received lecture fees from Astra Zeneca; and received sponsored grants from Actelion, Boehringer Ingelheim, Intermune, AstraZeneca, and Gilead. A.K. received a sponsored grant from InterMune. L.L. serves on the advisory board for Intermune and Actelion, and has received sponsored grants from Intermune, Actelion, Johnson and Johnson, and Gilead. P.W.N. received consultancy fees from Intermune, Actelion, Gilead, Novartis, and Boehringer Ingelheim. S.A.S. received lecture fees from Intermune and Boehringer-Ingelheim and sponsored grants from Actelion, Intermune, Boehringer Ingelheim, Novartis, AstraZeneca, Centocor, Sanovi-Aventis, Altana, and Wyeth. J.S. was a full-time employee of InterMune and Gilead. M.T. received consultancy fees from InterMune and expert witness fees from Guidepoint Global. D.V. serves on the advisory board for Actelion and received expert witness fees from Sanofi. T.E.K. received consultancy fees from Boehringer Ingelheim; serves on the advisory board for Actelion, InterMune, ImmuneWorks, and CV Therapeutics; and received lecture fees from Actelion.

INSPIRE Study Group: C. Agostini (Università degli Studi di Padova, Padova, Italy); C. Albera (Università di Torino, Orbassano, Italy); J. Allen (Ohio State University, Columbus); A. Anzueto (South Texas Veterans Healthcare System, San Antonio); J. Behr (Klinikum der Universität München-Großhadern, Munich, Germany); R. Bonnet (Klinik für Pneumologie, Zentralklinik Bad Berka, Bad Berka, Germany); R. Buhl (Johannes-Gutenberg Universität, Mainz, Germany); S. Burge (Birmingham Heartlands Hospital, Birmingham, UK); A. Chan (UC Davis Medical Center, Sacramento); C. Chan (Rosedale Medical Center, Toronto, Canada); P. Chanez (CHU Hôpital Arnaud de Villeneuve, Montpellier, France); J. Chapman (Cleveland Clinic Foundation, Cleveland); J. Cordier (Hôpital Louis Pradel, Bron, France); U. Costabel (Ruhrlandklinik, University of Duisburg/Essen Essen, Germany); H. Covelli (Pulmonary Consultants of North Idaho, Coeur d'Alene); N. Crimi (Azienda Ospedaliera Garibaldi, S. Luigi-Curro, Ascoli-Tomaselli, Catania, Italy); J. de Andrade (University of Alabama at Birmingham); P. Delaval (CHU Hôpital Pontchchaillou, Rennes, France); C. Dromer (Hôpital du Haut Lévèque, Pessac, France); J. Egan (Mater Misericordiae Hospital, Dublin, Ireland); R. Enelow (Yale University School of Medicine, New Haven); N. Ettinger (Cardiopulmonary Associates, Chesterfield); K. Flaherty (University of Michigan, Ann Arbor); A. Floreani (University of Nebraska, Omaha); S. Frankel (National Jewish Medical and Research Center, Denver, CO); A. Frost (Baylor College of Medicine, Houston); K. Gibson (University of Pittsburgh, Pittsburgh); M. Glassberg (University of Miami School of Medicine, Miami); M. Gottfried (Pulmonary Associates, Phoenix); S. Harari (Ospedale San Giuseppe, Milan, Italy); D. Helmersen (Peter Lougheed Center, Calgary, Canada); H. Hollingsworth (Boston University School of Medicine, Boston); M. Horton (Johns Hopkins School of Medicine, Baltimore); J. Jennings (Henry Ford Hospital, Detroit); M. Kallay (Highlands Hospital, Rochester); J. Lasky (Tulane University, New Orleans); A. Lee (Mayo Clinic, Jacksonville); C. Leonard (Wythenshawe Hospital, Manchester, UK); D. Lorch (PAB Clinical Research, Brandon); J. Lynch (Washington University School of Medicine, St. Louis); Y. Mageto (University of Texas, Southwestern, Dallas); S. Mette (Maine Medical Center, Portland); A. Millar (Southmead Hospital, Bristol, UK); F. Morell Brotad (Hospital General Vall d'Hebrón, Barcelona, Spain); J. Müller-Quernheim (Albert-Ludwigs-Universität, Freiburg, Germany); S. Nathan (Inova Research Center, Falls Church); I. Noth (University of Chicago, Chicago); M. Padilla (Mount Sinai Medical Center, New York); V. Poletti (Azienda Sanitaria di Forli, Forli, Italy); G. Raghu (University of Washington, Seattle); L. Richieldi (Azienda Policlinico di Modena, Modena, Italy); M. Robbins (University of Virginia, Charlottesville); J. Rolf (Kelowna General Hospital, Kelowna, Canada); J. Roman (Emory University School of Medicine, Atlanta); G. Rosen (Stanford University, Stanford); P. Rottoli (Azienda Ospedaliera Universitaria Senese, Sienna, Italy); C. Saltini (Azienda Ospedaliera Universitaria, Rome, Italy); T. Schaberg (Center of Pneumology, Rotenburg, Germany); T. Schaumberg (The Oregon Clinic, Portland); M. Scholand (University of Utah, Salt Lake City); N. Schönfeld (HELIOS Klinikum Emil von Behring, Berlin, Germany); S. Sharma (St. Boniface General Hospital, Winnipeg, Canada); P. Simonelli (New York Presbyterian Hospital, New York); M. Steele (Duke University, Durham); R. Sussman (Pulmonary and Allergy Associates, Summit); M. Thomeer (University Hospital Gasthuisberg, Leuven, Belgium); G. Tino (University of Pennsylvania, Philadelphia); D. Valeyre (Hôpital Avicenne, Bobigny, France); C. Vogelmeier (Klinik für innere Medizin, Marburg, Germany); B. Wallaert (Hôpital Albert Calmette, Lille, France); A. Wells (Royal Brompton Hospital, London, UK); M. Wencel (Wichita Clinic, Wichita); L. Wesselius (Mayo Clinic, Scottsdale); T. Whelan (University of Minnesota, Minneapolis); P. Wilcox (St. Paul's Hospital, Vancouver, Canada); P. Wolters (University of California, San Francisco); A. Xaubet (Hospital Clinic Provincial, Barcelona, Spain); D. Zisman (University of California, Los Angeles).

Acknowledgment: This study was funded by InterMune, Inc. The authors are indebted to Kenneth Glasscock for medical writing and editorial assistance and to the participating staff members and patients at all of the study centers.

References

1. Bjoraker JA, Ryu JH, Edwin MK, Myers JL, Tazelaar HD, Schroeder DR, Offord KP. Prognostic significance of histopathologic subsets in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 1998;157: 199–203.
2. King TE Jr, Tooze JA, Schwarz MI, Brown K, Cherniack RM. Predicting survival in idiopathic pulmonary fibrosis: scoring system and survival model. *Am J Respir Crit Care Med* 2001;164:1171–1181.
3. Latsi PI, du Bois RM, Nicholson AG, Colby TV, Bisirtzoglou D, Nikolakopoulou A, Veeraraghavan S, Hansell DM, Wells AU. Fibrotic idiopathic interstitial pneumonia: the prognostic value of longitudinal functional trends. *Am J Respir Crit Care Med* 2003;168: 531–537.
4. Lipkin DP, Scriven AJ, Crake T, Poole-Wilson PA. Six-minute walking test for assessing exercise capacity in chronic heart failure. *Br Med J (Clin Res Ed)* 1986;292:653–655.
5. Cahalin LP, Mathier MA, Semigran MJ, Dec GW, DiSalvo TG. The six-minute walk test predicts peak oxygen uptake and survival in patients with advanced heart failure. *Chest* 1996;110:325–332.
6. Gomberg-Maitland M, Huo D, Benza RL, McLaughlin VV, Tapson VF, Barst RJ. Creation of a model comparing 6-minute walk test to metabolic equivalent in evaluating treatment effects in pulmonary arterial hypertension. *J Heart Lung Transplant* 2007;26:732–738.
7. Guyatt GH, Pugsley SO, Sullivan MJ, Thompson PJ, Berman L, Jones NL, Fallen EL, Taylor DW. Effect of encouragement on walking test performance. *Thorax* 1984;39:818–822.
8. Bernstein ML, Despars JA, Singh NP, Avalos K, Stansbury DW, Light RW. Reanalysis of the 12-minute walk in patients with chronic obstructive pulmonary disease. *Chest* 1994;105:163–167.
9. Baughman RP, Sparkman BK, Lower EE. Six-minute walk test and health status assessment in sarcoidosis. *Chest* 2007;132:207–213.
10. Solway S, Brooks D, Lacasse Y, Thomas S. A qualitative systematic overview of the measurement properties of functional walk tests used in the cardiorespiratory domain [review]. *Chest* 2001;119: 256–270.
11. Chetta A, Aiello M, Foresi A, Marangio E, D'Ippolito R, Castagnaro A, Olivieri D. 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;18:170–175.
12. Hallstrand TS, Boitano LJ, Johnson WC, Spada CA, Hayes JG, Raghu G. The timed walk test as a measure of severity and survival in idiopathic pulmonary fibrosis. *Eur Respir J* 2005;25:96–103.
13. Eaton T, Young P, Milne D, Wells AU. Six-minute walk, maximal exercise tests: reproducibility in fibrotic interstitial pneumonia. *Am J Respir Crit Care Med* 2005;171:1150–1157.
14. Lederer DJ, Arcasoy SM, Wilt JS, D'Ovidio F, Sonett JR, Kawut SM. Six-minute-walk distance predicts waiting list survival in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2006;174:659–664.
15. Lettieri CJ, Nathan SD, Browning RF, Barnett SD, Ahmad S, Shorr AF. The distance-saturation product predicts mortality in idiopathic pulmonary fibrosis. *Respir Med* 2006;100:1734–1741.
16. Holland AE, Hill CJ, Conron M, Munro P, McDonald CF. Small changes in six-minute walk distance are important in diffuse parenchymal lung disease. *Respir Med* 2009;103:1430–1435.
17. Caminati A, Bianchi A, Cassandro R, Mirenda MR, Harari S. Walking distance on 6-MWT is a prognostic factor in idiopathic pulmonary fibrosis. *Respir Med* 2009;103:117–123.
18. Swigris JJ, Wamboldt FS, Behr J, du Bois RM, King TE, Raghu G, Brown KK. Six-minute walk test in idiopathic pulmonary fibrosis: longitudinal changes and minimal important difference. *Thorax* 2010; 65:173–177.
19. Guyatt GH, Osoba D, Wu AW, Wyrich KW, Norman GR. Methods to explain the clinical significance of health measures. *Mayo Clin Proc* 2002;77:371–383.
20. Schunemann HJ, Guyatt GH. Commentary–goodbye M(C)ID! Hello MID, where do you come from? *Health Serv Res* 2005;40:593–597.
21. Man-Son-Hing M, Laupacis A, O'Rourke K, Molnar FJ, Mahon J, Chan KB, Wells G. Determination of the clinical importance of study results. *J Gen Intern Med* 2002;17:469–476.
22. du Bois RM, Albera C, Bradford WZ, Costabel U, Kartashov A, Noble PW, Szwarcberg J, Thomeer M, Valeyre D, Weycker D, *et al.* 6-minute walk test distance (6MWD) is a reliable, valid, and responsive outcome measure that predicts mortality in patients with IPF [abstract]. *Am J Respir Crit Care Med* 2010;181:A1103.
23. King TE, Albera C, Bradford WZ, Costabel U, Hormel P, Lancaster L, Noble PW, Sahn SA, Szwarcberg J, Thomeer M, *et al.* Effect of

interferon gamma-1b on survival in patients with idiopathic pulmonary fibrosis (INSPIRE): a multicentre, randomised, placebo-controlled trial. *Lancet* 2009;374:175–264.
24. American Thoracic Society, European Respiratory Society. Idiopathic pulmonary fibrosis: diagnosis and treatment. International consensus statement. *Am J Respir Crit Care Med* 2000;161:646–664.
25. American Thoracic Society, European Respiratory Society. American Thoracic Society/European Respiratory Society international multidisciplinary consensus classification of the idiopathic interstitial pneumonias. *Am J Respir Crit Care Med* 2002;165:277–304.
26. Eakin EG, Resnikoff PM, Prewitt LM, Ries AL, Kaplan RM. Validation of a new dyspnoea measure: the UCSD shortness of breath questionnaire. University of California, San Diego. *Chest* 1998;113:619–624.
27. Chang JA, Curtis JR, Patrick DL, Raghu G. Assessment of health-related quality of life in patients with interstitial lung disease. *Chest* 1999;116:1175–1182.
28. Cohen J. Statistical power analysis for behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
29. Wyrwich KW, Nienaber NA, Tierney WM, Wolinsky FD. Linking clinical relevance and statistical significance in evaluating intra-individual changes in health-related quality of life. *Med Care* 1999;37:469–478.
30. Wyrwich 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;52:861–873.
31. Nichol MB, Epstein JD. Separating gains and losses in health when calculating the minimum important difference for mapped utility values. *Qual Life Res* 2008;17:955–961.
32. Ries AL. Minimally clinically important difference for the UCSD Shortness of Breath Questionnaire, Borg Scale, and Visual Analog Scale. *COPD* 2005;2:105–110.
33. Yost KJ, Eton DT. Combining distribution- and anchor-based approaches to determine minimally important differences: the FACIT experience. *Eval Health Prof* 2005;28:172–191.
34. American Thoracic Society. ATS Statement. Guidelines for the six-minute walk test. *Am J Respir Crit Care Med* 2002;166:111–117.
35. Collard HR, King TE, Bartelson BB, Vourlekis JS, Schwarz MI, Brown KK. Changes in clinical and physiologic variables predict survival in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2003;168:538–542.
36. Flaherty KR, Mumford JA, Murray S, Kazerooni EA, Gross BH, Colby TV, Travis WD, Flint A, Toews GB, Lynch JP, *et al.* Prognostic implications of physiologic and radiographic changes in idiopathic interstitial pneumonia. *Am J Respir Crit Care Med* 2003;168:543–548.
37. Zappala CJ, Latsi PI, Nicholson AG, Colby TV, Cramer D, Renzoni EA, Hansell DM, du Bois RM, Wells AU. Marginal decline in FVC is associated with a poor outcome in idiopathic pulmonary fibrosis. *Eur Respir J* 2010;35:830–836.
38. Redelmeier DA, Bayoumi AM, Goldstein RS, Guyatt GH. Interpreting small differences in functional status: the six minute walk test in chronic lung disease patients. *Am J Respir Crit Care Med* 1997;155:1278–1282.
39. Puhan MA, Mador MJ, Held U, Goldstein R, Guyatt GH, Schünemann HJ. Interpretation of treatment changes in 6-minute walk distance in patients with COPD. *Eur Respir J* 2008;32:637–643.