

Short- and Long-Term Reliability of the 6-Minute Walk Test in People With Idiopathic Pulmonary Fibrosis

Anne E Holland PhD, Catherine J Hill PhD, Leona Dowman MSc PhD,
Ian Glaspole MD PhD, Nicole Goh MD PhD, Annemarie L Lee MPhysio PhD, and
Christine F McDonald MD PhD

BACKGROUND: The aim of this work was to investigate the short- and long-term test-retest reliability of the 6-min walk distance (6MWD), peak heart rate, and nadir oxygen desaturation in idiopathic pulmonary fibrosis (IPF). **METHODS:** A reliability study of 70 adults with IPF was undertaken within out-patient pulmonary rehabilitation programs at 2 tertiary hospitals. Participants completed 2 baseline 6-min walk tests using a standard protocol, with continuous measures of percutaneous S_{pO_2} and heart rate via pulse oximetry. The 6-min walk test was completed immediately following an intervention period and 6 months after. Reproducibility was assessed by intraclass correlation coefficient and Bland-Altman analysis. **RESULTS:** Participants with a mean \pm SD diffusing capacity of the lung for carbon monoxide of $48 \pm 14\%$ were included. The reliability of the 6MWD was high (intraclass correlation coefficient = 0.96) with a mean learning effect of 21 m (95% CI 12–30 m). The learning effect persisted at 8 weeks (mean 14 m, 95% CI 5–23 m) but not 6 months (mean 15 m, 95% CI –1 to 30 m). Using the best (greatest) 6MWD significantly reduced the proportion of participants who were classified as having a clinically important response to rehabilitation compared with using the first 6MWD (40% vs 54%, $P = .002$). Nadir S_{pO_2} was reproducible, with a mean difference of $0.7 \pm 2.2\%$, and limits of agreement of –4 to 5%. Peak heart rate was more variable, with mean difference 5 ± 9 beats/min and limits of agreement of –12 to 20 beats/min. **CONCLUSIONS:** The 6MWD is a reproducible measure of exercise capacity in people with IPF. Whereas the nadir S_{pO_2} may be accurately determined from one test, evaluating change in 6MWD with interventions may require 2 tests on each occasion. (ClinicalTrials.gov registration NCT0016828.) *Key words:* walk test; pulmonary fibrosis; reliability. [Respir Care 2018;63(8):994–1001. © 2018 Daedalus Enterprises]

Introduction

Idiopathic pulmonary fibrosis (IPF) is a chronic and progressive fibrosing parenchymal lung disease that oc-

curs primarily in older adults and has no known cause.¹ It is characterized by progressive dyspnea, with declining health-related quality of life and impaired functional ex-

Drs Holland and Lee are affiliated with the Department of Physiotherapy, and Drs Glaspole and Goh are affiliated with the Department of Allergy, Immunology, and Respiratory Medicine, Alfred Health, Melbourne, Australia. Drs Holland, Dowman, and Lee are affiliated with the Department of Physiotherapy, La Trobe University, Melbourne, Australia. Drs Holland, Hill, Goh, Lee, and McDonald are affiliated with the Institute for Breathing and Sleep, Drs Hill and Dowman are affiliated with the Department of Physiotherapy, and Drs Goh and McDonald are affiliated with the Department of Respiratory and Sleep Medicine, Austin Health, Melbourne, Australia. Dr Glaspole is affiliated with the Department of Medicine, and Dr Lee is affiliated with the Department of Physiotherapy, Monash University, Melbourne, Australia.

This work was supported by the Pulmonary Fibrosis Foundation and American Thoracic Society Foundation. The authors have disclosed no conflicts of interest.

Dr Holland presented part of this work at the American Thoracic Society Annual Scientific Meeting, held May 17–22, 2013, in Philadelphia, Pennsylvania.

Correspondence: Anne Holland PhD, La Trobe University Clinical School, Level 4, The Alfred Centre, 99 Commercial Road, Melbourne 3004, Australia. E-mail: A.Holland@alfred.org.au.

DOI: 10.4187/respcare.05875

ercise capacity.^{2,3} Exercise capacity is often measured using the 6-min walk test (6MWT), a practical tool that is applied in a wide range of patients with respiratory conditions.⁴ Measurement of the 6-min walk distance (6MWD) is an important predictor of prognosis and mortality in IPF.⁵⁻⁸ The 6MWD is commonly used to assess treatment responses in individual patients and to measure outcomes of clinical trials, including pharmacotherapy^{9,10} and pulmonary rehabilitation.¹¹⁻¹³

Preliminary reports have suggested that the 6MWD is a reliable measure in IPF.^{5,14} However, 2 studies that included 30 and 45 participants have reported a learning effect (11–18 m), with a greater 6MWD reached during a second test on the same day.^{14,15} The magnitude of this learning effect was less than that described for COPD (27 m)¹⁶ and in a cohort of subjects with mixed interstitial lung disease (41 m).¹⁷ Larger studies are required to confirm the magnitude and importance of the learning effect. However, to date (as of August 2017) there has been no evaluation of whether learning effect persists over time in patients who have previously performed the test.

Previous work in people with COPD illustrated that in participants who performed 2 tests at baseline, repetition of the 6MWT immediately following pulmonary rehabilitation was not required, with no learning effect between 2 tests.¹⁸ However, at 3 months following pulmonary rehabilitation, the learning effect had re-emerged.¹⁸ As a result, in clinical practice, many pulmonary rehabilitation programs perform 2 tests at baseline, but only one test at the end of the program. To date, the 6MWT protocol in studies of pulmonary rehabilitation in IPF has generally included 2 post-program tests¹⁹; it is not known whether this is necessary, with no previous study in this population. Conversely, in trials of pharmacologic treatments, generally only one 6MWT is performed at follow-up visits.²⁰ In addition, the 6MWD is commonly used to monitor change in functional exercise capacity over time,²¹ which is an important prognostic indicator.^{22,23} Given that the minimal important difference (MID) for 6MWD is 30 m,⁴ which could be similar in magnitude to the learning effect, it is important to establish whether two tests are required to be confident that observed change in 6MWD reflects real clinical improvement or decline. Alternatively, if 2 tests are not required at follow-up, this would significantly reduce the time burden associated with performing the 6MWT.

The 6MWT also provides information related to physiological responses to exercise, including heart rate and S_{pO_2} . Both peak heart rate measured during the 6MWT and nadir S_{pO_2} are useful indicators of prognosis in IPF.^{22,24,25} A precise measure of nadir S_{pO_2} is of greater prognostic value than observed desaturation to 88% or less.²⁶ Measurement of symptoms during the 6MWT is also recom-

QUICK LOOK

Current knowledge

The 6-minute walk distance (6MWD) is a predictor of mortality and morbidity in idiopathic pulmonary fibrosis (IPF) and is a common outcome measure of treatment response. The distance is a reliable measure, but the reliability of nadir percutaneous oxyhemoglobin saturation, peak heart rate, and symptoms is unknown. The magnitude of the learning effect of the 6MWD and its change over time in IPF is unknown.

What this paper contributes to our knowledge

The 6MWD is a reproducible measure of exercise capacity in people with IPF. However, evaluating the change in 6MWD with interventions may require 2 tests on each occasion of testing.

mended.⁴ However, the reliability of these additional measures obtained during 6MWT has never been established.

The aims of this study were to determine (1) the magnitude of the learning effect of the 6MWD and whether this persists over time and (2) the reliability of nadir S_{pO_2} , peak heart rate, and symptom scores obtained from the 6MWT.

Methods

Participants were a convenience sample taking part in 2 trials of pulmonary rehabilitation in interstitial lung disease at Alfred and Austin Health, with the protocol for the exercise training program standardized across sites for the trials.^{11,27} Exercise training consisted of 30 min of endurance (walking and cycling) and upper- and lower-limb strength training using functional tasks and free weights.^{11,27} Participants were included if they had a diagnosis of IPF according to definitions outlined in the International Consensus Statement¹ and were ambulant. Exclusion criteria were the presence of a non-IPF respiratory disease, comorbidities that were severe enough to preclude exercise training, or a recent history of syncope on exertion. Demographic data collected included measures of lung function (Medgraphics Elite Series devices, MGC Diagnostics Corporation, Saint Paul, Minnesota), including FVC and carbon monoxide diffusing capacity according to standardized guidelines,^{28,29} with percentage predicted values calculated^{30,31}; and right-ventricular systolic pressure derived from trans-thoracic echocardiogram. Data were collected at baseline assessment, at the end of the 8-week intervention period (end pulmonary rehabilitation), and at 6 month follow-up. During the intervention period and the follow-up

phase, the 6MWT was not practiced by participants. All participants gave written informed consent, and the study was approved by the Alfred, Austin, and La Trobe University human research ethics committees.

Two 6MWTs were performed on each occasion of assessment on the same day, with 30 min between each test.⁴ Most participants had previously performed 6MWTs as part of their routine monitoring. In addition, all clinicians undertaking this test ensured that the vital signs and symptoms of each participant had returned to baseline levels before commencing the second test. Each test was conducted indoors, using a 30-m straight track and a standard protocol.⁴ Participants were instructed to walk as far as possible, with standardized encouragement given at each minute and the distance recorded. Participants who used walking aids for daily mobility were permitted to use the devices during the test. Continuous monitoring of heart rate and S_{pO_2} was performed during the test using pulse oximetry (PalmSAT 2500, Nonin Medical, Plymouth, Minnesota) with peak heart rate and nadir S_{pO_2} recorded.³² No oxygen supplementation was used during the test. Perceived dyspnea and leg fatigue were recorded using the modified Borg scale³³ before and after each 6MWT. Chronotropic response reflects the ability of the heart to increase its rate commensurate with increased activity; in this study, the chronotropic response was determined by the difference between peak heart rate and resting heart rate.

Statistical analysis was performed using the statistical package SPSS 21.0 (SPSS, Chicago, Illinois). Data are described as mean \pm SD. The intraclass correlation coefficient (ICC type 2,1) was used to establish the reproducibility of the 6MWT, and the coefficient of variation was calculated. Values for ICC of <0.5 are indicative of poor reliability, values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.90 indicate good reliability, and values >0.90 indicate excellent reliability.³⁴ Bland and Altman plots were used to evaluate agreement between the first and second tests for walking distance, nadir S_{pO_2} , and peak heart rate.³⁵ The effect of repeating the 6MWT on the clinical interpretation of change following pulmonary rehabilitation was evaluated by examining the mean change in 6MWD calculated using (1) the best 6MWD at both baseline and follow-up, (2) the first 6MWD at both baseline and follow-up, and (3) the best 6MWD at baseline compared with the first 6MWD at follow-up. The proportion (percentage) of participants who achieved the MID of 30 m⁴ for each comparison was calculated. Logistic regression assessed determinant factors of a clinically important change (30 m) in walked distance between the first and second test at baseline. A P value $< .05$ was considered statistically significant.

Table 1. Participant Demographics

| Characteristics | Mean \pm SD | 95% CI | Range |
|------------------------------------|-----------------|-----------|---------|
| Age, y | 72 \pm 8 | 65–80 | 55–90 |
| FVC, % predicted | 76.1 \pm 19.2 | 61.5–90.5 | 40–124 |
| D_{LCO} , % predicted | 47.7 \pm 14.2 | 34.8–57.0 | 19–91 |
| RVSP, mm Hg | 35.1 \pm 12.4 | 27–42 | 16–71 |
| Best 6MWD at baseline, m | 403 \pm 136 | 333–499 | 114–768 |
| Best 6MWD at baseline, % predicted | 75 \pm 23 | 65–88 | 45–92 |

D_{LCO} = diffusing capacity of the lung for carbon monoxide

RVSP = right-ventricular systolic pressure on trans-thoracic echocardiogram

6MWD = 6-min walk distance

Results

A total of 70 participants were included (31 males), of whom 33 underwent pulmonary rehabilitation. Of the 70 participants, 42 were reassessed immediately following pulmonary rehabilitation at week 8, and 36 were assessed at 6 months follow-up. The demographics of participants are outlined in Table 1, with the baseline best 6MWD being 75% of normal.³⁶

Reliability of the 6MWT at Baseline

At baseline, participants walked a mean \pm SD distance of 378 \pm 127 m on the first test and 399 \pm 139 m on the second test. The average difference between the two tests was 21 m (95% CI 12–30 m, 18% change) with 69% of participants walking farther on the second test. The Bland-Altman plot showed limits of agreement of -54 to 96 m (Fig. 1A). The 6MWD was reproducible, with an ICC of 0.96 (Table 2).

On average, the nadir S_{pO_2} was 88 \pm 6% during the first test and 87 \pm 6% during the second test. The change in oxyhemoglobin saturation was reproducible (Table 2), with an ICC of 0.94, and the limits of agreement between the two 6MWTs were between -4 to 5% (Fig. 1B).

On average, the peak heart rate was 107 \pm 19 beats/min during the first test and 112 \pm 19 beats/min during the second test. This peak heart rate was reproducible (Table 2), with an ICC of 0.96 and limits of agreement of -12 to 20 beats/min (Fig. 1C). There was a difference in chronotropic response (peak heart rate minus resting heart rate) between test 1 and test 2 of 2.8 beats/min (95% CI 1–6 beats/min).

The difference in end Borg dyspnea measures was 0.1 points (95% CI -0.4 to 1.1 points), whereas the difference in leg fatigue scores post-test was -0.5 points (95% CI -1.0 to -0.2 points). The reproducibility of change in Borg dyspnea and leg fatigue scores was good to excellent (dyspnea ICC 0.80, leg fatigue ICC 0.94).

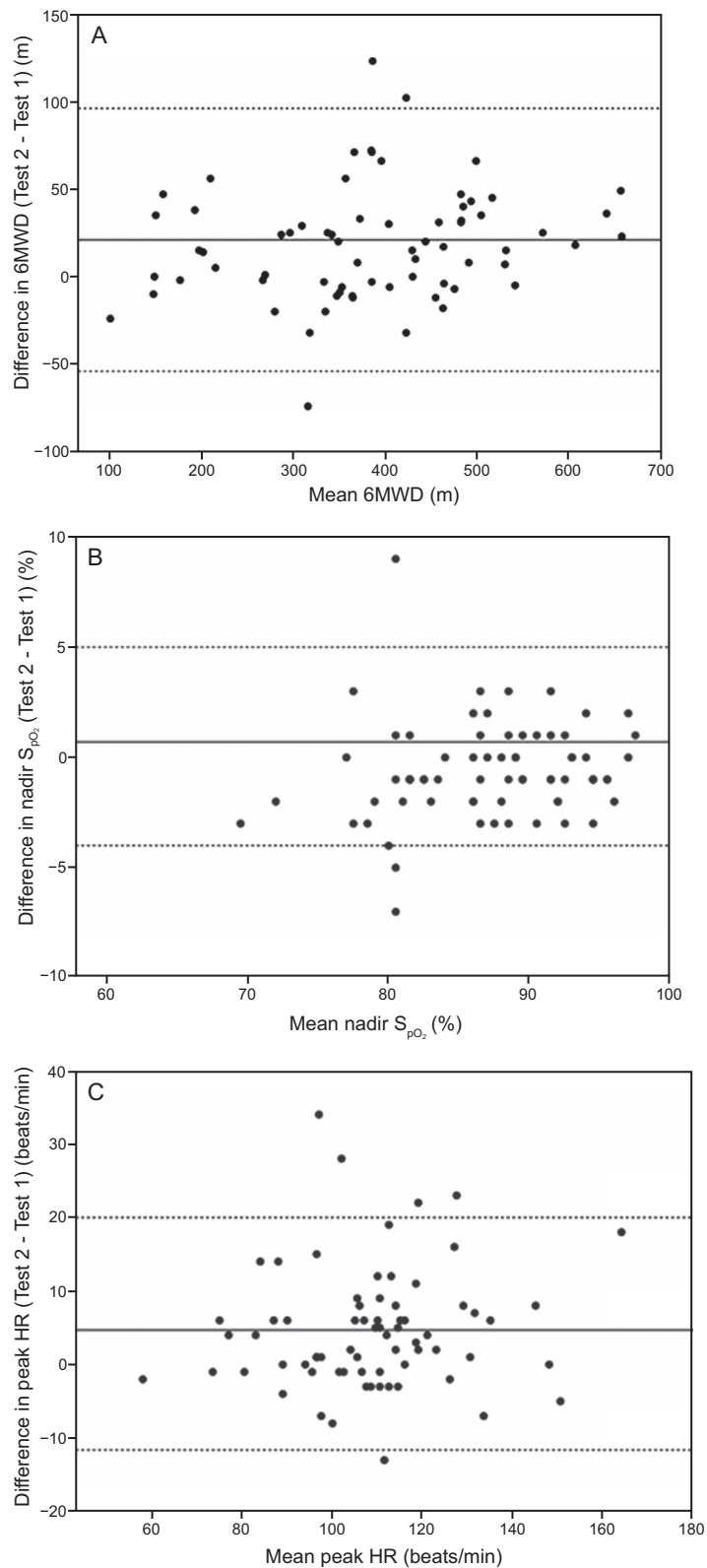


Fig. 1. Bland-Altman plot for baseline 6-min walk distance (6MWD) (A), nadir oxyhemoglobin saturation (S_{pO_2}) (B), and peak heart rate (HR) (C). The solid line represents mean differences, and dotted lines show the limits of agreement.

Table 2. Reproducibility of 6-Minute Walk Distance, Nadir S_{pO_2} , and Peak Heart Rate

| Parameters | 6MWD (m) | | | Nadir S_{pO_2} (%) | | | Peak Heart Rate (beats/min) | | |
|---------------------------|----------------------|---------------------|----------------------|----------------------|---------------------|----------------------|-----------------------------|---------------------|----------------------|
| | Baseline (n = 70) | 8 Weeks (n = 42) | 6 Months (n = 36) | Baseline (n = 70) | 8 Weeks (n = 42) | 6 Months (n = 36) | Baseline (n = 70) | 8 Weeks (n = 42) | 6 Months (n = 36) |
| Mean difference* \pm SD | 21 \pm 38 | 14 \pm 22 | 15 \pm 38 | 0.7 \pm 2.2 | -0.5 \pm 2.1 | -0.3 \pm 2.4 | 4.7 \pm 8.3 | -3.6 \pm 4.7 | -0.8 \pm 14 |
| Limits of agreement | -54 to 96 | -29 to 57 | -59 to 89 | -4 to 5 | -4 to 4 | -4 to 4 | -12 to 20 | -13 to 5 | -28 to 27 |
| ICC (95% CI) | 0.96 (0.93-0.98) | 0.98 (0.96-0.99) | 0.95 (0.88-0.98) | 0.94 (0.91-0.96) | 0.95 (0.89-0.98) | 0.90 (0.78-0.96) | 0.96 (0.90-0.97) | 0.96 (0.92-0.98) | 0.82 (0.59-0.92) |
| Coefficient of variation | 0.08 | 0.05 | 0.08 | 0.02 | 0.02 | 0.03 | 0.08 | 0.04 | 0.12 |

* Mean difference between 2 tests conducted on the same day.

6MWD = 6-min walk distance

ICC = intraclass correlation coefficient

Reliability of 6MWT After the Intervention Period and at 6 Months Follow-Up

After the intervention phase, there was a significant difference in 6MWD between test 1 and test 2 of 14 m (95% CI 5–23 m), with 81% of participants walking farther on the second test. The limits of agreement were narrower compared with baseline (Table 2). The mean difference in nadir S_{pO_2} was -0.5% (95% CI -1.3 to 0.3%), whereas the mean difference in peak heart rate was -3.6 beats/min (95% CI -5.0 to 0.8 beats/min). The average difference in chronotropic response was 5 beats/min (95% CI -4 to 14 beats/min), with a lower response on the second test. The difference in end Borg dyspnea score was -0.2 points (95% CI -0.7 to 0.3 points), whereas the difference in leg fatigue scores post-test was 0.2 points (95% CI -0.5 to 0.9 points). The reproducibility of change in Borg dyspnea and leg fatigue scores was moderate to excellent (dyspnea ICC 0.58, leg fatigue ICC 0.94). The impact of the intervention phase on 6MWD appeared larger if only the first 6MWD was considered at each time point (mean improvement 38 m vs 29 m using the best 6MWD at each time point). If only one test was performed at baseline and immediately following the intervention phase, then 54% of participants would be classed as achieving the MID, whereas only 40% of participants achieved the MID if 2 tests were performed at each time point, and the best results were compared (Table 3). In contrast, use of 2 tests at baseline and only one 6MWT at 8 weeks significantly reduced the mean change in 6MWD immediately following the intervention phase (13 m).

At 6 months there was a non-statistically significant difference in 6MWD between tests of 15 m (-1 to 30 m), with 64% of participants performing better on the second test. Good to excellent levels of reliability remained evident for 6MWD, nadir S_{pO_2} , and peak heart rate (Table 2). The difference in chronotropic response was 0 beats/min (95% CI -5 to 5 beats/min). The difference in end Borg dyspnea measures was 0.3 points (95% CI -0.3 to 0.8 points), whereas the difference in end leg fatigue was -0.1 points (95% CI -0.7 to 0.4 points). The reproducibility of Borg dyspnea and leg fatigue scores was moderate to excellent (dyspnea ICC 0.73, leg fatigue ICC 0.97). If only one test was performed at baseline and at 6 months follow-up, then 36% of participants would be classed as achieving the MID, whereas 40% of participants achieved the MID if 2 tests were performed at each time point, and the best results were compared (Table 3).

A model of logistic regression showed that a higher odds ratio for a clinically important improvement of ≥ 30 m in the second 6MWT in comparison with the first occurred in subjects with a lower dyspnea score at the end of the first 6MWT (odds ratio 0.601 [95% CI 0.399–0.905, $P = .02$]), and there was a trend for younger participants

Table 3. Impact of Using One or 2 Tests on the Estimate of Change Over Time

| Follow-Up Period | Best 6MWD at Follow-Up – Best 6MWD Baseline | First 6MWD at Follow-Up – First 6MWD Baseline | First 6MWD at Follow-Up – Best 6MWD Baseline |
|--------------------------------------|---|---|--|
| End of intervention period (8 weeks) | | | |
| Change in 6MWD, mean (95% CI) m | 29 (11–47) | 38 (23–55) | 13 (–5 to 31)* |
| Proportion achieving MID, % | 40 | 54 | 36 |
| 6 months after intervention period | | | |
| Change in 6MWD, mean (95% CI) m | 10 (–15 to 34) | 16 (–8 to 40) | –11 (–34 to 11) |
| Proportion achieving MID, % | 40 | 36 | 24 |

Mean change is from baseline 6MWD.

* Difference between first 6MWD at follow-up – best 6MWD at baseline versus best 6MWD at follow up – best 6MWD baseline, $P < .05$.

6MWD = 6-min walk distance

MID = minimum important difference

being more likely to have improvement exceeding the MID on their second test (odds ratio 0.94 [95% CI 0.87–1.011, $P = .10$]).

Discussion

This study shows that the 6MWT provides reliable measures of functional walking capacity, peak heart rate, and nadir S_{pO_2} in subjects with IPF. The variability in peak heart rate was greater than the variability in nadir S_{pO_2} . This is the first report with longitudinal data that demonstrates persistence of the learning effect for the 6MWD in this patient population. Although this diminished over time, it had significant impact on the interpretation of change following an intervention period. As a result, a second test may be necessary to confidently determine whether an intervention phase (which may include pulmonary rehabilitation) has resulted in a real clinical change. One test may be sufficient if measuring change over longer time periods, in the absence of an intervention.

The 6MWD is reproducible in IPF, with measures of reliability similar to previous reports (ICCs ranging from 0.83 to 0.98).^{5,14} A learning effect was evident at initial assessment, the magnitude of which was greater than that previously described in studies of subjects with IPF.^{14,15} The larger learning effect documented here may be due to a greater number of subjects included, reflecting the broader range of disease severity and functional performance seen in IPF. The magnitude of the learning effect and the limits of agreement are in a range similar to that of the MID for the 6MWD.⁴ This, together with the fact that 81% of subjects improved on their second test, emphasizes the clinical relevance of this learning effect and suggests that to accurately assess functional exercise capacity using the 6MWD, such as before commencement of a pulmonary rehabilitation program to ensure accurate exercise prescription, it is preferable that 2 tests be performed. Reporting less dyspnea at the end of the first 6MWT was a signifi-

cant determinant of a change in 6MWT in the second test being >30 m, which offers some guidance when applying these tests in clinical practice.

The learning effect for 6MWD persists over time, with increased distance on the second test still evident following either 8 weeks of rehabilitation or no intervention. This finding has novel and significant implications for the conduct of both clinical trials and clinical practice that apply the 6MWT as an outcome measure in individuals with IPF. A recent review article²⁰ highlighted that in trials of new pharmacologic agents for IPF, the 6MWT is generally performed only once at follow-up visits. Whereas the magnitude of the learning effect and its limits of agreement were smaller at the end of an intervention period than at the beginning, performing only one test on each occasion increases the risk of overestimating treatment effects, because the effects of the intervention phase cannot be distinguished from the effects of learning (Table 3). Conversely, the common clinical practice of performing 2 baseline tests and comparing this with a single post-intervention-phase 6MWD risks underestimating the treatment effect (Table 3). Therefore, with this consistent presence of a learning effect at baseline and immediately following an intervention period, it is important that 2 tests be performed on both occasions, with the best of the two tests used, rather than a priori selecting the first or the second test.

These results contrast with the only previous study that examined the learning effect of the 6MWT immediately following pulmonary rehabilitation in people with COPD.¹⁸ In this study, the learning effect was no longer significant following 8 weeks of pulmonary rehabilitation.¹⁸ More significantly, this directly challenges the technical standard for this test, which indicates that one test is sufficient at follow-up for people with chronic respiratory disease.⁴ Familiarity with the 6MWT is associated with motivation,³⁷ increased confidence, and reduced anxiety.³⁸ Whereas exertional symptoms improve following pulmo-

nary rehabilitation in IPF,^{11,13} this improvement did not appear to impact the learning effect at the end of rehabilitation in this group.

Exercise capacity declined over 6 months, a finding similar to those of previous studies.^{12,20} Whereas the test remained highly reproducible at 6 months, the learning effect was still evident, similar to findings in COPD.¹⁸ However, the magnitude was reduced and no longer statistically significant, and as a result, we found that average change in 6MWD from baseline was similar when using either a single test or the longer of 2 tests. With the wide use of the 6MWD to determine disease prognosis and mortality in IPF,^{7,8,14,23,25} it is important that an accurate assessment be completed. The results of this first longitudinal study in IPF suggest that when the 6MWD is used for risk stratification or to monitor disease course, in the absence of an intervention, a single test may be acceptable. As the tolerance of the 6MWT is variable in people with IPF,³⁹ this knowledge directly informs clinical practice and research studies with long-term follow-up using this outcome.

The nadir S_{pO_2} is highly reproducible, with the reliability and limits of agreement similar to those reported in COPD.^{16,18} This contrasts with previous studies in subjects with IPF, in which considerable variation for $S_{pO_2} \leq 88\%$ or a decline $\geq 4\%$ was evident.¹⁴ This may be due to the differing protocols for measuring S_{pO_2} during the test. Eaton et al¹⁴ measured S_{pO_2} at the beginning and end of the 6MWT, whereas in the current study, S_{pO_2} was monitored continuously. We have previously shown that end test S_{pO_2} and nadir S_{pO_2} are not the same in subjects with interstitial lung disease, including those with IPF.³² The findings of this study suggest that if the purpose of the walking test performed clinically is to determine the level of desaturation, and if continuous monitoring of oxygenation is included, only one 6MWT is required.

This is the first study to measure the reproducibility of the heart rate response during the 6MWT in IPF. Although the peak heart rate is reliable, variation within individuals is evident with wide limits of agreement between tests conducted on the same day. Variability in heart rate response during a 6MWT has been demonstrated previously in COPD.^{16,18} The importance of a reliable estimate of heart rate response during the 6MWT is illustrated by the finding that heart rate recovery ≤ 13 beats/min in the first minute following the test is a predictor of mortality in IPF.⁴⁰ We have also reported that an increase in heart rate of < 20 beats/min during a 6MWT has prognostic significance.²⁴ The relevance of the variation in heart rate responses between repeated tests may therefore depend on the threshold being used.

There are some limitations to this work. Although this is the largest study examining the reliability of the 6MWD in IPF, it included a relatively modest number of participants

recruited as a convenience sample, with fewer participants at the end of rehabilitation and at 6 months follow-up. In addition, a smaller proportion of participants originally allocated to the control group (no pulmonary rehabilitation) were available at both follow-up assessment times. With these small control participant numbers, it is difficult to comment on whether the reproducibility of the 6MWD differs between those undertaking rehabilitation and those who did not (and therefore whether rehabilitation influenced these findings). Further study is required to clarify this area. We examined the reliability of oxygen saturation and heart rate measures obtained during pulse oximetry, reflecting the measurements obtained in clinical practice. It is possible that heart rate measures may have been more reliable using other methods, such as a frequency meter; however, to date, no studies have made such a comparison. Finally, we included participants with IPF only. As a result, these results cannot be extrapolated to those with other types of interstitial lung disease, including those with connective tissue disease who may have musculoskeletal and circulatory manifestations that could affect the reliability of the test.

Conclusions

Measures of 6MWD, peak heart rate, and degree of oxygen desaturation can be reliably obtained from a 6MWT in people with IPF. A clear learning effect for 6MWD persists over short-term follow-up; this has significant implications for the accurate assessment of clinically important change with intervention. As a result, 2 6MWTs are required when the 6MWD is used to evaluate treatment effects, such as following pulmonary rehabilitation. One test may be sufficient to obtain a reliable measure of nadir oxyhemoglobin saturation.

REFERENCES

1. Raghu G, Collard HR, Egan JJ, Martinez FJ, Behr J, Brown KK, 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(6):788-824.
2. Swigris JJ, Gould MK, Wilson SR. Health-related quality of life among patients with idiopathic pulmonary fibrosis. *Chest* 2005; 127(1):284-294.
3. Nishiyama O, Taniguchi H, Kondoh Y, Kimura T, Ogawa T, Watanabe F, Nishimura K. Health-related quality of life in patients with idiopathic pulmonary fibrosis: what is the main contributing factor? *Respir Med* 2005;99(4):408-414.
4. Holland AE, Spruit MA, Troosters T, Puhan MA, Pepin V, Saey D, et al. An official European Respiratory Society/American Thoracic Society Technical Standard: field walking tests in chronic respiratory disease. *Eur Respir J* 2014;44(6):1428-1446.
5. du Bois RM, Weycker D, Albera C, Bradford WZ, Costabel U, Kartashov A, 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(9):1231-1237.

6. 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(1):96-103.
7. 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(1):117-123.
8. 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(6):659-664.
9. King TE Jr, Behr J, Brown KK, du Bois RM, Lancaster L, de Andrade JA, et al. BUILD-1: a randomized placebo-controlled trial of bosentan in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2008;177(1):75-81.
10. Noble PW, Albera C, Bradford WZ, Costabel U, Glassberg MK, Kardatzke D, et al. Pirfenidone in patients with idiopathic pulmonary fibrosis (CAPACITY): two randomised trials. *Lancet* 2011;377(9779):1760-1769.
11. Holland AE, Hill CJ, Conron M, Munro P, McDonald CF. Short term improvement in exercise capacity and symptoms following exercise training in interstitial lung disease. *Thorax* 2008;63(6):549-554.
12. Kozu R, Jenkins S, Senjyu H. Effect of disability level on response to pulmonary rehabilitation in patients with idiopathic pulmonary fibrosis. *Respirology* 2011;16(8):1196-1202.
13. Nishiyama O, Kondoh Y, Kimura T, Kato K, Kataoka K, Ogawa T, et al. Effects of pulmonary rehabilitation in patients with idiopathic pulmonary fibrosis. *Respirology* 2008;13(3):394-399.
14. 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(10):1150-1157.
15. Kozu R, Jenkins S, Senjyu H, Mukae H, Sakamoto N, Kohn S. Peak power estimated from 6-minute walk distance in Asian patients with idiopathic pulmonary fibrosis and chronic obstructive pulmonary disease. *Respirology* 2010;15(4):706-713.
16. Hernandez NA, Wouters EF, Meijer K, Annegarn J, Pitta F, Spruit MA. Reproducibility of 6-minute walking test in patients with COPD. *Eur Respir J* 2011;38(2):261-267.
17. Jenkins S, Cecins NM. Six-minute walk test in pulmonary rehabilitation: do all patients need a practice test? *Respirology* 2010;15(8):1192-1196.
18. Spencer LM, Alison JA, McKeough ZJ. Six-minute walk test as an outcome measure: are two six-minute walk tests necessary immediately after pulmonary rehabilitation and at three-month follow-up? *Am J Phys Med Rehabil* 2008;87(3):224-228.
19. Dowman L, Hill CJ, Holland AE. Pulmonary rehabilitation for interstitial lung disease. *Cochrane Database Syst Rev* 2014;(10):CD006322.
20. Brown AW, Nathan SD. The value and application of the 6-minute-walk test in idiopathic pulmonary fibrosis. *Ann Am Thorac Soc* 2018;15(1):3-10.
21. Tomioka H, Imanaka K, Hashimoto K, Iwasaki H. Health-related quality of life in patients with idiopathic pulmonary fibrosis—cross-sectional and longitudinal study. *Intern Med* 2007;46(18):1533-1542.
22. Collard HR, King TE Jr, 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(5):538-542.
23. Flaherty KR, Andrei AC, Murray S, Fraley C, Colby TV, Travis WD, et al. Idiopathic pulmonary fibrosis: prognostic value of changes in physiology and six-minute-walk test. *Am J Respir Crit Care Med* 2006;174(7):803-809.
24. Holland AE, Hill CJ, Glaspole I, Goh N, Dowman L, McDonald CF. Impaired chronotropic response to 6-min walk test and reduced survival in interstitial lung disease. *Respir Med* 2013;107(7):1066-1072.
25. 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(10):1734-1741.
26. Lama VN, Flaherty KR, Toews GB, Colby TV, Travis WD, Long Q, et al. Prognostic value of desaturation during a 6-minute walk test in idiopathic interstitial pneumonia. *Am J Respir Crit Care Med* 2003;168(9):1084-1090.
27. Dowman LM, McDonald CF, Hill CJ, Lee AL, Barker K, Glaspole I, et al. The evidence of benefits of exercise training in interstitial lung disease: a randomised controlled trial. *Thorax* 2017;72(7):610-619.
28. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, et al. Standardisation of spirometry. *Eur Respir J* 2005;26(2):319-338.
29. Macintyre N, Crapo RO, Viegi G, Johnson DC, van der Grinten CP, Brusasco V, et al. Standardisation of the single-breath determination of carbon monoxide uptake in the lung. *Eur Respir J* 2005;26(4):720-735.
30. Knudson RJ, Slatin RC, Lebowitz MD, Burrows B. The maximal expiratory-flow volume curve: normal standards, variability and effects of age. *Am Rev Respir Dis* 1976;113(5):587-600.
31. Roca J, Rodriguez-Roisin R, Cobo E, Burgos F, Perez J, Clausen JL. Single-breath carbon monoxide diffusing capacity prediction equations from a Mediterranean population. *Am Rev Respir Dis* 1990;141(4 Pt 1):1026-1032.
32. Fiore C, Lee A, McDonald C, Hill C, Holland A. Should oxyhaemoglobin saturation be monitored continuously during the 6-minute walk test? *Chron Respir Dis* 2011;8(3):181-184.
33. Borg G. Psycho physical bases of perceived exertion. *Med Sci Sports Exerc* 1982;14(5):377-381.
34. Portney LG, Watkins MP. Foundations of clinical research: applications to practice. New Jersey: Prentice Hall; 2000:585-619.
35. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;327(8476):307-310.
36. Jenkins S, Cecins N, Camarri B, Williams C, Thompson P, Eastwood P. Regression equations to predict 6-minute walk distance in middle-aged and elderly adults. *Physiother Theory Pract* 2009;25(7):516-522.
37. Eiser N, Willsher D, Doré CJ. Reliability, repeatability and sensitivity to change of externally and self-paced walking tests in COPD patients. *Respir Med* 2003;97(4):407-414.
38. Stevens D, Elpern E, Sharma K, Szidon P, Ankin M, Kesten S. Comparison of hallway and treadmill six-minute walk tests. *Am J Respir Crit Care Med* 1999;160(5 Pt 1):1540-1543.
39. Holland AE, Dowman L, Fiore J Jr, Brazzale D, Hill CJ, McDonald CF. Cardiorespiratory responses to 6-minute walk test in interstitial lung disease: not always a submaximal test. *BMC Pulm Med* 2014;14:136.
40. Swigris JJ, Swick J, Wamboldt FS, Sprunger D, du Bois R, Fischer A, et al. Heart rate recovery after 6-min walk test predicts survival in patients with idiopathic pulmonary fibrosis. *Chest* 2009;136(3):841-848.