The minimum clinically important improvement for the Incremental Shuttle Walking Test

S J Singh†*, P W Jones**, R Evans*, M D L Morgan*

Corresponding Author:

Sally Singh
*Pulmonary Rehabilitation Research Group,
Department of Respiratory Medicine and Thoracic Surgery,
Glenfield Hospital,
University Hospitals of Leicester NHS Trust,
Groby Road,
Leicester,
United Kingdom.
LE3 9QP

And

†Faculty of Health and Life Sciences
Coventry University
Priory St
CV1 5FB

Email: sally.singh@uhl-tr.nhs.uk
Tel: 00 44 116 256 3181
Fax: 00 44 116 236 7768

**St Georges, University of London
Cranmer Terrace
London
United Kingdom
SW17 0RE
Abstract

Background - The Incremental Shuttle Walking Test (ISWT) is used to assess exercise capacity in chronic obstructive pulmonary disease (COPD) and is employed as an outcome measure for pulmonary rehabilitation. This study was designed to establish the minimum clinically important difference (MCID) for the ISWT.

Methods - 372 patients (205 male) performed an ISWT before and after a 7 week outpatient pulmonary rehabilitation programme. After completing the course subjects were asked to identify, from a 5-point Likert scale, the perceived change in their exercise performance immediately upon completion of the ISWT. The scale ranged from 'better' to 'worse'.

Results - The mean (SD) age was 69.4 (8.4) years, FEV1 1.06 (0.53) litres, and FEV1/FVC ratio 50.8 (18.1) %. Baseline shuttle walking test distance was 168.5(14.6) m. After rehabilitation this increased to 234.7 (125.3) m. The mean difference and 95% confidence interval for the difference was 65.9 (58.9 to72.9) m. For a subject to feel they were ‘slightly better’ the mean improvement was 47.5 (38.6 to 56.5) m and to report feeling ‘better’ the mean increase was 78.7 (70.5 to 86.9) m. Patients who reported that their exercise tolerance was ‘about the same’ increased by a mean of 18.0 (4.5 – 31.5) m.

Conclusion - Two levels of improvement were identified. The minimum clinically important improvement for the ISWT is 47.5m. In addition patients were able to distinguish an additional benefit at 78.7m.

Introduction

Improving exercise performance is an important therapeutic goal for patients with Chronic Obstructive Pulmonary Disease (COPD) and other chronic respiratory diseases (1). In this context a laboratory exercise test may provide a precise physiological outcome but field walking tests are often employed as a pragmatic alternative. The most popular field walking tests are the unpaced six-minute walking distance (6MWD) and the incremental shuttle walk test (ISWT). The ISWT has been used in an increasing number of studies, because the externally paced incremental format is similar to the laboratory incremental exercise test (1). The ISWT has been used as an outcome measure in trials of pulmonary rehabilitation (2;3), pharmaceutical studies (4) and other disease populations as a prognostic marker (5).

In the context of a therapeutic trial, laboratory or field exercise tests are generally used as a proxy outcome that may reflect domestic functional performance. Any positive changes in exercise performance must therefore be perceived as beneficial by the patient. However the small statistically significant improvements reported in trials may not necessarily translate to useful recognised benefit. The concept of the minimal clinically important difference (MCID) has therefore been
developed to demonstrate the threshold of change that may have some meaning for the patient. The MCID has been defined as the smallest difference in a score that patients perceive to be beneficial in the absence of troublesome side effects and excessive costs’ (6).

The minimal clinically important differences have been identified for a number of common disease specific health status questionnaires, including the Chronic Respiratory Questionnaire (CRQ) (7) the St George’s Questionnaire (SGRQ) (8) and the Baseline and Transitional Dyspnoea Questionnaire (BDI/TDI). Less information is available for interpreting improvements in functional performance status. A MCID value for the six-minute walking test has been identified as 54m (9) but no MCID has yet been developed for the ISWT. The aim of this study was to calculate a MCID for incremental shuttle-walking test using pulmonary rehabilitation as the therapeutic intervention. A secondary aim of the study was to test whether more than one level of improvement could be distinguished by the patient following rehabilitation.

**Method**

All patients had a confirmed diagnosis of COPD and were appropriate referred to the out-patient based pulmonary rehabilitation programme. Prior to commencing the course patients were assessed by a respiratory physician. Patients were excluded from rehabilitation if there was significant locomotor, neurological or psychiatric limitation to exercise. Subjects were invited for an initial assessment. Exercise tolerance was assessed with the incremental shuttle walking test (10). At this initial assessment subjects completed two shuttle incremental shuttle walking tests with adequate rest of 20-30 minutes between the two tests. After assessment patients attended the hospital out-patient rehabilitation programme. The duration was seven weeks, involving two hospital outpatient visits per week each lasting two hours (one hour of physical training and one hour of education). A daily home walk was also prescribed at a speed representing 85% VO₂ peak, derived from the ISWT performance (11). In addition to brisk walking the subjects performed peripheral muscle training exercises three times a week with free weights, twice at home and once at hospital. They kept a diary of their exercises to allow the rehabilitation team to monitor compliance and progress. After seven weeks of rehabilitation the incremental shuttle-walking test was repeated. At this assessment a question was asked immediately upon completion of the test asking the subject to rate how much they felt their exercise tolerance had changed using the following question, Compared to last time how would you rate your exercise tolerance?

- Better □
- Slightly better □
- About the same □
- Slightly worse □
- Worse □

The score was assigned a numerical value from 1-5 according to descriptors. At this stage the subjects were not informed of any objective change in distance walked after the course of rehabilitation. Local ethical committee approval was granted to collect this additional data at the time of the follow-up assessment.
Statistical analysis
Data were analysed using the SPSS (v14). Baseline variables were normally distributed. To estimate the difference in the ISWT to within a precision of plus or minus 15m (as represented by the 95% confidence interval) and assuming a standard deviation of 36.7m (generated from previously published data (12) assuming equal group sizes then 46 patients would be needed per group. The mean change in ISWT distance achieved by pulmonary rehabilitation for each response of the simple question was calculated with mean and 95% CI. Since the data could be analysed in terms of categorical data (number of shuttles) or as a continuous variable (distance walked), both parametric and non-parametric analyses were carried out.
Results

Data from 372 patients completing rehabilitation are reported. Data collection was continued until at least 46 patients had been recruited into response category 1, 2 or 3. The baseline demographics were mean (SD) age 69.4 (8.4) years, FEV1 1.06 (0.53) litres, FEV1/FVC ratio was 50.8(18.1) % and 55% were male. The mean baseline shuttle distance was 168.5 (14.6) m. After rehabilitation this increased to 234.7 (125.3) m. The mean improvement was 65.9 m (95% confidence interval 58.9 to 72.9 m). As anticipated there was no relationship between baseline ISWT performance and the improvement in shuttle distance following rehabilitation. Figure 1 is a Bland and Altman plot for the 5 groups identifying the mean baseline/post ISWT against the ISWT difference.

The distribution of responses to question about perceived improvement was: “better” 50.5% (n=188), “slightly better”, 29.9% (n=111), “about the same” 14.8% (n=55), “slightly worse” 4.3% (n=16) and “worse” 0.5% (n=2). One-way analysis of variance (ANOVA) between all five groups identified a significant difference in mean distance achieved (p<0.0001), post hoc analysis identified the differences were between groups 1, 2, 3 and 4. However the numbers of patients who felt their exercise tolerance was “slightly worse” or “worse” were too small to give statistically significant results, so they were excluded from subsequent analyses. The baseline characteristics of the three groups are identified in table 1.

Table 1

<table>
<thead>
<tr>
<th></th>
<th>Better Mean(SD)</th>
<th>Slightly better Mean(SD)</th>
<th>About the same Mean(SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (male: female)</td>
<td>103:85</td>
<td>61:50</td>
<td>34:21</td>
</tr>
<tr>
<td>Age (yrs)</td>
<td>67.9 (8.3)</td>
<td>70.4 (7.7)</td>
<td>70.3 (7.3)</td>
</tr>
<tr>
<td>FEV1 (% predicted)</td>
<td>42.1(24.3)</td>
<td>43.0 (23.6)</td>
<td>45.6 (22.6)</td>
</tr>
<tr>
<td>FEV1/FVC</td>
<td>51.0 (17.2)</td>
<td>50.9 (18.0)</td>
<td>53.9 (23.7)</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>166.4 (9.4)</td>
<td>164.2 (9.8)</td>
<td>167.0 (9.1)</td>
</tr>
<tr>
<td>Weight</td>
<td>72.8 (17.1)</td>
<td>71.2 (17.3)</td>
<td>76.3 (17.9)</td>
</tr>
<tr>
<td>BMI</td>
<td>26.2 (5.3)</td>
<td>26.5 (6.5)</td>
<td>27.5 (7.1)</td>
</tr>
<tr>
<td>ISWT (m)</td>
<td>183.2 (113.2)</td>
<td>148.5 (105.6)</td>
<td>178.9 (109.6)</td>
</tr>
</tbody>
</table>

Analysis of variance identified there was no significant difference in baseline characteristics (p>0.05). The mean improvement in those who perceived their exercise tolerance was “better” was 78.7m (95% CI 70.5 to 86.9 m) and in those “slightly better” it was 47.5 m (95% CI 38.6 to 56.5 m) and about the same was 18.0 (95% CI 4.6 to 31.5m) (figure 2). The differences in magnitude of change between “better” versus “slightly better” was 31.2 m (95% CI 11.6 to 50.7 m), for “better” versus “about the same” the difference was 60.7m (95% CI 33.6 to 87.6 m), and between “slightly better” versus “about the same” the difference was 29.5 m (95% CI 0.75 to 58.2 m). The effect size (mean within-patient change expressed as a proportion of the
between-patient standard deviation at baseline) was 0.31 for those who were “better” and 0.22 for those who were “slightly better”.

When analysing the data as whole shuttles, for responses 1 to 3, using the Kruskal-Wallis test there was a significant difference between categories of response. Expressed as whole shuttles, for a patients to rate their exercise tolerance as “better” they needed to improve by 8 shuttles, to feel “slightly better” 5 shuttles and report feeling “about the same” 2 shuttles. The MCID to identify improvements in the ISWT is 47.5 m when assessed at a population level in which an average shuttle distance could be calculated or 5 shuttles, since using the ISWT distance is measured in number of whole shuttles completed.

To test whether the improvement was independent of baseline performance, the patients were divided into quartiles based on their baseline ISWT data: 0-80 m; 90-150 m; 160-250 m and 250 m or greater. The increase in distance covered in those who were “slightly better” was not significantly different between the quartiles (ANOVA, p=0.9).

Discussion

The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change for this test. Patients who perceived their performance in the post rehabilitation shuttle test to be slightly better than their baseline performance had on increased their ISWT distance by 48m. When assessed at the level of an individual patient the MCID was 5 shuttles. We have also identified a further category of improvement; an average increase in exercise performance of 78 m (or 8 shuttles) identified a grater level of perceived improvement.

A number of methods are used to define the MCID (13), the most widely used is the patient preference-based estimate, which usually takes the form of a study in which patients judge their current performance against their recall of their performance at some time in the past. Its chief strength is that it is patient centred, and its chief weakness is that it requires a retrospective estimate. Redelmeier et al used a different approach to produce the MCID for the 6MWD and asked cohorts of patients to compare their performance against their peers (14). Their MCID estimate was based upon social interaction between patients, not on direct observations of performance. This method overcame issues about patient recall, but strictly speaking the method provided only an indirect assessment of the MCID, since it depended on the association between the 6MWD and the patients’ overall function within the environment of a rehabilitation class, furthermore it relies on an individual rating an important change between subjects equivalent to an improvement regarded as important to themselves. By contrast, we anchored our patients’ estimates much more closely upon their performance during the test. Issues of retrospective estimates of change are complex, but we believe that our estimates are reliable since we showed a clear rank-order relationship between the size of the perceived benefit and the measured improvement.
In the 55 patients who judged that they had no improvement, the change in ISWT was only 18m, but the lower 95% confidence interval did not include no change. This suggests that patients failed to rate very small changes in exercise performance. This may be due to ‘response shift’, first proposed by Sprangers & Schwartz (15). This theory hypothesises that interaction with healthcare providers or a sustained change in health state may alter the way in which an individual may assess their state. Alternatively the patients’ expectations of benefit may influenced their perceptions, and the classic Hawthorn affect may have been operating – in which subjects who are being observed may change their behaviour. In addition there may have been an improvement in the patients that was not reflected in terms of improved ISWT performance. Either way, the effect appears to have been small and the 5-shuttle MCID should provide a reliable threshold for determining the response rate to a therapy. Uniquely our study also has identified a second more stringent measure of improvement for use by investigators.

Competing interests- the authors have no competing interests.

Acknowledgements: Dr John Bankart, Medical Statistician, University of Leicester for his statistical advice.

Legends
Figure 1  Bland and Altman plot of the average ISWT distance (m) compared to the mean difference (m) (lines of agreement +/- 2 SD’s).

Figure 2  Mean difference (m) and 95% confidence interval by response 1-3.
Reference List


The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence (or non-exclusive for government employees) on a worldwide basis to the BMJ Publishing Group Ltd and its Licensees to permit this article to be published in Thorax editions and any other BMJPGL products to exploit all subsidiary rights, as set out in our licence http://thorax.bmjjournals.com/ifora/licence.pdf
The minimum clinically important improvement for the Incremental Shuttle Walking Test

Sally J Singh, Paul Jones, Rachael Evans and Mike Morgan

Thorax published online April 4, 2008

Updated information and services can be found at: http://thorax.bmj.com/content/early/2008/04/04/thx.2007.081208

These include:

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections

Airway biology (1100)
Lung function (773)

Notes

To request permissions go to: http://group.bmj.com/group/rights-licensing/permissions
To order reprints go to: http://journals.bmj.com/cgi/reprintform
To subscribe to BMJ go to: http://group.bmj.com/subscribe/