Minimum clinically important improvement for the incremental shuttle walking test

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

ABSTRACT
Background: The incremental shuttle walking test (ISWT) is used to assess exercise capacity in patients with 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 for the ISWT.

Methods: 372 patients (205 men) 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, forced expiratory volume in 1 s (FEV1) 1.06 (0.53) l and FEV1/forced vital capacity (FVC) ratio 50.8 (18.1)%. The baseline shuttle walking test distance was 168.5 (114.6) m which increased to 234.7 (125.3) m after rehabilitation (mean difference 65.9 m (95% CI 58.9 to 72.9)). In subjects who felt their exercise tolerance was “slightly better” the mean improvement was 47.5 m (95% CI 38.6 to 56.5) compared with 78.7 m (95% CI 70.5 to 86.9) in those who reported that their exercise tolerance was “better” and 18.0 m (95% CI 4.5 to 31.5) in those who felt their exercise tolerance was “about the same”.

Conclusion: Two levels of improvement were identified. The minimum clinically important improvement for the ISWT is 47.5 m. In addition, patients were able to distinguish an additional benefit at 78.7 m.
Exercise tolerance had changed using the following question: “Compared to last time, how would you rate your exercise tolerance?” Responses were categorised as (1) better, (2) slightly better, (3) about the same, (4) slightly worse or (5) worse; each response was assigned a numerical value from 1 to 5. At this stage the subjects were not informed of any objective change in distance walked after the course of rehabilitation.

**Statistical analysis**

Data were analysed using SPSS V.14. Baseline variables were normally distributed. To estimate the difference in the ISWT to within a precision of ±15 m (as represented by the 95% confidence interval (CI)) and assuming a standard deviation of 36.7 m (generated from previously published data), 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 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 who completed rehabilitation are reported. Data collection was continued until at least 46 patients had been recruited into response categories 1, 2 or 3. At baseline the mean (SD) age was 69.4 (8.4) years, FEV₁ 1.06 (0.58) l, FEV₁/FVC ratio was 50.8 (18.1)% and 55% were men. The mean (SD) baseline shuttle distance was 168.5 (114.6) m which increased to 234.7 (125.3) m after rehabilitation (mean improvement 65.9 m (95% CI 58.9 to 72.9)).

As anticipated, there was no relationship between baseline ISWT performance and the improvement in shuttle distance following rehabilitation. Figure 1 is a Bland-Altman plot for the five groups identifying the mean baseline/post-ISWT against the ISWT difference.

The distribution of responses to the question about perceived improvement was “better” in 50.5% (n = 188), “slightly better” in 29.9% (n = 111), “about the same” in 14.8% (n = 55), “slightly worse” in 4.3% (n = 16) and “worse” in 0.5% (n = 2). One-way analysis of variance (ANOVA) between all five groups identified a significant difference in the mean distance achieved (p < 0.001); post hoc analysis showed that 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” (groups 4 and 5) were too small to give statistically significant results so they were excluded from subsequent analyses. The baseline characteristics of groups 1, 2 and 3 are shown in table 1.

Analysis of variance showed that there was no significant difference in baseline characteristics between groups 1, 2 and 3 (p > 0.05). The mean improvement in those who perceived their exercise tolerance was “better” was 78.7 m (95% CI 70.5 to 86.9) compared with 47.5 m (95% CI 38.6 to 56.5) in those who perceived their exercise tolerance to be “slightly better” and 18.0 m (95% CI 4.6 to 31.5) in those whose exercise tolerance was perceived as being “about the same” (fig 2). The difference in the magnitude of change between groups 1 and 2 was 51.2 m (95% CI 11.6 to 50.7) compared with 60.7 m (95% CI 33.6 to 87.6) between groups 1 and 3 and 29.5 m (95% CI 0.75 to 58.2) between groups 2 and 3. The effect size (mean within-patient change expressed as a proportion of the between-patient standard deviation at baseline) was 0.31 for patients in group 1 and 0.22 for those in group 2.

When analysing the data as whole shuttles for patients in groups 1–3, a significant difference between categories of response was found using the Kruskal-Wallis test. Expressed as whole shuttles, for patients to rate their exercise tolerance as “better” they needed to improve by 8 shuttles compared with 5 shuttles to feel “slightly better” and 2 shuttles to report feeling “about the same”. The MCID to identify improvements in the ISWT performance and the improvement in shuttle distance following rehabilitation. Figure 1 is a Bland-Altman plot for the five groups identifying the mean baseline/post-ISWT against the ISWT difference.

The mean (SD) baseline shuttle distance was 168.5 (114.6) m which increased to 234.7 (125.3) m after rehabilitation (mean improvement 65.9 m (95% CI 58.9 to 72.9)).

As anticipated, there was no relationship between baseline ISWT performance and the improvement in shuttle distance following rehabilitation. Figure 1 is a Bland-Altman plot for the five groups identifying the mean baseline/post-ISWT against the ISWT difference.

The distribution of responses to the question about perceived improvement was “better” in 50.5% (n = 188), “slightly better” in 29.9% (n = 111), “about the same” in 14.8% (n = 55), “slightly worse” in 4.3% (n = 16) and “worse” in 0.5% (n = 2). One-way analysis of variance (ANOVA) between all five groups identified a significant difference in the mean distance achieved (p < 0.001); post hoc analysis showed that 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” (groups 4 and 5) were too small to give statistically significant results so they were excluded from subsequent analyses. The baseline characteristics of groups 1, 2 and 3 are shown in table 1.

Analysis of variance showed that there was no significant difference in baseline characteristics between groups 1, 2 and 3 (p > 0.05). The mean improvement in those who perceived their exercise tolerance was “better” was 78.7 m (95% CI 70.5 to 86.9) compared with 47.5 m (95% CI 38.6 to 56.5) in those who perceived their exercise tolerance to be “slightly better” and 18.0 m (95% CI 4.6 to 31.5) in those whose exercise tolerance was perceived as being “about the same” (fig 2). The difference in the magnitude of change between groups 1 and 2 was 51.2 m (95% CI 11.6 to 50.7) compared with 60.7 m (95% CI 33.6 to 87.6) between groups 1 and 3 and 29.5 m (95% CI 0.75 to 58.2) between groups 2 and 3. The effect size (mean within-patient change expressed as a proportion of the between-patient standard deviation at baseline) was 0.31 for patients in group 1 and 0.22 for those in group 2.

When analysing the data as whole shuttles for patients in groups 1–3, a significant difference between categories of response was found using the Kruskal-Wallis test. Expressed as whole shuttles, for patients to rate their exercise tolerance as “better” they needed to improve by 8 shuttles compared with 5 shuttles to feel “slightly better” and 2 shuttles to report feeling “about the same”. 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

![Figure 1](https://example.com/figure1.png)

**Figure 1** Bland-Altman plot of the mean incremental shuttle walk test (ISWT) distance in metres compared with the mean difference (m) (lines of agreement ± 2SD).

<table>
<thead>
<tr>
<th>Group</th>
<th>Better (group 1)</th>
<th>Slightly better (group 2)</th>
<th>About the same (group 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M/F</td>
<td>103.85</td>
<td>61.50</td>
<td>34.21</td>
</tr>
<tr>
<td>Age (years)</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>

Data are mean (SD). BMI, body mass index; FEV₁, forced expiratory volume in 1 s; FVC, forced vital capacity; ISWT, incremental shuttle walk test.
the ISWT distance is measured in numbers 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. 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 in health state. The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change in health state. The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change in health state.

A number of methods are used to define the MCID, the ISWT is measured in numbers 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. The increase in distance covered in those who were ‘slightly better’ was not significantly different between the quartiles (ANOVA, p = 0.9).

The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change in health state. The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change in health state.

A number of methods are used to define the MCID, the ISWT is measured in numbers 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. The increase in distance covered in those who were ‘slightly better’ was not significantly different between the quartiles (ANOVA, p = 0.9).

The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change in health state. The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change in health state.

A number of methods are used to define the MCID, the ISWT is measured in numbers 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. The increase in distance covered in those who were ‘slightly better’ was not significantly different between the quartiles (ANOVA, p = 0.9).

The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change in health state. The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change in health state. The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change in health state.

A number of methods are used to define the MCID, the ISWT is measured in numbers 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. The increase in distance covered in those who were ‘slightly better’ was not significantly different between the quartiles (ANOVA, p = 0.9).

The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change in health state. The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change in health state. The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change in health state.

A number of methods are used to define the MCID, the ISWT is measured in numbers 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. The increase in distance covered in those who were ‘slightly better’ was not significantly different between the quartiles (ANOVA, p = 0.9).

The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change in health state. The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change in health state. The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change in health state.

A number of methods are used to define the MCID, the ISWT is measured in numbers 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. The increase in distance covered in those who were ‘slightly better’ was not significantly different between the quartiles (ANOVA, p = 0.9).

The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change in health state. The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change in health state. The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change in health state. The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change in health state.

A number of methods are used to define the MCID, the ISWT is measured in numbers 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. The increase in distance covered in those who were ‘slightly better’ was not significantly different between the quartiles (ANOVA, p = 0.9).

The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change in health state. The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change in health state. The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change in health state. The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change in health state. The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change in health state. The ISWT is used widely to assess exercise capacity and we have now identified a threshold for a minimally important change in health state.