An official systematic review of the European Respiratory Society/American Thoracic Society: measurement properties of field walking tests in chronic respiratory disease


Affiliations: For a full list of the authors’ affiliations please refer to the Acknowledgements.

Correspondence: Sally J. Singh, Centre for Exercise and Rehabilitation Science, University Hospitals of Leicester NHS Trust, Groby Road, Leicester, LE3 9QP, UK. E-mail: sally.singh@uhl-tr.nhs.uk

ABSTRACT This systematic review examined the measurement properties of the 6-min walk test (6MWT), incremental shuttle walk test (ISWT) and endurance shuttle walk test (ESWT) in adults with chronic respiratory disease.

Studies that report the evaluation or use of the 6MWT, ISWT or ESWT were included. We searched electronic databases for studies published between January 2000 and September 2013.

The 6-min walking distance (6MWD) is a reliable measure (intra-class correlation coefficients ranged from 0.82 to 0.99 in seven studies). There is a learning effect, with greater distance walked on the second test (pooled mean improvement of 26 m in 13 studies). Reliability was similar for ISWT and ESWT, with a learning effect also evident for ISWT (pooled mean improvement of 20 m in six studies). The 6MWD correlates more strongly with peak work capacity ($r = 0.59–0.93$) and physical activity ($r = 0.40–0.85$) than with respiratory function ($r = 0.10–0.59$). Methodological factors affecting 6MWD include track length, encouragement, supplemental oxygen and walking aids. Supplemental oxygen also affects ISWT and ESWT performance. Responsiveness was moderate to high for all tests, with greater responsiveness to interventions that included exercise training.

The findings of this review demonstrate that the 6MWT, ISWT and ESWT are robust tests of functional exercise capacity in adults with chronic respiratory disease.

@ERSpublications
Systematic review: support for use of the 6MWT, ISWT and ESWT in adults with chronic respiratory disease http://ow.ly/Bq2Mz

This article has supplementary material available from erj.ersjournals.com

Received: Aug 17 2014 | Accepted: Aug 31 2014 | First published online: Oct 30 2014

Support statement: The realisation of this systematic review would not have been possible without the financial support of the European Respiratory Society and American Thoracic Society. Sally J. Singh is supported by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care East Midlands (CLAHRC EM), and work took place at the University Hospitals of Leicester NHS Trust. Support was also provided by the NIHR Leicester Respiratory Biomedical Research Unit. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Dept of Health.

Conflict of interest: Disclosures can be found alongside the online version of this article at erj.ersjournals.com

Copyright ©ERS 2014
Overview
The aim of this systematic review was to examine the measurement properties for the 6-min walk test (6MWT), incremental shuttle walk test (ISWT) and endurance shuttle walk test (ESWT) in adults with chronic respiratory disease. A companion paper describes the standard operating procedures for the tests [1].

The key findings of this systematic review are as follows.

1) The 6-min walking distance (6MWD) is a valid and reliable measure of exercise capacity for people with chronic lung disease. The 6MWD correlates more strongly with measures of peak work capacity and physical activity than with respiratory function or quality of life, which supports its conceptualisation as a test of functional exercise performance.

2) The ISWT offers a different protocol to the 6MWT as it is incremental and externally paced. The ISWT is a valid and reliable measure of cardiopulmonary exercise capacity in chronic obstructive pulmonary disease (COPD), where there is a strong relationship between ISWT distance and peak oxygen uptake ($V_{O2}$peak) or work rate on a cardiopulmonary exercise test (CPET).

3) A learning effect is observed for the 6MWT and the ISWT. The second test usually is the better compared with the first, but this is inconsistent.

4) The ESWT is a test of endurance capacity. It is externally paced and is performed along the same course as the ISWT. Two tests do not appear to be necessary if the second test is conducted on the same day.

5) Reference equations have been proposed for the 6MWD and the ISWT. Age, height and weight are included in most equations. The influence of race and ethnicity is unclear.

6) The safety profile of the 6MWT, ISWT and ESWT is good, with few significant adverse events reported in people with chronic respiratory disease.

7) Performance on the 6MWT can be influenced by the course layout and track length. Such variations have not been reported for the ISWT or the ESWT.

8) Walking aids used during the 6MWT can improve walking distance; those with greater levels of disability appear to gain the most.

9) Supplemental oxygen increases the distance walked on the 6MWT, ISWT and ESWT. The difference is greatest when the cylinder is not carried by the patient.

10) There is a strong association between shorter 6MWD and an increased risk of mortality for patients with COPD, interstitial lung disease (ILD), pulmonary hypertension and those awaiting transplantation. Similar observations are made for the ISWT (mainly for COPD), but there are fewer data in comparison with the 6MWT.

11) The primary outcome measure for the 6MWT and ISWT is distance walked, recorded in metres or feet. The ESWT is usually reported as time.

12) The minimal important difference (MID) for 6MWD in adults with chronic respiratory disease lies between 25 and 33 m. A change of 47.5 m (approximately five shuttles) is reported to be clinically important for the ISWT. The MID for the ESWT is reported as 65 s or 85 m after bronchodilation. A MID for the ESWT after rehabilitation has yet to be confirmed.

13) The 6MWD, ISWT and the ESWT are sensitive to change, particularly after a course of rehabilitation.

Background
Field walking tests play a key role in evaluating functional exercise capacity, assessing prognosis, determining outcome of clinical trials and evaluating response to treatment across a wide range of chronic respiratory diseases. Since the previous American Thoracic Society (ATS) statement on the 6MWT was published in 2002 [2], new information regarding field walking tests has been reported in a range of key areas. This includes the emergence of the ISWT and ESWT as frequently used tests of exercise performance in people with chronic respiratory disease. There is also a new body of knowledge relevant to the good conduct and utility of the 6MWT in both research and clinical settings. This growth in knowledge provides an opportunity to update the standards for field test performance, to enhance test quality and interpretation.

The overall aim of this systematic review was to inform the development of technical standards for the 6MWT, ISWT and ESWT in adults with chronic respiratory disease. The Technical Standard has been published separately [1]. The specific questions addressed in this review were as follows. 1) Are the 6MWT, ISWT and ESWT reliable and valid tests of exercise capacity in people with chronic respiratory disease? 2) Which methodological factors affect test performance in adults with chronic respiratory disease? 3) What
is the relationship of 6MWT, ISWT or ESWT performance to clinical outcomes in people with chronic respiratory disease? 4) Which test parameters, apart from distance, should be reported in people with chronic respiratory disease? 5) What kind of monitoring is required during the 6MWT, ISWT and ESWT in people with chronic respiratory disease? 6) Which reference equations can be used for the 6MWT, ISWT and ESWT? 7) Can the 6MWT, ISWT and ESWT identify clinically meaningful change in people with chronic respiratory disease?

Methods
An ad hoc Task Force was assembled to develop technical standards for the performance of the 6MWT, ISWT, and ESWT based upon a systematic review of the evidence. Co-chairs were selected by the Pulmonary Rehabilitation Assembly and Proficiency Standards Committee of the ATS and the Rehabilitation and Chronic Care Assembly and Allied Respiratory Professionals Assembly of the European Respiratory Society (ERS), then approved by the leadership of both societies. Members of the Task Force were selected by the co-chairs on the basis of their expertise in application of field walking tests in research and/or clinical practice. All potential conflicts of interest were disclosed and managed according to the policies and procedures for joint ATS/ERS projects.

Inclusion criteria
Studies that reported the evaluation or use of the 6MWT, ISWT or ESWT in adults with chronic respiratory disease were included. There were no restrictions on study design. Review papers were not included except in question 7, to determine responsiveness of field tests to interventions. Studies published in any language were included, provided an English abstract was available. Translations of the full text were obtained where necessary. Studies included adults with chronic respiratory disease of any type, diagnosed according to the investigators’ definition. Paediatric studies were not included. As the performance, predictive value and measurement properties of the field tests may vary according to population, analyses were conducted separately for each disease group. Some review questions had additional, specific inclusion criteria and outcomes of interest, which are detailed in the online supplementary material.

Search methods
EMBASE, MEDLINE, CINAHL, PEDro and the Cochrane Library were initially searched for articles published between the year 2000 and May 10, 2012; the search was then updated in September 2013. The search strategies can be found in the online supplementary material. Reference lists of all primary studies and review articles were checked for additional references. Selected studies published before 2000 were included where necessary.

Selection of studies
Two reviewers excluded studies that clearly did not meet the inclusion criteria based on title or abstract. Studies that met the inclusion criteria, or studies where it was not possible to determine inclusion based on title and abstract, were retrieved in full text to determine their inclusion. Full text papers were reviewed independently by two reviewers to determine whether they were suitable for inclusion. Disagreements were resolved by consensus or a third reviewer where required.

Data were extracted by two reviewers independently. Differences were resolved by consensus or a third reviewer where required. Where possible, a pooled estimate of effect was calculated. Where it was not possible to combine data, a narrative synthesis was performed.

Results
6MWT
The 6MWT is a self-paced test of walking capacity. Patients are asked to walk as far as possible in 6 min along a flat course. Walking distance is the primary outcome and is recorded in metres or feet. Standardised instructions and encouragement are commonly given during the test [2].

The initial search returned a total of 2990 records, after removal of duplicates. After removal of records that were not related to the 6MWT or adults with respiratory disease, 1392 records remained. The updated search in September 2013 returned a total of 520 records, of which 161 were pertinent to the 6MWT or individuals with chronic respiratory disease. Further details on selection of studies for each section are given in the online supplementary material.

Reliability of the 6MWD
Reliability is the extent to which a test provides the same result on repeated testing occasions. 30 studies were included in this section: 19 studies in COPD, six studies in cystic fibrosis (CF), six studies in ILD...
of which two were also included in the COPD section) and one study in pulmonary arterial hypertension (PAH). Characteristics of the included studies are given in the online supplementary material (tables S1, S4 and S6).

The 6MWD is a reliable measure, with intra-class correlation coefficients (ICCs) ranging from 0.72 to 0.99 (seven studies) (table S2) [3–8]. Coefficients of variation were small, with narrow ranges in COPD (0.0475–0.073) [3, 9–11], ILD (0.042–0.083) [11–13] and CF (0.0409–0.043) [14, 15]. There were no discernible differences in reliability across diagnostic groups. These data indicate that the vast majority of variation in the 6MWD can be attributed to between-patient variation, rather than within-patient variation. However, the limits of agreement were large (table 1). These studies showed that, although the 6MWD is a reliable measurement, the results of the second test cannot be predicted from the first test.

Learning effect on 6MWD

There is strong evidence of a learning effect for the 6MWD when two or more tests are conducted. 13 studies in patients with COPD showed a pooled mean improvement on the second 6MWT of 26.3 m (table 1). This estimate did not change when including only the subgroup of studies where the two tests were conducted within 24 h (26.1 m). The largest study to address this issue (n=1514) reported a 95% confidence interval for the learning effect of 24–29 m [8]. We can be confident that the mean learning effect on the second 6MWT lies within this range and additional data are unlikely to change this conclusion. Fewer data were available in other chronic respiratory diseases (tables S5 and S7) [11, 13–15, 23–26].

In COPD, the proportion of individuals who walked further on the second 6MWT ranged from 50% to 87% (table S3) [6, 8, 13, 22]. The proportion of individuals who had a clinically significant improvement in 6MWD on their second walk ranged from 15% for an improvement of 54 m [6] to 28% for an improvement of 42 m [8]. In ILD, one study reported that 86% of participants increased their 6MWD on the second test [13].

Effects of test repetition and familiarity on 6MWD

The learning effect appears to be moderated by test repetition and practice, at least in the short term. Three studies in COPD reported that there was a statistically significant increase in walk distance between walks 1 and 3 on a single day [4, 7, 10]. However, one study reported no significant difference between 6MWD for walks 2 and 3 (mean increase 3 m). After three walks, further repetition did not consistently improve 6MWD [4, 10]. In patients with COPD who had performed the 6MWT twice prior to rehabilitation, the post-rehabilitation difference was smaller, although the learning effect may have returned by 3 months (table S3) [22].

<table>
<thead>
<tr>
<th>First author [ref.]</th>
<th>Subjects n</th>
<th>Timepoint</th>
<th>Mean difference (95% CI) m</th>
<th>Limits of agreement m</th>
<th>Mean difference (95% CI) as % of baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEACH [4]</td>
<td>30</td>
<td>Same day</td>
<td>17.1 (14–20)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>POULAIN [16]</td>
<td>10</td>
<td>6 days</td>
<td>-5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>REJESKI [17]</td>
<td>30</td>
<td>1 week</td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RODRIGUES [18]</td>
<td>35</td>
<td>1 day</td>
<td>35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROOMI [19]</td>
<td>15</td>
<td>2–10 days</td>
<td>0.65 (-1.5–16.6)</td>
<td>-62–64</td>
<td></td>
</tr>
<tr>
<td>STEVENS [20]</td>
<td>21</td>
<td>Same day</td>
<td>33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TROOSTERS [21]</td>
<td>20</td>
<td>NA</td>
<td>15 (2.3–27.7)</td>
<td>2.6</td>
<td></td>
</tr>
<tr>
<td>IRIBERRI [10]</td>
<td>30</td>
<td>3 days</td>
<td>55</td>
<td></td>
<td>10.0</td>
</tr>
<tr>
<td>SCIURBA [6]</td>
<td>470</td>
<td>1 day</td>
<td>20.1 (16.1–24.1)</td>
<td>-60–107</td>
<td>7.0</td>
</tr>
<tr>
<td>EISER [7]</td>
<td>23</td>
<td>Same day</td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPENCER [22]</td>
<td>44</td>
<td>Same day</td>
<td>27 (12–42)</td>
<td>-71–148</td>
<td></td>
</tr>
<tr>
<td>JENKINS [13]</td>
<td>245</td>
<td>Same day</td>
<td>37 (33–41)</td>
<td></td>
<td>11.0 (9–12)</td>
</tr>
<tr>
<td>KOZU [11]</td>
<td>45</td>
<td>1 day</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HERNANDES [8]</td>
<td>1514</td>
<td>1 day</td>
<td>27 (24–29)</td>
<td>-60–120</td>
<td>7.0</td>
</tr>
<tr>
<td>Pooled mean</td>
<td></td>
<td></td>
<td>26.3</td>
<td></td>
<td>6.9</td>
</tr>
</tbody>
</table>

NA: not available.
Reliability of other measures taken during the 6MWT

Oxyhaemoglobin saturation measured with pulse oximetry

Nine studies in COPD, ILD and CF reported the reliability of arterial oxygen saturation measured by pulse oximetry (\(\text{SpO}_2\)) measures during the 6MWT (table S8). Mean differences were generally small (1–2%) (table S8) [15, 18, 22, 24]. In COPD and CF, the ICCs for \(\text{SpO}_2\) ranged from 0.81 to 0.97 [8, 15, 27]. However, in systemic sclerosis-associated ILD (SSc-ILD), the ICCs ranged from 0.24 to 0.64, with the most reliable measures obtained using forehead oximetry [24]. This may reflect the difficulty in achieving reliable \(\text{SpO}_2\) measures in this population due to underlying vascular disease. Another study in people with ILD, most of whom had idiopathic pulmonary fibrosis (IPF), also showed substantial measurement variation, with a coefficient of variation of 0.283 [12]. Kappa values for detecting desaturation (\(\text{SpO}_2 \leq 88\%\) or fall \(\geq 4\%\)) were variable, ranging from 0.52 to 0.93 [28].

Heart rate

Six studies in COPD and CF reported the reliability of heart rate (HR) measures (peak HR or change from rest) during the 6MWT (table S9) [8, 15, 16, 18, 22, 27]. Mean differences between tests ranged from -4 bpm to +8 bpm. The ICCs ranged from 0.28 to 0.87 and coefficients of variation ranged from 0.04 to 0.11. There were no clear differences according to underlying disease.

Symptom scores

10 studies in COPD, ILD and CF reported the reliability of symptom scores during the 6MWT (table S10) [7, 8, 12, 15, 16, 18, 22–24, 27].

Dyspnoea

Eight studies reported on reliability of the modified Borg dyspnoea scale [29], with ICCs from 0.59 to 0.92 and mean differences of <1 point. The modified Borg dyspnoea scale may be more reliable than the 15-count dyspnoea scale (ICC 0.66) [27] and visual analogue scale for dyspnoea (coefficient of variation 0.22) [16]; however, few studies have investigated these measures (table S10).

Fatigue

Three studies reported the reliability of the Borg fatigue scale. One study reported an ICC of 0.59 [8]. Another reported a kappa of 0.52 for post-6MWT fatigue, compared with 0.71 for post-6MWT dyspnoea (table S10) [15].

Validity of the 6MWD

Validity is the extent to which a test measures the concept that it is intended to measure. 68 studies were included in this section: 34 studies in COPD; 12 studies in ILD, one of which is also included in the COPD section; seven studies in SSc, one of which is also included in the ILD section; three studies in CF; eight studies in PAH, one of which is also included in the SSc section and one in the ILD section; and eight studies in other disease groups. Full details of included studies are given in the online supplementary material (tables S11–S15).

Relationship of 6MWD to measures of peak exercise capacity

The relationship between 6MWD and \(V'\text{O}_2\)\(_\text{peak}\) on a progressive incremental CPET was moderate to strong, with correlation coefficients ranging from 0.4 to 0.8 (table 2). This relationship was consistent across patient groups. The correlation coefficients for the relationship between 6MWD and peak work on CPET ranged from 0.58 to 0.93 with no difference across patient groups (table 2).

Is the 6MWT a maximal test?

A comparison of peak cardiopulmonary responses on the 6MWT and CPET has been made in 10 studies in COPD, ILD and PAH (table 3). These studies generally included people with moderate to severe lung disease (e.g. in COPD studies, average forced expiratory volume in 1 s (FEV1) ranged from 37% to 52% predicted) and with moderately impaired functional exercise capacity (average 6MWD ranged from 403 to 539 m). The studies excluded people requiring supplemental oxygen or gait aids, due to the nature of the testing protocols. There was no difference between tests in \(V'\text{O}_2\)\(_\text{peak}\) or peak HR for seven out of the eight studies where these were reported. However, the peak carbon dioxide production (\(V'\text{CO}_2\)\(_\text{peak}\)), peak ventilation (\(V'\text{Epeak}\)) and respiratory exchange ratio were significantly lower during the 6MWD in six out of the seven studies where these variables were measured. There were no discernible differences between patient groups.
Relationship of 6MWD to disease severity

There were weak to moderate correlations between 6MWD and measures of disease severity such as FEV$_1$ (in COPD; correlation coefficients 0.31–0.70) (table S16) [17, 30–32, 50–58], forced vital capacity (FVC) or diffusing capacity of the lung for carbon monoxide (D$_{LCO}$) (in ILD and SSc; correlation coefficients 0.06–0.61) (tables S17 and S18) [12, 23, 59–65]. In COPD, two studies reported that the relationship between FEV$_1$ and 6MWD was stronger in more severe disease [54, 66]. In CF, one study reported moderate relationships between 6MWD and FEV$_1$ (r = 0.53) and FVC (r = 0.62) [67]. In PAH, correlations between 6MWD and mean pulmonary artery pressure ranged from -0.2 to -0.62 [43, 63, 68, 69].

Relationship of 6MWD to patient-reported outcomes

Although the 6MWD was consistently associated with symptoms and health-related quality of life (HRQoL) across all disease groups, these relationships were generally of weak to moderate strength (correlation coefficients 0.01–0.65) (tables S19 and S20) [17, 19, 30, 32, 50, 51, 56, 58, 61, 62, 67, 70, 71].

Relationship of 6MWD to physical activity

There were moderate to strong relationships between 6MWD and objective measures of physical activity in patients with COPD, CF, ILD and PAH (correlation coefficients 0.38–0.85) (table 4).

Multivariate models for predicting 6MWD

A wide variety of independent predictors of 6MWD were identified, with no consistency within or between disease groups (tables S21 and S22) [19, 30, 32, 50, 51, 58, 60, 64, 82]. These models predicted between 20% and 70% of the variation in 6MWD, suggesting that there are other contributors to 6MWD that have not yet been defined.

---

**TABLE 2 Relationship between 6-min walking distance and peak oxygen uptake (V'0$_{peak}$) or peak work in chronic respiratory disease**

<table>
<thead>
<tr>
<th>First author [ref.]</th>
<th>Diagnosis</th>
<th>Subjects n</th>
<th>V'0$_{peak}$</th>
<th>Peak work</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Pearson’s r</td>
<td>Spearman’s rho</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pearson’s r</td>
</tr>
<tr>
<td>WIJKSTRA [30]</td>
<td>COPD</td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>REJESKI [17]</td>
<td>COPD</td>
<td>209</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>CHUANG [31]</td>
<td>COPD</td>
<td>27</td>
<td>0.40</td>
<td></td>
</tr>
<tr>
<td>Oga [32]</td>
<td>COPD</td>
<td>36</td>
<td></td>
<td>0.64</td>
</tr>
<tr>
<td>CARTER [33]</td>
<td>COPD</td>
<td>124</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>SATAKE [34]</td>
<td>COPD</td>
<td>12</td>
<td>0.73</td>
<td></td>
</tr>
<tr>
<td>TURNER [35]</td>
<td>COPD</td>
<td>20</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>STORBORG [36]</td>
<td>COPD</td>
<td>50</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td>HILL [37]</td>
<td>COPD</td>
<td>50</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>LUXTON [38]</td>
<td>COPD</td>
<td>22</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>DIAZ [39]</td>
<td>COPD</td>
<td>81</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>KOZU [11]</td>
<td>COPD</td>
<td>45</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td>SILLEN [40]</td>
<td>COPD</td>
<td>2906</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>EATON [12]</td>
<td>IPF</td>
<td>29</td>
<td></td>
<td>0.78</td>
</tr>
<tr>
<td>KOZU [11]</td>
<td>IPF</td>
<td>35</td>
<td></td>
<td>0.80</td>
</tr>
<tr>
<td>HOLLAND [41]</td>
<td>ILD</td>
<td>14</td>
<td></td>
<td>0.93</td>
</tr>
<tr>
<td>DALE [42]</td>
<td>Asbestos-related pleural disease</td>
<td>25</td>
<td>0.53</td>
<td>0.58</td>
</tr>
<tr>
<td>MIYAMOTO [43]</td>
<td>PAH</td>
<td>27</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>DEREK [44]</td>
<td>PAH</td>
<td>20</td>
<td>0.48</td>
<td></td>
</tr>
<tr>
<td>FOWLER [45]</td>
<td>EIPAH</td>
<td>17</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td>BALDI [46]</td>
<td>LAM</td>
<td>40</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>CAHALIN [5]</td>
<td>End-stage lung disease, pre transplant</td>
<td>30</td>
<td>0.73</td>
<td></td>
</tr>
<tr>
<td>ROSS [47]</td>
<td>Mixed chronic lung diseases, many pre transplant</td>
<td>48</td>
<td>0.59</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>First author [ref.]</th>
<th>Subjects n</th>
<th>Patient group</th>
<th>$V'O_2$peak mL·min$^{-1}$</th>
<th>HR peak bpm</th>
<th>$V'CO_2$peak mL·min$^{-1}$</th>
<th>RER</th>
<th>$V'$peak L·min$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>6MWT</td>
<td>CPET</td>
<td>6MWT</td>
<td>CPET</td>
<td>6MWT</td>
</tr>
<tr>
<td>TROOSTERS [21]</td>
<td>20</td>
<td>COPD</td>
<td>1400 ± 290</td>
<td>1410 ± 180</td>
<td>126 ± 13</td>
<td>130 ± 13</td>
<td>1300 ± 310</td>
</tr>
<tr>
<td>SATAKE [34]</td>
<td>12</td>
<td>COPD</td>
<td>14.5 ± 2.1*</td>
<td>13.9 ± 3.6*</td>
<td>109 ± 19</td>
<td>112 ± 16</td>
<td>12.56 ± 2.02*</td>
</tr>
<tr>
<td>TURNER [35]</td>
<td>20</td>
<td>COPD</td>
<td>1394 ± 404</td>
<td>1383 ± 437</td>
<td>126 ± 15</td>
<td>124 ± 12</td>
<td>1069 ± 378</td>
</tr>
<tr>
<td>LUXTON [38]</td>
<td>22</td>
<td>COPD</td>
<td>1195 ± 15</td>
<td>1208 ± 537</td>
<td>121 ± 9</td>
<td>128 ± 22</td>
<td>1069 ± 378</td>
</tr>
<tr>
<td>HILL [37]</td>
<td>50</td>
<td>COPD</td>
<td>1186 ± 314</td>
<td>1208 ± 437</td>
<td>121 ± 9</td>
<td>128 ± 22</td>
<td>1069 ± 378</td>
</tr>
<tr>
<td>HILL [48]</td>
<td>26</td>
<td>COPD</td>
<td>1168 ± 344</td>
<td>1186 ± 314</td>
<td>128 ± 17</td>
<td>128 ± 19</td>
<td>1008 ± 270</td>
</tr>
<tr>
<td>BLANCO [49]</td>
<td>13</td>
<td>ILD</td>
<td>14.3 ± 2.7*</td>
<td>12.9 ± 3.1*</td>
<td>119 ± 17</td>
<td>135 ± 17</td>
<td>881 ± 290</td>
</tr>
<tr>
<td>HOLLAND [41]</td>
<td>14</td>
<td>ILD</td>
<td>15.1 ± 3.5*</td>
<td>17.5 ± 2.6*</td>
<td>119 ± 17</td>
<td>135 ± 17</td>
<td>881 ± 290</td>
</tr>
<tr>
<td>DIBDECK [44]</td>
<td>20</td>
<td>PAH</td>
<td>14.2 ± 2.7*</td>
<td>12.9 ± 3.1*</td>
<td>119 ± 17</td>
<td>135 ± 17</td>
<td>881 ± 290</td>
</tr>
<tr>
<td>BLANCO [49]</td>
<td>14</td>
<td>PAH</td>
<td>16 ± 6*</td>
<td>16 ± 6*</td>
<td>142 ± 24</td>
<td>146 ± 23</td>
<td>1000 ± 240</td>
</tr>
</tbody>
</table>

Data are presented as mean ± SD, unless otherwise stated. $V'O_2$peak: peak oxygen uptake; HR: heart rate; $V'CO_2$peak: peak carbon dioxide production; RER: respiratory exchange ratio; $V'$peak: peak ventilation; COPD: chronic obstructive pulmonary disease; ILD: interstitial lung disease; PAH: pulmonary arterial hypertension. *: incremental cycle ergometer test; †: data given in mL·kg$^{-1}$·min$^{-1}$. #: p<0.05 versus 6MWT data.
Relationship of 6MWD to patient performance ratings

Studies in COPD and ILD demonstrated that patient perceptions of differences in their walking capacity were reflected by significant differences in the 6MWD, both in comparison with other patients [83] and pre-versus post-rehabilitation [84, 85].

Technical factors affecting 6MWT performance

The search strategy revealed 43 papers that were screened in full text, with 20 papers meeting the inclusion criteria.

Hallway versus treadmill

In two studies, a 6MWT performed on an externally paced treadmill was 13–20% lower than that performed in the hallway [20, 86]. This may be the result of the poor walking efficiency during treadmill walking in subjects unaccustomed to this activity [20].

Course location and layout

One study compared an indoor course with an outdoor course and reported little difference in 6MWD (mean difference 4 m) [87]. One large multicentre study showed a small benefit of circular versus linear track layout (mean difference 19 m; 5%) [6], and another also showed a slight increase in 6MWD on a circular track (mean difference 13 m; 3%) [88].

Track length

The effect of track length was examined in two studies. In one nonrandomised cross-sectional study in patients with severe COPD (candidates for lung volume reduction surgery), no differences in average outcome were found between tracks of 15–50 m [6]. In contrast, a randomised crossover study comparing track lengths of 30 m and 10 m in patients with moderate COPD found a mean increase in 6MWD of 49.5 m on the longer course (95% CI 39.4–59.6 m) [89]. The greater number of turns required for the very short track length probably contributes to this finding.

Walking aid

Six studies investigated the effect of a walking aid on 6MWD in people with COPD [90–95]. In all five studies investigating the effect of rollators, there was an improvement in 6MWD, ranging 2–46 m or 1–14%
(weighted mean 6.2% improvement). Two studies concluded that the benefits were seen most in those patients with more impaired walking distance. An additional study found that use of a modern draisine resulted in a longer 6MWD than the rollator (mean difference 83 m) [95].

Oxygen
All five studies comparing walking with and without oxygen showed an increased 6MWD when walking with oxygen supplementation in patients with COPD (mean difference 12–59 m) [96–100]. This increase in 6MWD may be further amplified when walking with oxygen is compared with walking with compressed air, due to the handicap of carrying the weight of the compressed air (mean difference 17–109 m) [99, 101].

Method of carrying the oxygen
One study showed that people with COPD who used a wheeled cart to carry their oxygen cylinder walked on average 23 m further than those who carried it on their shoulder [102]. A separate study demonstrated that if the patient carried the oxygen cylinder themselves, the beneficial effect of oxygen was smaller than if the oxygen was carried by an investigator (24 m versus 35 m improvement) [100].

Instructions
A study in 24 patients with ILD and PAH found an average increase in 6MWD of 52.7 m when the patient was asked to walk as “fast” as they could for 6 min, rather than the standard instruction of as “far” as they could in 6 min [103]. It was not clear whether the patients were naïve to the test or whether they were blinded to the purpose of the study.

Encouragement
No papers since 2000 have reported on the effects of encouragement. However, the effect of encouragement is generally recognised and was studied by Guyatt et al. [3] in 1984. These authors reported an average effect of encouragement every 30 s of 30.5 m.

Medication
Improvement in 6MWD following bronchodilator administration has been demonstrated in patients with COPD [104–107], although its magnitude appears small and may be clinically insignificant (mean difference 6–7 m) [106, 108].

Relationship of 6MWD to clinical outcomes
We identified 35 studies that assessed the association of 6MWD with mortality and/or hospitalisation in patients with chronic lung disease. 14 studies included patients with COPD [109–122], eight studies included patients with ILD [69, 123–129], nine studies included patients with PAH [43, 130–137] and four included studies patients with other chronic respiratory disease or on the list for transplantation [138–141]. Average 6MWD at baseline was between 300 and 450 m (fig. 1). Studies differed by population, source of recruitment, number of covariates in the statistical models and modelling of 6MWD in the analyses (i.e. as continuous or categorical variable). Therefore, we did not perform meta-analysis. Instead, we determined the proportion of studies that found a statistically significant association, defined by a p-value of ≤0.05. We have referred to p-values from multivariate statistical models whenever possible.

In 13 (93%) out of the 14 COPD studies, a lower 6MWD was significantly associated with increased mortality (table S23) [109–122]. Two studies also assessed the association of 6MWD with hospitalisation and both of them (100%) found statistically significant associations [109, 120]. In patients with ILD, four (50%) out of eight studies found statistically significant associations of 6MWD with mortality (table S24) [69, 123–129]. Six (66%) out of nine studies in patients with pulmonary hypertension found statistically significant associations of 6MWD with mortality (table S25) [43, 130–137]. One assessed the association with hospitalisation and found a statistically significant association [134]. Finally, all four studies (100%) that included patients with other respiratory disease or on the list for transplantation found statistically significant associations of 6MWD with mortality (table S26) [138–141].

In summary, evidence from 27 (77%) out of 35 studies shows that lower 6MWD is consistently associated with increased mortality. There is less evidence for the association with hospitalisation, although it was consistently present in all the studies where it was assessed (n=3).

Measurements and reporting for 6MWT
The primary outcome measure for the 6MWT is distance (6MWD), reported in metres or feet. Other measures include exertional dyspnoea and fatigue, oxyhaemoglobin saturation, HR, 6MWD × body weight (6-min walk work), and measures derived from a combination of these variables.
Dyspnoea

11 studies were included in this section: eight in COPD and three in ILD. Several authors have reported a relationship between 6MWD and modified Borg dyspnoea score in people with COPD [142–145], with correlation coefficients ranging from 0.39 to 0.49. Two studies showed that modified Borg score was an independent predictor of 6MWD [143, 145]. In 50 people with COPD and mean FEV1 45% predicted, the reduction in inspiratory capacity following a 6MWT was associated with the change in modified Borg score during the test ($r^2 = 0.21, p = 0.0006$) [146]. Other studies have reported close relationships between modified Borg scores during the 6MWT in COPD and the severity of dyspnoea during daily life ($r = 0.54$) [147] and HRQoL ($r = -0.69$–$-0.43$) [58, 148].

In people with IPF, $SpO_2$ at the end of 6MWT was an independent predictor of modified Borg dyspnoea score ($r^2 = 0.27, p < 0.0005$) [149]. In contrast, in sarcoidosis, the strongest predictor of modified Borg dyspnoea score during 6MWT was nonvolitional measures of respiratory muscle strength [150]. Modified Borg dyspnoea score after the 6MWT was also an independent predictor of HRQoL in people with sarcoidosis [151].

Exertion or subjective fatigue

Subjective fatigue is an important symptom in people with a range of chronic lung diseases. 23 studies related to subjective fatigue were retrieved in full text and 12 were included in this section, all related to COPD [148–159].

All dimensions of subjective fatigue are higher in COPD compared with healthy aged subjects; almost half of the patients (47.2%) report fatigue every day compared with 13.5% of healthy individuals [159]. This is reflected during the 6MWT, where COPD patients experienced more subjective fatigue compared with healthy elderly subjects [55]. Patients with higher tumour necrosis factor-$\alpha$ reported more fatigue after 6MWD ($p = 0.054$) [153]. Dimensions of subjective fatigue have been related to pulmonary function, skeletal muscle force and quality of life in COPD patients [154, 155]. Subjective fatigue measured during the 6MWT using the Borg exertion scale is associated with lower 6MWT, slower gait speed, more severe lung disease, more dyspnoea on exertion and lower HRQoL (table S27) [148, 156–158]. Although there are fewer
data available to assess the impact of subjective fatigue during 6MWT compared with dyspnoea, existing studies support its role as an important contributor to 6MWD and disease status in COPD.

Heart rate responses
Nine studies were retrieved in full text. Three studies were excluded after full text review because of lack of pertinent data on HR responses. Six studies were included: three studies in PAH, two studies in IPF and one study in COPD.

HR responses may contribute to performance on the 6MWT in people with chronic respiratory disease. In a study using factor analysis, “HR pattern” was one of four factors that together explained 78.4% of the variance in 6MWD in 83 patients with severe COPD [145]. The components of “HR pattern” were HR at baseline and 6 min, maximum HR, increase in HR, HR at 2 min into recovery, and the HR recovery (HRR) at 2 min. Impairment in the HR response during the 6MWT, also known as “chronotropic response”, has been shown to predict 6MWD in idiopathic PAH (n = 46) and nonidiopathic PAH (n = 37) [160].

HRR refers to the reduction in HR with rest after the exercise period. In IPF, an abnormal HRR at 1 min (HRR1), defined as ≤13 beats, was a strong predictor of mortality [128]. Abnormal HRR1 was also the strongest predictor of pulmonary hypertension in IPF on right heart catheterisation in multivariable analysis that included 6MWD (odds ratio 4.0, 95% CI 1.17–13.69; p = 0.02) [161]. In PAH, HRR following 6MWT was consistently slowed in PAH patients compared with controls, and the best cut-off of HRR1 to separate PAH patients from controls was 18 beats [162]. In 75 patients with IPAH, a HRR1 < 16 beats was found to be a strong predictor of clinical worsening [134].

In summary, the measurement of HR appears valuable and should be included in 6MWT assessments. Consideration should be given to measurement of HRR1; however, there is no universally accepted cut-off for this measure.

Oxyhaemoglobin saturation
Oxygen desaturation is commonly observed during the 6MWT in patients with lung diseases [163]. In a group of individuals with COPD and moderate to severe lung disease (mean FEV1 43% predicted), the prevalence of exercise-induced desaturation on 6MWT was 62%, where desaturation was defined as a drop of >4% in \( S_{\text{pO2}} \) to < 90% [164]. Desaturation during a 6MWT provides information regarding disease severity and prognosis in COPD [114, 115], ILD [123, 165], PAH [130] and SSc patients with pulmonary manifestations [82]. Desaturation during the 6MWT has been associated with more severe lung disease, higher levels of dyspnoea, poorer muscle strength, impaired daily physical activity and desaturation during daily life activities across a range of chronic respiratory diseases (table S28). These relationships support the clinical importance of quantifying desaturation during the 6MWT.

The 6MWT has been found to be more sensitive in identifying exercise-induced desaturation compared with cycle testing. In a study by POULAIN et al. [16], 28% of patients with COPD exhibited desaturation on 6MWT that was not observed on CPET. Other authors have reported similar findings [35, 37, 38]. The phenomenon was reproducible and not protocol dependent, emphasising the clinical interest of the 6MWT [16].

The 6MWT may be safe without continuous monitoring of \( S_{\text{pO2}} \) [166]. However, continuous monitoring may be necessary to accurately record the extent of desaturation in patients with ILD and COPD. FIORE et al. [167] found that, although the end-6MWT \( S_{\text{pO2}} \) and the nadir \( S_{\text{pO2}} \) were similar in most patients with COPD and ILD, the end \( S_{\text{pO2}} \) did not provide an accurate estimate of nadir \( S_{\text{pO2}} \) in patients who rested during the test. A recent study in COPD reported similar findings, with 46% of patients exhibiting a “desaturation–resaturation” pattern, with nadir \( S_{\text{pO2}} \) occurring relatively early in the test [168]. Therefore, constant monitoring of \( S_{\text{pO2}} \) during the 6MWT is needed to obtain an accurate measure of exercise-induced desaturation.

Novel desaturation indices
A number of indices derived from desaturation measurements have been proposed in patients with ILD, with the aim of improving the predictive value of information gained from the 6MWT.

The distance–saturation product (DSP) (given in m%) is defined as the product of the final distance walked in metres and the lowest \( S_{\text{pO2}} \) when breathing air. In 80 patients with IPF, a DSP ≤200 m% was associated with a markedly increased hazard ratio for mortality (6.5 (95% CI 3.1–12.7)) [125]. A subsequent study has shown that the DSP was not useful for detecting pulmonary hypertension in people with IPF who were listed for lung transplantation (DSP in those with pulmonary hypertension 272 m%, versus 286 m% in those without) [169].
A retrospective study in 59 individuals with sarcoidosis found that the DSP had a stronger relationship with measures of lung function, oxygenation and dyspnoea (standardised beta 0.45–0.59) than the 6MWD (standardised beta 0.19–0.52) [170]. The relationship between DSP and prognosis was not evaluated. The DSP also predicts HRQoL in sarcoidosis [151]. A model containing the DSP and end-6MWT modified Borg dyspnoea score was better able to predict SF-36 (short-form health survey) scores ($R^2=0.33$) than a similar model where 6MWD was substituted for DSP ($R^2=0.26$).

The desaturation area is defined as the total area above the curve between the $SpO_2$ observed at each minute of the 6MWT and 100%. A higher desaturation area indicates higher overall desaturation during the 6MWT. FLAHERTY et al. [165] showed that, in a group of patients with IPF, for every 10-point increase in desaturation area there was an increased hazard ratio for mortality (1.30 (95% CI 0.97–1.75)). This measure has not been widely used.

The desaturation–distance ratio is a combination of 6MWD and desaturation area. One study in 49 people with ILD has shown a strong relationship between desaturation–distance ratio and measures of disease severity [171]. The prognostic value of this relatively complex measure has not been studied.

In summary, a number of newer measures incorporating desaturation are available for use in ILD. They are not routinely collected during the 6MWT and their utility has not been confirmed.

### 6-min walk work

Body weight directly affects the work/energy required to perform the 6MWT. The 6-min walk work is the product of 6MWD and body weight, which may provide a better estimate of the work required to perform the test than distance alone. Three studies have investigated the use of 6-min walk work as an outcome of the 6MWT in COPD. The 6-min walk work correlates more strongly with $V'_{O_2}$peak than 6MWD alone, with correlation coefficients of 0.67–0.81 versus 0.40–0.54, respectively [31, 33]. A similar pattern is seen with $DLCO$ (0.60–0.70 versus 0.35–0.46) [31, 33]. The relationship between FEV1 and 6-min walk work has been reported as weak to moderate ($r=0.2–0.52$) [31, 33, 172]. None of the studies investigated the sensitivity of 6-min walk work to change over time. Additional studies are needed to better characterise the utility of 6-min walk work in adults with chronic respiratory disease.

### Adverse events during the 6MWT

Reported complications associated with the performance of the 6MWT are unusual. Of the research that has been published to date, only two papers have specifically addressed the issue of complications associated with the test itself. Patients taking part in outpatient pulmonary rehabilitation (n=741) completed the 6MWT with continuous monitoring of oxygen saturation and HR [163]. Adverse events were noted in 43 (6%) of tests. One test was terminated by a patient developing chest pain and another by a patient developing persistent tachycardia. In six tests, the patient developed symptoms and the test was discontinued. In 35 tests, the patient was instructed to discontinue walking because the $SpO_2$ fell below 80%.

In contrast, in a much smaller study in patients with ILD (n=19) and $DLCO$ 41% predicted, the 6MWD was allowed to continue if $SpO_2$ fell below 80% [166]. No clinically significant cardiac arrhythmias were noted, despite marked desaturation in several participants.

Many patients who undergo a 6MWT develop oxygen desaturation [164]; however, significant adverse events appear to be very rare. Nevertheless, many previous studies have terminated the test before significant desaturation occurs [113, 123, 124, 163]. As a result, the safety of the 6MWT during severe desaturation ($SpO_2 <80\%$) has not been clearly documented.

### Reference equations for the 6MWT

30 papers were reviewed in full text and 17 were included in this section, with reference equations shown in table 5. These studies were performed using a wide variety of populations and methodologies (table S29). This gives rise to substantial variation in the predicted 6MWD. Within groups of healthy people, independent predictors of the 6MWD include height, age, sex and weight (table 5). However, large differences in the predicted 6MWD occur among the published studies.

### Site-to-site variation

In the Asociación Latinoamericana de Tórax (ALAT) study [173], Brazilian adults walked $>100\,$m farther than subjects from Venezuela (638 versus 510 m). Some authors have concluded that different countries require specific equations for the 6MWD [173]. However, even between cities within a single country, large differences in mean 6MWD have been seen (e.g. 510 m from Zaragoza, Spain versus 613 m from Tenerife, Spain) [173]. A portion of the site-to-site differences were due to lower mean 6MWD from sites with a
lower mean exercise intensity (end-of-exercise HR as a percentage of estimated maximum HR(%max HR)), probably due to variations in test administration.

Effect of race/ethnicity
Many authors have concluded that differences in mean 6MWD between their study and previous studies of Caucasians were due to racial or ethnic differences [174, 178, 182, 185, 190]. However, because of differences in protocols, and thus in the exercise intensity, it is difficult to interpret differences between studies performed at a single site with a single race/ethnic group. Only one study included two different racial groups at the same sites, allowing a direct comparison [187]. Using a regression model, this study of healthy elderly people reported a mean 6MWD for African-Americans 39 m lower than for Caucasian Caucasians due to racial or ethnic differences [174, 178, 182, 185, 190]. However, because of differences in protocols, and thus in the exercise intensity, it is difficult to interpret differences between studies performed at a single site with a single race/ethnic group. Only one study included two different racial groups at the same sites, allowing a direct comparison [187]. Using a regression model, this study of healthy elderly people reported a mean 6MWD for African-Americans 39 m lower than for Caucasian subjects. Similar within-study comparisons have not been reported for other age groups or other racial/ethnic groups.

Methodological differences
The mean walking distances obtained from the healthy participants in the Cardiovascular Health Study [187] were relatively low because participant instructions were to “walk from end to end of the hallway at your own pace.” Many studies published after the ATS guidelines [2] have reached higher values of 6MWD. In these studies, assessors asked the subjects to reach the farthest possible distance in 6 min, which may prompt a higher walking speed.

Exercise intensity
Mean exercise intensity (%max HR) varied from 44% to 81% between studies that reported this outcome [191]. Exercise intensity may be influenced by the technician and by the patient’s motivation. Higher exercise intensity provides a higher 6MWD (after adjusting for other factors using linear regression models). Some investigators have produced 6MWD prediction equations which include measured %max HR [185].
However, caution is needed when using this approach for patients aged ≥40 years, since patients with clinically apparent cardiopulmonary disease often have a higher HR at a given level of exercise (when compared with healthy people), which would cause their %max HR to be elevated, causing an overestimation of their exercise intensity. Conversely, those taking beta-blockers for hypertension or coronary artery disease have attenuation of %max HR [176], causing an underestimation of their exercise intensity.

Number of tests
The authors of some 6MWD reference equations performed only one test, while investigators of other studies always performed two tests and reported the highest 6MWD obtained (table S29). The mean improvement in 6MWD when performing two 6MWTs is approximately 26 m [8], so this does not fully explain the much larger mean differences between studies.

Identifying meaningful change in the 6MWD
Minimal important difference of the 6MWD
We evaluated the full text of 22 studies (table S30) and finally included studies that assessed the MID in patients with COPD (six studies), ILD (three studies) and PAH (two studies). Table S31 shows the study and participant characteristics. The >5600 patients included in these studies had moderate to severe lung disease and received rehabilitation, drug treatments, lung volume reduction surgery or no treatment. Mean baseline 6MWD ranged from 343 to 403 m.

Table 6 summarises the methods and outcomes of studies used to determine the MID. The majority of estimates are based on distribution-based methods because correlations with potential anchors were too low for anchor-based methods. Three studies in COPD, all three in ILD and one study in PAH also provided anchor-based estimates. Figure 2 shows all estimates for the MID. There is relatively little variability across patient groups. Estimates based on anchor-based methods ranged from 21.6 to 38.6 m, with a median estimate across studies of 24.8 m. Distribution-based estimates varied more (range 26–57 m), with a median estimate of 32.1 m, which can be explained by different effect-size-based methods used in different studies. The study by REDELMEIER et al. [83] used an entirely different approach, where patients estimated the difference between themselves and other patients during rehabilitation. The estimate from that study was 54 m. Across all estimates (based on any method), the median estimate was 30 m.

The available evidence suggests a MID of 30 m for adult patients with chronic respiratory disease. There is some variability across studies and methods to determine the MID; however, based on the large evidence base now available, we can be confident that the MID lies between 25 and 33 m.

Responsiveness of the 6MWD
Responsiveness is the capacity of a measure to detect meaningful change over time. Relatively few studies have been explicitly designed to assess the responsiveness of 6MWD but a large number of randomised trials provide insights into how responsive 6MWD is to treatment effects. Therefore, we selected systematic reviews from the Cochrane Collaboration that evaluated interventions where change in 6MWD is expected to occur. We chose the Cochrane reviews on exercise interventions for COPD and ILD and the Cochrane review on endothelin receptor antagonists for PAH, to illustrate the responsiveness of 6MWD.

For COPD, the reviews by LACASSE et al. [199] and PUHAN et al. [200] summarised the trials on pulmonary rehabilitation versus usual care in patients with stable COPD and after a COPD exacerbation, respectively. In the review of trials in stable COPD patients, 16 trials reported on 6MWD, of which 15 reported improvements of 6MWD with rehabilitation compared with usual care. Most trials were underpowered to detect statistically significant effects, but the meta-analysis showed an average improvement of 48 m (95% CI 32–65 m; p<0.0001) [199]. The effect size (pooled estimate divided by pooled standard deviation of change in the control group) as a measure for responsiveness was 0.51. In the review in patients after a COPD exacerbation, six trials reported on 6MWD, of which four reported statistically significant improvements of 6MWD with rehabilitation compared with usual care and two trials reported no improvements. The meta-analysis showed an improvement of 78 m (95% CI 12–143 m; p<0.0001) with a corresponding effect size of 1.07 (median standard deviation of change in control group of 73 m) [200].

The systematic review on physical exercise training in patients with ILD identified two trials that reported on 6MWD [201]. Both trials comparing exercise with usual care showed statistically significant improvement of 6MWD and the meta-analysis showed an improvement of 39 m (95% CI 15–62 m). The corresponding effect size was 0.65 (median standard deviation of change in control group of 58 m).

11 trials reported on the effects of selective and nonselective endothelin receptor antagonists versus placebo on 6MWD in patients with PAH [202]. 10 trials showed an improvement of 6MWD but in only six trials was the effect statistically significant. The meta-analysis showed a statistically significant improvement of
<table>
<thead>
<tr>
<th>First author [ref.]</th>
<th>Patient population</th>
<th>Study design</th>
<th>Intervention</th>
<th>Anchor-based</th>
<th>Distribution-based</th>
<th>Other method</th>
<th>Proposed MID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Redelmeier [83]</td>
<td>Moderate to severe COPD</td>
<td>Prospective, single arm</td>
<td>Rehabilitation</td>
<td>NA</td>
<td>Between-patient ratings: 54 m</td>
<td>NA</td>
<td>54 m</td>
</tr>
<tr>
<td>Wiste [192]</td>
<td>Severe to very severe COPD</td>
<td>Prospective, single arm within RCT</td>
<td>None (test-retest)</td>
<td>NA</td>
<td>SEM: 45.3 m; ERES: 45.3 m; CRM: 80.0 m; average across three estimates: 57.4 m</td>
<td>NA</td>
<td>54 m (triangulation)</td>
</tr>
<tr>
<td>HOLLAND [84]</td>
<td>Moderate to severe COPD</td>
<td>Prospective, single arm</td>
<td>Rehabilitation</td>
<td>25 (20–61) m</td>
<td>SRM: 1.2; SEM: 25.5 m; average across two estimates: 26.0 m</td>
<td>NA</td>
<td>25 m (triangulation)</td>
</tr>
<tr>
<td>PUHAN [194]</td>
<td>Severe to very severe COPD</td>
<td>RCT LVRS or medical treatment</td>
<td>None</td>
<td>18.9 (18.1–20.1) m; 24.2 (23.4–25.4) m; 24.6 (24.2–27.4) m; 26.4 (25.4–27.4) m; average across four estimates: 25.5 m</td>
<td>NA</td>
<td>26 m (mean of all MIDs)</td>
<td></td>
</tr>
<tr>
<td>POLKEY [195]</td>
<td>Mild to very severe COPD</td>
<td>Prospective, observational</td>
<td>Rehabilitation</td>
<td>NA</td>
<td>30 m</td>
<td>NA</td>
<td>30 m</td>
</tr>
<tr>
<td>HOLLAND [85]</td>
<td>Moderate to severe DPLD (50% IPF)</td>
<td>Prospective, single arm</td>
<td>Rehabilitation</td>
<td>30.5 m</td>
<td>SEM: 33 m</td>
<td>NA</td>
<td>30–33 m</td>
</tr>
<tr>
<td>SWIGRIS [196]</td>
<td>Moderate to severe IPF</td>
<td>RCT Bosentan</td>
<td>NA</td>
<td>21.6 m</td>
<td>ES: 41.3 m; SEM: 30.6 m; ERES: 25.7 m; average across three estimates: 27.7 m</td>
<td>NA</td>
<td>28 m (mean of all MIDs)</td>
</tr>
<tr>
<td>DU BOIS [62]</td>
<td>Moderate to severe IPF</td>
<td>RCT Interferon-β</td>
<td>NA</td>
<td>24 m</td>
<td>SEM: 45 m</td>
<td>NA</td>
<td>No specific estimates (range 24–45 m)</td>
</tr>
<tr>
<td>GILBERT [197]</td>
<td>Moderate to severe PAH</td>
<td>RCT Sildenafil (three dosages)</td>
<td>NA</td>
<td>21.6 m</td>
<td>ES: 31.2 m; SEM: 24.7 m; ERES: 25.7 m; average across three estimates: 27.7 m</td>
<td>NA</td>
<td>41 m (median across all estimates)</td>
</tr>
<tr>
<td>MATHAI [198]</td>
<td>Moderate to severe PAH</td>
<td>RCT Tadalafil or placebo</td>
<td>NA</td>
<td>38.6 m</td>
<td>ES: 38.5 m; SEM: 25.1 m; ERES: 30.8 m; average across three estimates: 31.2 m</td>
<td>NA</td>
<td>33 m (triangulation)</td>
</tr>
</tbody>
</table>

Data are presented as value (95% CI), unless otherwise stated. COPD: chronic obstructive pulmonary disease; NA: not available; RCT: randomised controlled trial; SEM: standard error of the baseline measurement; ERES: empirical rule effect size; CRM: coefficient of repeated measurements; ERES: empirical rule effect size; CRM: coefficient of repeated measurements; SEM: standard error of the baseline measurement; ERES: empirical rule effect size.
34 m (95% CI 25–43 m) with a corresponding effect size of 0.38 (median standard deviation of change in control group of 89 m).

In summary, 6MWD appears to be responsive to treatment effects in patients with COPD, ILD and PAH, although the majority of the data relate to rehabilitation and few trials were available to evaluate responsiveness to pharmacological treatments. The standardised response means (mean change/standard deviation of change) ranged from 0.2 to 1.2 (table 6). Single trials were often underpowered to show statistically significant results, which may be attributable to a relatively large variability in 6MWD in patients with chronic respiratory disease. Trials either need to include a large number of patients (several hundred) or show a large effect (>100 m) to detect a statistically significant signal. The meta-analyses confirmed this as they showed highly statistically significant results with effects between 34 and 78 m and effect sizes between 0.38 and 1.07.

ISWT

The ISWT was first described in 1992 to assess exercise capacity in patients with COPD [203]. Compared with the 6MWT there is a smaller but emerging literature on the conduct and properties of the test in chronic respiratory disease. The ISWT offers a different protocol to the 6MWT as it is both incremental and externally paced, conducted around a 10-m course.

The search returned a total of 224 records, after removal of duplicates. After removal of records that were not related to the ISWT or individuals with respiratory disease, 168 records remained. Three additional references were identified in the updated search in September 2013.

Reliability of the ISWT

There are a number of studies (n=7) describing the reliability of the ISWT (table 7). The majority of the data are from patients with COPD. One study reported an ICC of 0.88 with a 95% CI of 0.83–0.92 [205]. Following subgroup analysis on sex, age and severity, the ICC remained high (range 0.80–0.93). A second study reported inter-subject variability as 88.9% [7]. These two studies suggest that the association between test–retest ISWT distances is strong, with the majority of variability being attributable to between-subject differences rather than within-subject differences.

Learning effect on ISWT

There is a significant learning effect between the first two ISWTs, with mean differences ranging 9–25 m (pooled mean difference 20 m; n=640 subjects) (table 7). This is of sufficient magnitude to recommend two tests at first exposure. The learning effect may be attenuated over subsequent repeats of the test [7].
<table>
<thead>
<tr>
<th>First author [ref.]</th>
<th>Diagnosis</th>
<th>Mean age years</th>
<th>Severity</th>
<th>Methods</th>
<th>ICC</th>
<th>Mean difference m</th>
<th>Coefficient of repeatability</th>
<th>Least square differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arnardóttir [204]</td>
<td>COPD</td>
<td>64</td>
<td>Moderate to severe</td>
<td>Repeat test 7 days later in subgroup, no practice tests, test 1 versus test 2</td>
<td></td>
<td>9 ± 38</td>
<td>74.5*</td>
<td></td>
</tr>
<tr>
<td>Campo [205]</td>
<td>COPD</td>
<td>68.6</td>
<td>Mild to severe</td>
<td>Test-retest, 1 week apart (practice walks performed), operators not blinded, 10 min between practice and a test</td>
<td>0.88 [0.83–0.92]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dyer [206]</td>
<td>Chronic airflow limitation</td>
<td>76</td>
<td>Mild to moderate</td>
<td>Test 1 versus test 3, within 2 weeks, Ino practice walk on day 1, operators not blinded</td>
<td></td>
<td>6 ± 17.89</td>
<td>35.1 m</td>
<td></td>
</tr>
<tr>
<td>Dyer [207]</td>
<td>Mixed chronic lung diseases</td>
<td>70</td>
<td>Mild to severe</td>
<td>Repeatability of test 1 versus test 2 (i.e. practice versus actual) on separate days</td>
<td></td>
<td>23.5 ± 47.3</td>
<td>92.7*</td>
<td></td>
</tr>
<tr>
<td>Eiser [7]</td>
<td>COPD</td>
<td>69</td>
<td>Moderate to severe</td>
<td>Nine ISWTs performed over 3 weeks [3 tests-day⁻¹], difference within and between days</td>
<td>0.89</td>
<td></td>
<td>Inter-subject 10 m, intra-subject 11 m, inter-subject variability 88.9%, intra-subject variability 9.2%</td>
<td></td>
</tr>
<tr>
<td>Hill [48]</td>
<td>COPD</td>
<td>66.5</td>
<td>Mild to severe</td>
<td>Practice and actual performed on same day, difference calculated as learning effect</td>
<td></td>
<td>25 ± 35</td>
<td>68.6*</td>
<td></td>
</tr>
<tr>
<td>Mckegough [208]</td>
<td>COPD</td>
<td>72</td>
<td>Moderate to severe</td>
<td>Test 1 versus test 2 (pre-exercise training), test 3 versus test 4 (post-exercise training), no practice walks</td>
<td></td>
<td></td>
<td>Test 1 versus 2: 20 [9–31], test 3 versus 4: 19 [6–30]</td>
<td></td>
</tr>
<tr>
<td>Zainulldin [209]</td>
<td>COPD</td>
<td>70</td>
<td>Mild to severe</td>
<td>Two tests on the same day, 30 min rest between</td>
<td></td>
<td>18 [14–22]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data are presented as mean ± SD or value [95% CI], unless otherwise stated. ICC: intra-class correlation coefficient; COPD: chronic obstructive pulmonary disease. *: coefficient of repeatability calculated from published data as SD of differences × 1.96.
However, there is a suggestion that over longer periods (>8 weeks) the learning effect may reset to a level similar to that recorded at first test exposure [208]. This observation requires further corroboration. There are no data available that indicate whether the repeatability of the test varies with baseline performance; therefore, it is important that two tests are performed regardless of functional capacity.

Reliability of other measures taken during the ISWT
Two studies evaluated reliability of end-test \( \text{SpO}_2 \) and HR measures for ISWTs repeated within the same day [48, 208]. These variables appear to repeat well with only small differences reported (mean difference: HR 0–4% and \( \text{SpO}_2 \) 0%). Two studies report measurements of symptom scores recorded during repeat ISWTs [7, 208]. The modified Borg dyspnoea score may be more reproducible than the visual analogue score [7]. The mean ± sd difference between Borg fatigue scores was 0 ± 2 units [208].

Validity of the ISWT
The aim of developing the ISWT was to derive a field-based exercise test that could reflect the response provoked by a CPET. Validation of the ISWT, therefore, has focused largely on direct comparison with incremental exercise tests conducted in the laboratory. Studies that met the inclusion criteria are shown in table 8. The data show a strong relationship between \( V'\text{O}_2 \) or work rate on CPET and ISWT (\( r = 0.75–0.88 \)) [35, 38, 204, 210], with no difference in measured \( V'\text{O}_2 \) between the tests [48, 209, 211]. Other physiological parameters (\( V'\text{CO}_2 \), \( V'E \) and derivatives) were generally lower during the ISWT (table 8). Minute-by-minute analysis of both the ISWT and CPET demonstrated a linear response in \( V'\text{O}_2\text{peak} \), suggesting that the two tests provoke a similar cardiopulmonary response [48]. The current literature suggests that the ISWT is a valid measure of cardiopulmonary exercise capacity in COPD. It should also be noted that most of the data are obtained from individuals with COPD with moderate to severe disease and usually confined to small study populations.

The ISWT has been shown to be related to other outcomes in people with chronic respiratory disease. In COPD, demographics such as age, sex and lung function are weakly related to ISWT distance (\( r = 0.29–0.38 \)) [206, 213, 214]. The distance walked in the ISWT correlates moderately with the sniff nasal inspiratory pressure in patients with lung cancer (\( r = 0.42 \)) [215]. The ISWT distance also has a moderate correlation with quadriceps muscle strength in COPD (\( r = 0.47 \)) [214] and lung cancer (\( r = 0.39 \)) [215]. Two studies described a weak to moderate relationship of ISWT distance with objectively measured levels of physical activity in daily life (\( r = 0.17–0.58 \)) [216, 217].

Technical factors affecting ISWT performance
There were no data available on track layout. This would be expected, as the course to conduct the ISWT is clearly defined and not subject to variation. No data were available for differences that might be associated with different operators or the impact of encouragement. One paper examined seasonal variation and found no important differences in performance, although these tests were conducted indoors [218]. Compared with walking with a compressed air cylinder, oxygen supplementation during ISWT increased distance by 33 m (95% CI 18–47 m) in patients who were hypoxic at rest or had exercise-induced desaturation [219]. However, there was no difference when the oxygen-supplemented walk was compared with the baseline walk without a cylinder (mean difference 4 m, 95% CI -11–18 m) [219], highlighting the importance of whether the cylinder is supported by the individual or the clinician/researcher.

Relationship of ISWT to clinical outcomes
The ISWT may be a useful marker of clinical outcomes, although the data are reported from a small number of studies (table 9). The ISWT appears to be a significant predictor of survival and re-admission in people with COPD, those with a lower performance having a greater risk of admission [220, 222, 223]. A threshold of \( <170 \) m has been suggested to be associated with higher mortality [222]; however, these data have not yet been replicated.

Measurements and reporting for ISWT
Complications associated with the performance of the ISWT are unusual. We were unable to find specific published reports of complications or adverse events associated with performance of the ISWT in clinical trials. The safety of the test has only been described in cardiac disease [224], with no adverse events reported. One study that included 50 participants with mild-to-severe COPD and who had experienced no exacerbation in the previous 6 weeks reported that the mean ± sd change in saturation was -4.6 ± 6.2% after the ISWT [225].

Performance on the ISWT is defined by the distance achieved. This is expressed in 10-m increments. The level and number of shuttles achieved can also be reported as outcomes for ISWT. Other outcomes reported

---

ERS/ATS SYSTEMATIC REVIEW | S.J. SINGH ET AL.

DOI: 10.1183/09031936.00150414
### TABLE 8 Relationship of incremental shuttle walk test (ISWT) to other exercise tests in people with chronic respiratory disease

<table>
<thead>
<tr>
<th>First author [ref.]</th>
<th>Diagnosis</th>
<th>Mean age years</th>
<th>Disease severity</th>
<th>Subjects n</th>
<th>Comparison test</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arnardottir [204]</td>
<td>COPD</td>
<td>64</td>
<td>Moderate to severe</td>
<td>93*</td>
<td>CPET</td>
<td>r=0.88, p&lt;0.0001 ISWT distance × body weight versus Wpeak</td>
</tr>
<tr>
<td>Campo [205]</td>
<td>COPD</td>
<td>68.6</td>
<td>Mild to severe</td>
<td>30</td>
<td>6MWT, 12MWT, CPET</td>
<td>r=0.82 ISWT versus 6MWT; r=0.74 ISWT versus 12MWT; r=0.68 estimated V̇O₂ versus V̇O₂max</td>
</tr>
<tr>
<td>Hill [48]</td>
<td>COPD</td>
<td>67</td>
<td>Mild to severe</td>
<td>24</td>
<td>CPET</td>
<td>No difference in V̇O₂ between ISWT and CPET (mean 1227 versus 1186 mL·min⁻¹)</td>
</tr>
<tr>
<td>Luxton [38]</td>
<td>COPD</td>
<td>65</td>
<td>Moderate to severe</td>
<td>25 (22*)</td>
<td>CPET</td>
<td>r=0.75 ISWT distance versus CPET Wpeak; r=0.85 ISWT distance × weight versus CPET Wpeak; significant mean difference: V̇CO₂ (p=0.001), V̇E (p=0.006), SpO₂ (p&lt;0.001), RPE (p&lt;0.001), Borg dyspnoea (p=0.001)</td>
</tr>
<tr>
<td>Onorati [210]</td>
<td>COPD</td>
<td>70</td>
<td>Moderate</td>
<td>13</td>
<td>CPET</td>
<td>r=0.92 ISWT V̇O₂peak versus V̇O₂peak cycling; r=0.86 ISWT V̇O₂peak versus ISWT distance; r=0.72 ISWT distance versus V̇O₂peak cycling</td>
</tr>
<tr>
<td>Palange [211]</td>
<td>COPD</td>
<td>71</td>
<td>Moderate</td>
<td>9 (7*)</td>
<td>CPET</td>
<td>Mean V̇O₂peak 1145 versus 1025 mL·min⁻¹ (NS) for CPET versus ISWT; significant mean difference: V̇CO₂peak (p&lt;0.001), RER (p&lt;0.001), HR/V̇O₂ (p&lt;0.05), V̇E/V̇O₂peak (p&lt;0.001), V̇E/V̇CO₂peak (p&lt;0.001), lactate (p&lt;0.001), dyspnoea (p&lt;0.05), leg effort (p&lt;0.05)</td>
</tr>
<tr>
<td>Turner [35]</td>
<td>COPD</td>
<td>64</td>
<td>Moderate to severe</td>
<td>20</td>
<td>CPET</td>
<td>r=0.73 ISWT distance versus ISWT V̇O₂peak; r=0.79 ISWT distance versus CPET V̇O₂peak</td>
</tr>
<tr>
<td>Win [212]</td>
<td>Lung cancer</td>
<td>69</td>
<td>Operable</td>
<td>125</td>
<td>Treadmill CPET</td>
<td>r=0.67 V̇O₂peak CPET versus ISWT distance</td>
</tr>
<tr>
<td>Zainuldin [209]</td>
<td>COPD</td>
<td>70</td>
<td>Mild to severe</td>
<td>34</td>
<td>CPET</td>
<td>No difference in V̇O₂peak CPET versus ISWT: mean (95% CI) difference 0.024 (-0.036–0.085) L·min⁻¹; on ISWT, significantly lower V̇CO₂, V̇E/MVV, SpO₂, symptoms</td>
</tr>
</tbody>
</table>

All ISWTs had a cohort study design. COPD: chronic obstructive pulmonary disease; CPET: cardiopulmonary exercise test; Wpeak: peak work rate; 6MWT: 6-min walk test; 12MWT: 12-min walk test; V̇O₂: oxygen uptake; V̇O₂max: maximum oxygen uptake; V̇CO₂: carbon dioxide production; V̇E: minute ventilation; SpO₂: arterial oxygen saturation measured by pulse oximetry; RPE: rate of perceived exertion; V̇O₂peak: peak oxygen uptake; NS: nonsignificant; V̇CO₂peak: peak carbon dioxide production; RER: respiratory exchange ratio; HR: heart rate; MVV: maximum voluntary ventilation. *: convenience sample; †: completers.

include HR and blood pressure responses, SpO₂, V̇O₂, V̇E, respiratory rate, inspiratory capacity, dyspnoea and leg fatigue, monitored before and during the test. Despite systematic effort, we were unable to find published reports showing how different monitoring protocols affect detection of changes in performance or clinical outcomes during shuttle walking tests in patients with chronic respiratory diseases.

Four studies demonstrated the utility of SpO₂ during the ISWT. Significant desaturation during the test has been reported [219], which may be greater than during the 6MWT [225]. Two studies recorded desaturation following the ISWT compared to a CPET [35, 226]. One study reported improvement in SpO₂ during walking exercise [35, 219], it seems reasonable to continuously monitor SpO₂ and HR during the test, with recording of the lowest SpO₂ and the peak HR.

**Reference equations for ISWT**

There have been three papers describing reference values for the ISWT. In total, 411 healthy adults have been recruited, with two studies conducted in South America contributing the most data [227, 228]. The largest study has a median age of 50 years, which is at least a decade younger than that observed in pulmonary rehabilitation. The reference equation from this group included age, sex and body mass index (BMI) as the important variables, explaining 71% [228] and 50% [227] of the variability. The third study,
from a single centre in the UK [229], included measures of strength and physical activity. The addition of these measures to FEV₁ and BMI did not improve the reference equation, which explained only 50% of the variance. This study did not find a sex difference in performance of the ISWT once corrected for age.

**Identifying meaningful change in the ISWT**

Minimal important difference of the ISWT

Based on the titles and abstracts, only two studies, conducted in participants with confirmed COPD, were retained [230, 231]. The first report provided a MID estimate for the ISWT in 372 patients with moderate to very severe airflow limitation and a mean ±SD baseline ISWT distance of 168.5 ± 114.6 m undertaking 7 weeks of pulmonary rehabilitation [230]. The MID estimate was determined with an anchor-based method, where change in ISWT distance was related to the patients’ perception of change from baseline, rated on a 5-point Likert scale. A mean change of 47.5 m (approximately five shuttles) was associated with the rating “slightly better”, while a mean change of 78.7 m (approximately eight shuttles) was associated with the next rating (“better”). The authors concluded that the MID for the ISWT was 47.5 m. Similar results were reported in a study of 261 patients with COPD undertaking pulmonary rehabilitation [231]. A mean change of 45.7 m was associated with the rating “a little better”.

**Responsiveness of the ISWT**

Three studies were explicitly designed to evaluate the responsiveness to treatment of the ISWT, following bronchodilation [206], ambulatory oxygen therapy [219] or pulmonary rehabilitation [232]. Characteristics of the included studies can be seen in table S32. Overall, the studies suggested that the ISWT was responsive to the interventions tested, with standardised response means (mean change/standard deviation of change) ranging from 0.72 to 1.55 (table 10).

Two systematic reviews aimed to study the effects of pulmonary rehabilitation after acute exacerbation on subsequent hospital admissions, mortality, HRQoL and exercise capacity compared with usual care in
<table>
<thead>
<tr>
<th>First author [ref.]</th>
<th>Studied population</th>
<th>Subjects n</th>
<th>Mean age years</th>
<th>FEV1</th>
<th>Field tests</th>
<th>Mean field test distance or time at baseline</th>
<th>Mean change in walking distance or time</th>
<th>SRM</th>
</tr>
</thead>
<tbody>
<tr>
<td>DYER [206]</td>
<td>Elderly with CAL</td>
<td>50</td>
<td>76</td>
<td>72%</td>
<td>ISWT</td>
<td>177.7 m</td>
<td>13.2±4.8% (1.28.5–61.3%)</td>
<td>NA</td>
</tr>
<tr>
<td>PEPIN [105]</td>
<td>COPD</td>
<td>17</td>
<td>65</td>
<td>56%</td>
<td>ESWT</td>
<td>405 s</td>
<td>164±177 s§</td>
<td>0.93</td>
</tr>
<tr>
<td>PEPIN [106]</td>
<td>COPD GOLD stage II–III</td>
<td>14</td>
<td>64</td>
<td>50%</td>
<td>ESWT</td>
<td>420 m</td>
<td>144±219 m§</td>
<td>0.66</td>
</tr>
<tr>
<td>Brouillard [107]</td>
<td>COPD</td>
<td>20</td>
<td>65</td>
<td>52%</td>
<td>ESWT</td>
<td>373 s</td>
<td>117±208 s§</td>
<td>0.56</td>
</tr>
<tr>
<td>Sandland [219]</td>
<td>Severe hypoxaemic COPD</td>
<td>41</td>
<td>71</td>
<td>0.85 L</td>
<td>ESWT</td>
<td>192.4 m</td>
<td>33.2±46.3 m§</td>
<td>0.72</td>
</tr>
<tr>
<td>Revill [233]</td>
<td>COPD with exertional desaturation</td>
<td>23</td>
<td>67</td>
<td>33%</td>
<td>ESWT</td>
<td>273 s</td>
<td>91 [43–139] s*</td>
<td>NA</td>
</tr>
<tr>
<td>Eaton [234]</td>
<td>COPD</td>
<td>20</td>
<td>71</td>
<td>0.95 L</td>
<td>ESWT</td>
<td>313 m</td>
<td>302±387 m§</td>
<td>0.78</td>
</tr>
<tr>
<td>Leung [232]</td>
<td>COPD GOLD stage I–IV, walking training group</td>
<td>17</td>
<td>71</td>
<td>56%</td>
<td>ESWT</td>
<td>397 s</td>
<td>439±346 s*</td>
<td>1.27</td>
</tr>
<tr>
<td></td>
<td>COPD GOLD stage I–IV, cycling training group</td>
<td>15</td>
<td>72</td>
<td>53%</td>
<td>ESWT</td>
<td>375 s</td>
<td>45±29 m*</td>
<td>1.55</td>
</tr>
</tbody>
</table>

Data are presented as mean±standard error of the measurement (95% CI), unless otherwise stated. FEV1: forced expiratory volume in 1 s; SRM: standardised response mean (mean change/SD of change); CAL: chronic airflow limitation; NA: not available and/or incalculable; COPD: chronic obstructive pulmonary disease; GOLD: Global Initiative for Chronic Obstructive Lung Disease; 6MWT: 6-min walk test. §: p<0.05; *: p-value NA.
COPD [235, 236]. The meta-analysis, including 128 COPD patients, reported a weighted mean difference of 64.35 m (95% CI 41.28–87.43 m) in ISWT distance, favouring rehabilitation.

In summary, very few systematic reviews have reported changes in ISWT after interventions of known effectiveness, but studies designed specifically to assess responsiveness to treatment suggest that the ISWT is responsive.

ESWT
The ESWT is a test of endurance capacity, first described in 1999 [237]. It is externally paced and is performed along the same course as the ISWT. The pace of the ESWT is traditionally calculated at a predefined percentage of peak performance on the ISWT (e.g. 70–85% estimated \( V_{\text{O2peak}} \)) [238]. However, recent data have indicated that the speed can also be accurately derived using 85% of walking speed on the ISWT [239].

The search returned a total of 60 records after removal of duplicates. After removal of studies unrelated to ESWT or respiratory disease, a total of 56 records remained.

Reliability of the ESWT
Three studies have examined the reliability of the ESWT, all in people with COPD (table 11) [48, 208, 240]. The differences between tests repeated on the same day were generally small and statistically nonsignificant (pooled mean difference +26 s). Two tests do not appear to be necessary, although it is acknowledged that the number of studies is limited. Measurements of HR, \( \text{SpO2} \) and modified Borg dyspnoea scale appear to repeat well during the test [48, 240]. The reliability of the test in other chronic respiratory diseases has not been studied.

Validity of the ESWT
There have been no studies describing the validity of the ESWT compared with laboratory-based exercise tests within the search timeframe. No studies have reported on the relationship with ESWT and other clinical outcomes. One study in COPD has reported that FEV1 is an independent predictor of ESWT time; however, it did not correlate with either muscle strength or mass [214].

Technical factors affecting ESWT performance
No data were available for differences that might be associated with different operators or the impact of encouragement. Both supplemental oxygen or heliox and oxygen may have an impact on ESWT performance [219, 233, 241, 242]. The studies reflect a difference in conduct, regarding whether the cylinder was carried by the participant or the operator. The changes in performance with oxygen compared with air range from 70 to 174 s [219, 242] or 32−76% of walking time [241]. An additional study reported large improvements in ESWT distance with application of supplemental oxygen (mean increase 275 m, 95% CI 197−352 m) [243]. However, all participants were known oxygen “responders” (increase in ESWT of \( \geq 10\% \) with oxygen), the oxygen cylinder was carried by the operator and the study was not blinded.

### Table 11 Reliability of the endurance shuttle walk test (ESWT) in chronic obstructive pulmonary disease

<table>
<thead>
<tr>
<th>First author [ref.]</th>
<th>Mean age [years]</th>
<th>Disease severity</th>
<th>Method of repeatability/reliability</th>
<th>Result</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCKEOUGH [208]</td>
<td>72</td>
<td>Moderate</td>
<td>Test 1 versus test 2 (pre-exercise training), test 3 versus test 4 (post-exercise training)</td>
<td>Test 1 versus test 2 mean difference 2 s ((p=0.95)), test 3 versus test 4 mean difference 44 s ((p=0.07))</td>
<td>Repeatability pre- and post-8–12 weeks cycling or tai chi training, ESWT at 85% of best ISWT</td>
</tr>
<tr>
<td>REVILL [240]</td>
<td>68</td>
<td>Severe</td>
<td>Test 1 versus test 2</td>
<td>Test 1 versus test 2 mean difference 12 s ((95% \text{CI } -3–28 s))</td>
<td>ESWT at 85% on best ISWT, same Borg dyspnoea score at the end of both tests in 77% of sample, Bland–Altman shows limits of agreement (\pm 100 s)</td>
</tr>
<tr>
<td>HILL [48]</td>
<td>67</td>
<td>Mild to severe</td>
<td>Test 1 versus test 2 on same day</td>
<td>Test 1 versus test 2 mean difference 50 ± 83 s, n=18</td>
<td>Mean differences between tests: heart rate 2 bpm, ( \text{SpO2} ) 1%</td>
</tr>
</tbody>
</table>

ISWT: incremental shuttle walk test; \( \text{SpO2} \): arterial oxygen saturation measured by pulse oximetry.
a result, this degree of increase in ESWT performance with oxygen should not be inferred in all patients with COPD.

**Relationship of ESWT to clinical outcomes**
There were no studies that evaluated the association of the ESWT and survival or hospitalisation.

**Measurements and reporting for ESWT**
There have been no reports of adverse events associated with performing the ESWT in clinical practice or in the context of clinical trials.

The ESWT, like all tests of endurance capacity, should be reported as time (in seconds). Additional measures can be recorded. Three studies demonstrated the potential for the use of SpO2 to record exertional desaturation during ESWT, either breathing air or for assessment of ambulatory oxygen [219, 233, 242]. One study reported greater desaturation with ESWT compared with 6MWT [233]. No study reported the utility of HR measurement.

**Reference equations for ESWT**
No reference equations have been published for the ESWT.

**Identifying meaningful change in the ESWT**

**Minimal important difference of the ESWT**
One study aimed to determine the MID for the ESWT in patients receiving pulmonary rehabilitation (n=132, mean ± SD FEV1 48 ± 22% predicted, baseline ISWT distance 203 ± 129 m) or bronchodilation (n=69, FEV1 50 ± 12% predicted, baseline ISWT distance 483 ± 148 m) [104]. A valid MID estimate could not be obtained from the pulmonary rehabilitation data. Using an anchor-based method, the bronchodilation data indicated that a change of 65 s (95% CI 45–85 s) or 85 m (95% CI 60–115 m) was likely to be perceivable to patients.

**Responsiveness of the ESWT**
Seven studies were specifically designed to examine the responsiveness of the ESWT to bronchodilation [105–107], oxygen therapy [219, 233] and pulmonary rehabilitation [232, 234]. The responsiveness of the ESWT was moderate to high, with standardised response means (mean change/standard deviation of change) ranging from 0.52 to 1.27.

**Discussion**
Studies included in this review consistently showed that the 6MWD is a highly reliable measure in people with chronic lung disease (table S2). In people with COPD, there is an average improvement of 26 m between the first and second tests (table 1). The confidence interval around this estimate is narrow [8] and the estimate is unlikely to be altered substantially by future studies. This learning effect is large enough to be clinically important when the 6MWT is used to evaluate response to treatment or change over time, as failure to consider the learning effect may lead to erroneous interpretation of changes in 6MWD. As a result, consideration should be given to performing two 6MWTs where the test is being used to evaluate longitudinal changes.

This systematic review has demonstrated construct and criterion validity for the 6MWT in patients with chronic respiratory disease. Relationships are strongest with measures of maximal exercise performance and physical activity (tables 2 and 4), indicating that the test is primarily a test of physical capacity. However, the proportion of variance explained is modest, suggesting that these measures are not interchangeable. The 6MWT has historically been considered to be a test of submaximal exercise capacity [2]; however, direct comparisons of the physiological demands of the 6MWT and CPET reveal that, in patients with chronic respiratory disease, measures of peak exercise performance are similar between the tests. Nevertheless, the 6MWT has lower ventilatory requirements (table 3). This body of evidence lends support to the conceptualisation of the 6MWT as a test of functional exercise capacity.

The 6MWD has a clear association with clinical outcomes across a range of chronic respiratory diseases, with the majority of studies identified reporting statistically significant associations of 6MWD with mortality and/or hospitalisation (tables S23–S26 and fig. 1). It is possible that some studies, where 6MWD was not statistically significantly associated with mortality or hospitalisation, were not published. Thus, we cannot exclude some publication bias and the possibility that our summary of the evidence overestimates the strength of association of 6MWD with mortality and/or hospitalisation. Also, it was beyond the scope of this document to identify thresholds for 6MWD to categorise patients according to their risk for the...
outcomes. The studies included here addressed the question of the presence and absence of an association of 6MWD with mortality and hospitalisation, rather than finding optimal cut-offs for assessing prognosis.

We identified 17 reference equations for 6MWD, giving rise to wide variation in predicted walk distance. The factors contributing to this variation are likely to include methodological variations as well as local differences in test performance. As a result of this variation, reference equations generated in a local population, using a protocol which closely resembles that used when the reference equations were generated, should be applied where possible. When assessing change in 6MWD, the distance in metres should be used. As well as allowing assessment of meaningful change, this approach also allows the 6MWD to be used in estimation of prognosis in patients with clinically important respiratory disease.

Since the last ATS statement on the 6MWT published in 2002, there have been a number of new publications on the MID for the 6MWD (fig. 2). Encouragingly, recent papers have produced similar estimates, in the range 25–35 m, regardless of the methodology used. The strengths of the current evidence are that they come from a wide variety of clinical sites involving over 5600 participants with moderate to severe chronic respiratory disease and both anchor-based and distribution-based methods were used to estimate the MID. A limitation is the relatively narrowly distributed range of age and 6MWD, which may limit the applicability of the MID estimates to younger adult patients with moderate to severe lung disease. Another limitation is the lack of anchors that correlate strongly with 6MWD, so that even anchor-based estimates need to be interpreted with caution.

It is difficult to investigate whether the MID for 6MWT varies according to disease severity or other patient characteristics, because stratifying patients leads to smaller standard deviations and because correlations of changes in 6MWD with potential anchors often become too small to justify an anchor-based approach. One study reported a MID of 14% of baseline; however, this was less sensitive than the MID in metres [84]. On the basis of current data, a MID based on the absolute change in metres should be used. There is currently little evidence to suggest that the MID varies according to patient characteristics, including the type of chronic respiratory disease.

The relatively recent development of the ISWT and ESWT is reflected by the smaller number of studies identified in our search, in comparison with the number studying 6MWT. The available data indicated that both tests are valid, reliable and responsive (tables 7–11). Most research to date has focused on the use of the ISWT and ESWT in COPD, with more studies needed in other chronic respiratory conditions. There are few studies that have investigated the relationship of ISWT or ESWT to clinical outcomes, relationships that would reinforce the clinical importance of these measures. However, the small number of existing studies has shown that the ISWT is a significant predictor of survival and re-admission [220, 222, 223]. There was little information about the effects of methodological variations on performance of the ISWT and ESWT, probably because these tests are externally paced and performed in accordance with very stringent protocols. This should be considered to be a strength of these tests. Further data are required for reference values of the ISWT, to ensure that suitable equations are available for a wider proportion of the global community.

The measurement properties of the field tests outlined in this review indicate that all are suitable for use in patients with chronic respiratory disease. The choice of an individual test may depend on the circumstances and purpose of the test. For instance, the 6MWD shows strong associations with important long-term outcomes (fig. 1), so may be useful in circumstances where disease staging or estimation of prognosis is required. The availability of consistent MID estimates to which longitudinal change can be compared is also an advantage of this test (fig. 2). However, the shuttle tests may be more sensitive to the effects of common interventions, including bronchodilators [106, 233, 234]. The externally paced nature of the ISWT and ESWT may be an advantage in circumstances where methodological variation is a concern (e.g. multiple sites or operators). Finally, where the larger space requirements of the 6MWT preclude its use, the ISWT and ESWT may be useful alternatives.

The strengths of this review include the extensive body of evidence that was evaluated in a systematic fashion. Both the 6MWT and the shuttle tests were included, allowing comparison of the measurement properties and utility of these different field tests. Limitations include the restriction of the search to the year 2000 onwards, which presents the possibility that earlier, pertinent literature was not included, and the consideration of papers only where there was an abstract published in English. The scope of the review was carefully restricted to adults with respiratory disease; the measurement properties described here cannot be assumed to apply to other populations with chronic disease, or to children.

In conclusion, this systematic review has found robust evidence to support the reliability, validity and responsiveness of the 6MWT, ISWT and ESWT in people with chronic respiratory disease. There is a substantial body of evidence that supports the conceptualisation of the 6MWT as a test of functional exercise capacity for individuals with a range of chronic respiratory conditions. While studies examining the
ISWT and ESWT are more limited in number, the strong measurement properties of these tests are already emerging, particularly with regard to responsiveness to common interventions. There is a sufficient body of evidence to support the use of the 6MWT, ISWT and ESWT in research and in clinical practice for adults with chronic respiratory diseases.

Acknowledgements
The authors’ affiliations are as follows. Sally J. Singh: Centre for Exercise and Rehabilitation Science, University Hospitals of Leicester NHS Trust, Leicester, UK; Milo A. Puhun: Institute for Social and Preventive Medicine, University of Zurich, Zurich, Switzerland; Vasileios Andrianopoulos: Dept of Research and Education, CIRO+ Centre of Expertise for Chronic Organ Failure, Horn, The Netherlands; Nidia A. Hernandez and Fabio Pitta: Laboratory of Research in Respiratory Physiotherapy, Dept of Physiotherapy, Universidade Estadual de Londrina, Londrina, Brazil; Kate E. Mitchell: Centre for Exercise and Rehabilitation Science, University Hospitals of Leicester NHS Trust, Leicester, UK; Catherine J. Hill: Institute for Breathing and Sleep, Austin Health, Melbourne, Australia, and Dept of Physiotherapy, Austin Health, Melbourne, Australia; Annemarie L. Lee: Dept of Physiotherapy, Alfred Health, Melbourne, Australia, and Institute for Breathing and Sleep, Austin Health, Melbourne, Australia; Carlos Augusto Camillo and Thierry Troosters: Faculty of Kinesiology and Rehabilitation Sciences, Dept of Rehabilitation Sciences, Katholieke Universiteit Leuven, Leuven, Belgium; Martijn A. Spruit: Dept of Research and Education, CIRO+ Centre of Expertise for Chronic Organ Failure, Horn, The Netherlands, and REVAL – Rehabilitation Research Center, BIOMED – Biomedical Research Institute, Faculty of Medicine and Life Sciences, Hasselt University, Diepenbeek, Belgium; Brian W. Carlin: Drexel University School of Medicine, Pittsburgh, PA, US; and Sleep Medicine and Lung Health Consultants, Pittsburgh, PA, USA; Jack Wanger: ATS Proficiency Standards for Pulmonary Function Laboratories Committee, Rochester, MN, USA; Véronique Pepin: Aix maladies chroniques, Centre de recherche de l’Hôpital du Sacré-Cœur de Montréal, Montréal, QC, Canada, and Dept of Exercise Science, Faculty of Arts and Science, Concordia University, Montréal, QC, Canada; Didier Saey: Centre de recherche, Institut Universitaire de cardiologie et de pneumologie de Québec, Québec, QC, Canada, and Faculté de médecine, Université Laval, Québec, QC, Canada; David A. Kaminsky: University of Pittsburgh College of Medicine, Pittsburgh, VT, USA; Meredith C. McCormack: Johns Hopkins University, Pulmonary and Critical Care Medicine, Baltimore, MD, USA; Neil Maclntyre: Duke University, Durham, NC, USA; Bruce H. Culver: Pulmonary and Critical Care Medicine, University of Washington, Seattle, WA, USA; Frank C. Sciurba: University of Pittsburgh School of Medicine, Pittsburgh, PA, USA; Susan M. Revill: Clinical Diagnostix Ltd, Derby, UK; Verónica Delafosse: Health Sciences Library, Caulfield Hospital, Alfred Health, Caulfield, Australia; Anne E. Holland: Dept of Physiotherapy, La Trobe University, Melbourne, Australia, Dept of Physiotherapy, Alfred Health, Melbourne, Australia, and Institute for Breathing and Sleep, Austin Health, Melbourne, Australia.

References

DOI: 10.1183/09031936.00150414


Spencer LM, Alison IA, 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: 224–228.


Brasil Santos D, de Assis Viegas CA. Correlação dos graus de obstrução na DPOC com lactato e teste de caminhada de seis minutos [Correlation of levels of obstruction in COPD with lactate and six-minute walk test]. Rev Port Pneumol 2009; 15: 11–25.


DOI: 10.1183/09031936.00150414


Chuang ML, Lin IF, Chen SP. Kinetics of changes in oxyhemoglobin saturation during walking and cycling tests in COPD. *Respir Care* 2014; 59: 353–362.


McKeough ZJ, Leung RW, Alison JA. Shuttle walk tests as outcome measures: are two incremental shuttle walk tests and two endurance shuttle walk tests necessary? Am J Phys Med Rehabil 2011; 90: 35–39.


