International Journal of COPD

Dovepress
open access to scientific and medical research

Open Access Full Text Article

REVIEW

Comparison of laboratory- and field-based exercise tests for COPD: a systematic review

This article was published in the following Dove Press journal:
International Journal of COPD
19 March 2015

Iain Fotheringham[1]
Georgina Meakin[1]
Yogesh Suresh Punekar[2]
John H Riley[2]
Sarah M Cockle[2]
Sally J Singh[3]

[1]Value Demonstration Practice, Oxford PharmaGenesis, Oxford, [2]GlaxoSmithKline, Uxbridge, [3]Department of Respiratory Medicine, University Hospitals of Leicester NHS Trust, Leicester, UK

Abstract: Exercise tests are often used to evaluate the functional status of patients with COPD. However, to the best of our knowledge, a comprehensive systematic comparison of these tests has not been performed. We systematically reviewed studies reporting the repeatability and/or reproducibility of these tests, and studies comparing their sensitivity to therapeutic intervention. A systematic review identified primary manuscripts in English reporting relevant data on the following exercise tests: 6-minute walk test (6MWT) and 12-minute walk test, incremental and endurance shuttle walk tests (ISWT and ESWT, respectively), incremental and endurance cycle ergometer tests, and incremental and endurance treadmill tests. We identified 71 relevant studies. Good repeatability (for the 6MWT and ESWT) and reproducibility (for the 6MWT, 12-minute walk test, ISWT, ESWT, and incremental cycle ergometer test) were reported by most studies assessing these tests, providing patients were familiarized with them beforehand. The 6MWT, ISWT, and particularly the ESWT were reported to be sensitive to therapeutic intervention. Protocol variations (eg, track layout or supplemental oxygen use) affected performance significantly in several studies. This review shows that while the validity of several tests has been established, for others further study is required. Future work will assess the link between these tests, physiological mechanisms, and patient-reported measures.

Keywords: 6MWT, 12MWT, COPD, walk test, repeatability, reproducibility, shuttle walk test, cycle ergometer test

Introduction

COPD is a leading cause of death worldwide, and the prevalence of the disease is projected to increase as the population ages and as exposure to risk factors, such as smoking, continues.[1–3] COPD is characterized by symptoms of breathlessness and reduced exercise capacity.[4,5] Decrements in exercise capacity can result in reduced ability to perform activities of daily living, and the resultant inactivity and sedentary lifestyle can further exacerbate exercise impairment (the COPD "vicious circle").[6]

In clinical practice, spirometry is recommended by the Global Initiative for Chronic Obstructive Lung Disease (GOLD) for the diagnosis of COPD.[5] However, the results of spirometry alone poorly predict disability and quality of life in patients with COPD[7] and correlate only weakly with dyspnea, exercise capacity, and health status.[8–10] Recent guidelines on the diagnosis and treatment of COPD indicate that assessment of disease severity is improved by using additional functional criteria such as exercise capacity.[4,5,11] Quantification of the degree of functional impairment is therefore important for the assessment of response to treatment and as an outcome for clinical trials.

There are a number of laboratory- and field-based tests currently used for the assessment of exercise capacity, including the 6- and 12-minute walk tests (6MWT

Correspondence: Iain Fotheringham
Oxford PharmaGenesis,
Tubney Warren Barn, Tubney,
Oxford, OX13 5QJ, UK
Tel +44 1865 390144
Fax +44 1865 390145
Email iain.fotheringham@pharmagenesis.com

submit your manuscript | www.dovepress.com
Dovepress
http://dx.doi.org/10.2147/COPD.S70518

and 12MWT, respectively), the incremental and endurance shuttle walk tests (ISWT and ESWT, respectively), the incremental and endurance cycle ergometer tests (ICET and ECET, respectively), and the incremental and endurance treadmill tests (ITT and ETT, respectively). However, there is no consensus about which test is the most appropriate for use in patients with COPD. These tests have different primary outcomes (eg, endurance time, distance, oxygen consumption) that may reflect different physiological parameters. It is therefore difficult to compare results across studies, limiting interpretation of the published literature in this field. Furthermore, the relative merits of different tests have not been established.

The systematic review presented here therefore evaluated evidence of the "repeatability" (defined as consistency of results when multiple tests are conducted on the same day) and the "reproducibility" (consistency of results when tests are conducted on different days) of the tests. The review also assessed the relative properties of the eight commonly used exercise tests and their sensitivity to therapeutic intervention (such as rehabilitative, pharmacological, or surgical procedures). In addition, the effect of protocol variations within each test was assessed across studies. When possible, results were placed in the context of available minimal clinically important difference (MCID) values, which have thus far been ascertained for the 6MWT,[12] ISWT,[13] ESWT,[14] and ICET.[12] Investigation of these factors will be useful in guiding test selection in clinical practice and for outcome measures in clinical trials. As these tests are often also used as interventions, evaluation of exercise testing modalities in patients with COPD will also inform the clinical development of optimal exercise rehabilitation strategies.

Methods

Search strategy

Literature searches were conducted using Ovid® (Ovid Technologies Inc., New York, NY, USA), incorporating Ovid Medline® (US National Library of Medicine, Bethesda, MD, USA), for the period from 1948 to January 22, 2013, Ovid Embase™ (Elsevier Inc., Philadelphia, PA, USA) for 1974 to January 22, 2013, and The Cochrane Library (John Wiley and Sons Ltd, Hoboken, NJ, USA) for 1962 to January 22, 2013 (see Tables S1–S3). Search strings were constructed to identify studies reporting primary data on the outcomes of the following exercise tests in patients with COPD: the 6MWT, 12MWT, ISWT, ESWT, ICET, ECET, ITT, and ETT. The full search strings are presented in the "Supplementary materials" section.

Study selection

Study selection followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines for performing a systematic literature review.[15] Review articles and studies not published in English were excluded using search-engine filters. Studies confounded by comorbidities (such as cancers, diabetes, and non-COPD respiratory-tract diseases) were excluded on review of title/abstract. The remaining studies were screened based on titles and abstracts, and full articles were reviewed when their relevance was unclear from the abstract. Screening was performed by a single author (GM) and records were initially reviewed by title/abstract; a full paper review was subsequently undertaken for publications that could not be excluded by title/abstract. Included records were verified by a second author (IF). A 30% random sample of excluded records was also reviewed by the second author (IF). Disagreements were settled by consultation with the remaining authors.

When reviewing abstracts or full papers, records were excluded if they were reviews, were not in the English language, studied patients with confounding comorbidities (eg, cancers or diabetes), did not use an exercise test as an outcome measure, or examined an intervention other than our interventions of interest (pulmonary rehabilitation, bronchodilator therapy, and lung-volume reduction surgery). Specific inclusion criteria included any definition of COPD (including emphysema- and bronchitis-specific studies); interventions were included only in our comparison of sensitivity and limited to pulmonary rehabilitation, bronchodilation, and lung-volume reduction surgery. Included test outcomes are outlined in the "Data abstraction" section. Following screening, studies were subsequently included for assessment if they reported data on:

- repeatability (studies reporting data from two or more performances of the same test[s] on the same day under the same conditions)
- reproducibility (studies reporting data from two or more performances of the same test[s] on different days under the same conditions)
- comparisons of sensitivity (studies reporting responses of two or more tests to the following therapeutic interventions: pulmonary rehabilitation, bronchodilator therapy, or lung-volume reduction surgery)
- protocol variations (studies reporting two or more performances of a test when protocol parameters have been modified).

Data abstraction

Data were primarily abstracted by a single author (GM) and reviewed by all co-authors. A randomly generated selection of 30% of all articles was reviewed by a second author (IF) for quality-control purposes.

The following outcomes of exercise tests were recorded: distance or stages achieved for the 6MWT, 12MWT, and ISWT; duration of exercise for the ESWT, ECET, and ETT; and the highest recorded volume of oxygen consumption (peak VO_2) and/or maximum workload (W_{max}) for the ICET and ITT. Articles merited inclusion in this review if they reported: outcomes of the specified tests when performed repeatedly under the same conditions, either on the same day (repeatability) or on different days (reproducibility); changes in response before and after therapeutic intervention (comparison of sensitivity); or effects of within-test variations in protocol (protocol variation).

Studies comparing the sensitivity of tests were also assessed for expression by the authors of preference for any specific test. When distances were reported in feet, values were converted to meters using standard conversion criteria stated by the International Bureau of Weights and Measures (0.3048 meters per foot). Within each publication, tests for which results are available are referred to as "test 1", "test 2", etc; occasions on which a test has been described by the authors, but results are not reported (such as for practice tests), are referred to as "familiarizations".

Results

Overview of identified studies

The search methodology used to identify relevant articles is summarized in Figure 1. Of 1,781 unique articles screened, 71 were ultimately deemed eligible for inclusion in this review.

Studies assessing the repeatability and reproducibility of tests

Clinical practice is influenced by factors such as the repeatability and reproducibility of exercise tests in patients with COPD. These factors have been extensively assessed for the 6MWT and 12MWT; data are more limited for the ISWT, ESWT, ITT, and ETT (23 references for the 6/12MWT; 12 for the IWST, ESWT, ITT, and ETT; seven for the ICET and ECET; this made 37 references in total owing to overlap of these categories). Table 1 summarizes the results of studies assessing the repeatability and reproducibility of the 6MWT and 12MWT; Table 2 focuses on the ISWT, ESWT, ETT, and ITT; and Table 3 presents data on the ICET and ECET.

Six studies presented repeatability data for the 6MWT (Table 1).[16–21] Of these, five reported a significant increase in 6MWT distance from the first to the second test;[16–20] the remaining study found no differences between results, though patients had been previously familiarized with the tests.[21] The three studies clearly reporting the results of three 6MWTs performed on the same day found that there

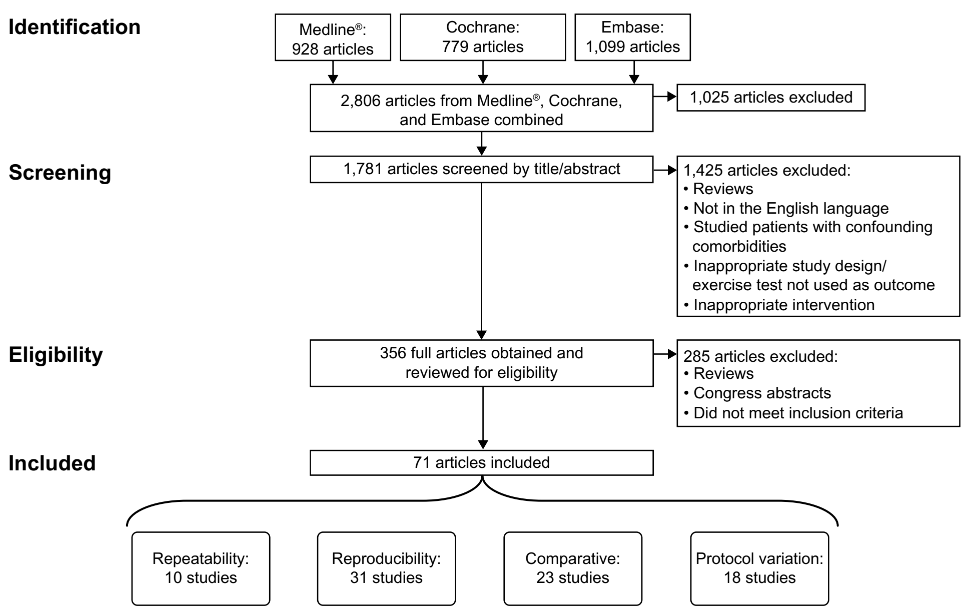

Figure 1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram detailing the identification and inclusion process of the articles. Some studies are included in more than one analysis category; consequently, the aggregate number of studies in the repeatability, reproducibility, comparative, and protocol variation groups adds up to more than 71.

Table 1 Repeatability and reproducibility of the 6- and 12-minute walk tests

Study	Patients, n (males, n)	Disease severity (mean ± SD FEV_1 [mean ± SD% predicted] or COPD grading, unless otherwise stated)	Comparison	Additional information
6-minute walk test				
Repeatability[a]				
Andersson et al[16]	47 (16)	1.20±0.49 (46.0±17.0)	First versus second test	Significant increase in distance from first to second test (Δ22 m [5.3%]) (ICC =0.94)
Bansal et al[17]	27 (15)	0.8±0.2 (38.1±14.3)	First versus second test	Significant increase in distance from first to second test, for linear (Δ14 m [3.5%]) and circuit (Δ12 m [3.0%]) track layouts
Eiser et al[18]	57 (30)	NR (35.0±12.0)	First versus second versus third test	Significant increase in distance from first to second test (Δ7 m [1.6%]); NSD from second to third test
Jenkins and Cecins[19]	245 (162)	1.06±0.5 (41.0±18.0)	First versus second test	Significant increase in distance from first to second test (Δ37 m [9.5%])
Stevens et al[20]	21 (9)	1.07±0.53 (NR)	First versus second versus third test	Hallway: significant increase in distance from first to second test (approximate Δ32.9 m [9.0%]); NSD in distance from second to third test Treadmill: significant increase in distance from first to second test (approximate Δ16.2 m [9.9%]); NSD in distance from second to third test
Vagaggini et al[21]	18 (15)	NR (48.0±14.0)	Familiarization then first versus second test	NSD in distance from first to second test after familiarization (Δ15 m [3.4%])
Reproducibility[b]				
Bansal et al[17]	27 (15)	0.8±0.2 (38.1±14.3)	First versus second test	NSD in distance from first to second day for straight (Δ12 m [2.9%]) and circular (25 m [6.0%]) track layouts
Behnke et al[26]	TG: 66 (51)	TG: 1.30±0.49 (41.9±13.9)	Familiarization then first versus mean of second and third tests	TG: significant increase in distance from first to mean of second and third test (Δ209.5 m [72.9%])
	CG: 22 (20)	CG: 1.45±0.51 (46.9±15.1)		CG: significant increase in distance from first to mean of second and third test (Δ57.5 m [18.9%])
Brooks et al[27]	18 (10)	1.0±0.3 (42.0±8.0)	First versus second test	NSD in distance from first to second test (Δ NR)
Chatterjee et al[28]	88 (37)	Median ± IQR FEV_1, L (mean ± SD%) 1.21±0.55 (52.0±19.4)	First versus second versus third test	NSD in distance from first to second test (Δ32 m [9.8%]); NSD in distance from second to third test (Δ4 m [1.1%])
Eiser et al[18]	57 (30)	NR (35.0±12.0)	First versus second versus third test	Significant decrease in mean distance from first to second day (Δ–9 m [–2.1%]); significant increase in mean distance from second to third day (Δ8 m [1.9%])
Kozu et al[29]	45 (38)	1.10±0.50 (45.0±12.0)	First versus second test	Significant increase in distance from first to second day (Δ13 m [4.3%])
Poulain et al[30]	10 (NR)	Moderate COPD	Familiarization then first versus second test	NSD in distance from first to second day after familiarization (Δ4.7 m [0.94%])
Rejeski et al[31]	209 (117)	1.57; 0.58 (57.1±17.0)	First versus second test	Significant increase in distance from first to second day (Δ22 m [4.6%]); significant correlation between tests (r=0.91)
Roomi et al[32]	15 (6)	NR (49.0±5.0)	First versus second test	NSD in distance from first to second day (Δ–1 m [0.0%])
Sciurba et al[33]	470 (287)	0.75±0.24 (26.3±7.4)	First versus second test	Significant increase in distance from first to second day (Δ20 m; baseline NR) (ICC =0.88)

Spencer et al[34]	44 (22)	NR (56.0±19.0)	First versus second test	Significant increase in distance from first to second day (Δ27 m [5.9%])
Troosters et al[35]	20 (NR)	1.36±0.46 (45.0±14.0)	First versus second test	NSD in distance from first to second day (Δ15 m [2.6%])
12-minute walk test				
Repeatability[a]				
O'Reilly et al[22]	10 (10)	Chronic bronchitis or radiological emphysema	Familiarization then first versus second test	NSD in distance from first to second test after familiarization (mean variation 3.1%; distance NR)
Reproducibility[b]				
Arnardóttir et al[36]	EIH group: 19 (8)	EIH group: 0.9±0.1 (34.6±2.4)	First versus second versus third test	EIH group: NSD in distance from first to second to third day
	Non-EIH group: 38 (19)	Non-EIH group: 1.0±0.1 (38.9±1.8)		Non-EIH group: significant increase in distance from first to second day (Δ92 m [12%]; baseline NR); significant increase in distance from second to third day (Δ28 m [4%]; baseline NR)
Beaumont et al[37]	12 (10)	1.03±0.27 (NR)	First versus second versus third test	Significant increase in distance from first to second day (Δ46 m [6.5%]); NSD in distance from second to third day (Δ37 m [4.9%])
Berger and Smith[38]	10 (10)	Moderate/severe COPD	Familiarization then first versus second test	NSD in distance from first to second test after familiarization (Δ13 m [2.2% using values from Beaumont et al[37]])
McGavin et al[39]	35 (35)	Chronic bronchitis	First versus second versus third test	Significant increase in distance from first to second day (Δ64 m [7.2%]); NSD in distance from second to third day (Δ41 m [4.3%])
Mungall and Hainsworth[40]	13 (13)	1.5±0.4 (NR); chronic bronchitis or radiological emphysema	First versus second versus third test (with three further tests NR in detail)	Significant increase in distance in third test compared with first and second; NSD distance on three subsequent tests (distances NR)
O'Reilly et al[22]	10 (10)	Chronic bronchitis or radiological emphysema	Familiarization then first versus second test	NSD in distance from first to second day after familiarization (Δ38 m [4.8%])
Swinburn et al[41]	17 (6)	0.77±0.30 (NR)	First versus second versus third versus fourth test	Progressive, significant increases in distance from first to fourth day (16%; baseline NR)

Notes: [a]Repeatability = similarity of test results when performed on the same day; [b]reproducibility = similarity of test results when performed on different days.

Abbreviations: CG, control group; CI, confidence interval; EIH, exercise-induced hypoxia; FEV_1, forced expiratory volume (L) in 1 second; ICC, intra-class correlation coefficient; IQR, interquartile range; NR, not reported; NSD, no significant difference; SD, standard deviation; TG, training group; m, meters; min, minutes.

Table 2 Repeatability and reproducibility of incremental shuttle walk test, endurance shuttle walk test, incremental treadmill test, and endurance treadmill test

Study	Patients, n (males, n)	Disease severity (mean ± SD FEV_1 [mean ± SD% predicted] or COPD grading, unless otherwise stated)	Comparison	Additional information
Incremental shuttle walk test				
Repeatability[a]				
Eiser et al[18]	57 (30)	NR (35.00±12.00)	First versus second versus third test	Significant increase in distance from first to second test (Δ13 m; 4.7%]); NSD in distance from second to third test (Δ6 m [2.1%])
McKeough et al[23]	53 (34)	NR (55.00±19.00)	First versus second test	Significant increase in distance from first to second test (Δ20 m [6.3%])[c]
	31 (NR)	NR (51.00±16.00)	Third versus fourth test	Significant increase in distance from third to fourth test (Δ18 m [4.8%])[d]
Vagaggini et al[21]	18 (15)	NR (48.00±14.00)	Familiarization then first versus second test	Significant increase in distance from first to second test, after familiarization (Δ40 m [14.6%])[e]
Reproducibility[b]				
Arnardóttir et al[44]	93 (26)	Moderate/severe COPD	First versus second test	NSD in distance from first to second day (Δ9 m [2.9%])
Campo et al[42]	30 (18)	Mild/moderate/severe/very severe COPD	Familiarization then first versus second test	NSD in distance from first to second day (ICC =0.88) (distances NR)
Eiser et al[18]	57 (30)	NR (35.00±12.00)	First versus second versus third test	Significant increase in distance from first to second day (Δ13 m; 4.6%]); NSD in distance from second to third day (Δ–1 m [0.3%])
Perrault et al[43]	43 (36)	1.40±0.50 (49.00±16.00)	Familiarization then first versus second test	NSD in VO_2, HR, V_E, V_T from first to second test at each of the four cadences (ICC >0.93); distance NR
Singh et al[45]	35 (25)	Group A: 0.50 (0.36–1.45) Group B: 1.10 (0.60–2.10)	First versus second versus third test	Downgraded protocol: NSD in distance on first, second, and third day Modified protocol: significant increase in distance from first to second day (Δ31 m [9.0%]); NSD in distance from second to third day
Endurance shuttle walk test				
Repeatability[a]				
Revill et al[24]	44 (33)	0.94±0.40 (37.00±13.00)	First versus second test	NSD in duration from first to second test (Δ12 s [6.2%])
McKeough et al[23]	53 (34)	NR (55.00±19.00)	First versus second test	NSD in duration from first to second test (Δ–2 s [5.8%])
	31 (NR)	NR (51.00±16.00)	Third versus fourth test	NSD in duration from third to fourth test (Δ44 s [8.7%])
Reproducibility[b]				
Revill et al[47]	44 (22)	Group A: 1.01±0.36 (35.00±4.00) Group B: 0.79±0.21 (34.00±4.00) Group C: 0.80±0.18 (35.00±8.00)	First versus second versus third test	Significant increase in duration from first to second day (Δ59 s; 23.5%]); NSD in duration from second to third day (Δ15 s [4.8%])
Revill et al[46]	23 (13)	0.81±0.27 (33.00±12.00)	Familiarization then first versus second test	NSD in duration from first to second test after familiarization (tests performed with supplemental oxygen)
Endurance treadmill test				
Reproducibility[b]				
Cooper et al[52]	470 (NR)	Moderate/severe/very severe COPD	First versus second test	NSD in duration from first to second test (Δ24 s [7.6%]) (ICC =0.85)
Incremental treadmill test				
Reproducibility				
Mathur et al[50]	8 (6)	0.69±0.16 (NR)	First versus second test	Numerical increase in peak VO_2 from first to second test (Δ1.0 mL/min/kg [8.3%]; statistical test NR)

Notes: [a]Repeatability = similarity of test results when performed on the same day; [b]reproducibility = similarity of test results when performed on different days; [c]$P<0.001$; [d]$P\leq0.05$; [e]$P\leq0.01$.
Abbreviations: FEV_1, forced expiratory volume (L) in 1 second; HR, heart rate; ICC, intra-class correlation coefficient; NR, not reported; NSD, no significant difference; SD, standard deviation; V_E, expired volume; VO_2, oxygen consumption; V_T, tidal volume; m, meters; min, minutes; s, time in seconds.

Table 3 Repeatability and reproducibility of cycling tests

Study	Patients, n (males, n)	Disease severity (mean ± SD FEV_1 [mean ± SD% predicted] or COPD grading, unless otherwise stated)	Comparison	Additional information
Incremental cycle ergometer test				
Repeatability[a]				
Brown et al[25]	11 (NR)	1.51±0.59 (NR)	First versus second test	NSD in peak VO_2 from first to second test (Δ53 mL/min [4.0%]); NSD in W_{max} from first to second test (Δ0 W [0%])
Reproducibility[b]				
Brown et al[25]	11 (NR)	1.51±0.59 (NR)	First versus third test	NSD in peak VO_2 from first to third test (Δ93 mL/min [7.1%]); NSD in W_{max} from first to third test (Δ0 W [0%])
Covey et al[48]	56 (40)	NR (49.00±16.00)	First versus second test	NSD in peak VO_2 from first to second day (Δ0.011 L/min [0.9%]); NSD in W_{max} from first to second day (Δ1 W [1.5%])
Cox et al[49]	11 (8)	Individually listed in paper	First versus second test	NSD in peak VO_2 from first to second day (Δ0.04 L/min [2.0%]); NSD in W_{max} from first to second day (Δ1 W [0.6%])
Mathur et al[50]	8 (6)	0.69±0.16 (NR)	First versus second test	Numerical increase in peak VO_2 from first to second test (Δ0.8 mL/min/kg [6.8%]; statistical test NR)
Poulain et al[30]	10 (NR)	Moderate COPD	Familiarization then first versus second test	NSD in %-predicted peak VO_2 from first to second day (Δ2.5%-predicted VO_2 [3.9%]); NSD in W_{max} from first to second day (Δ1 W [1.3%])
Swinburn et al[41]	17 (6)	0.77±0.30 (NR)	First versus second versus third versus fourth test	Progressive and significant increases in duration from first to fourth day (29%; baseline values NR) (VO_2/W_{max} NR)
Endurance cycle ergometer test				
Reproducibility[b]				
van't Hul et al[51]	60 (46)	Moderate/severe COPD	First versus second test	NSD in duration from first to second test (Δ–12 s [approximately –4.4%]) (ICC =0.85)

Notes: [a]Repeatability = similarity of test results when performed on the same day; [b]reproducibility = similarity of test results when performed on different days.
Abbreviations: FEV_1, forced expiratory volume (L) in 1 second; ICC, intra-class correlation coefficient; NR, not reported; NSD, no significant difference; SD, standard deviation; VO_2, oxygen consumption; W_{max}, maximum workload; W, watts; min, minutes; s, time in seconds.

was no significant difference between the second and third tests.[18,20,21] The only study that presented intra-class correlation coefficients (ICCs) for repeated 6MWTs on the same day reported excellent repeatability (ICC =0.94), but also observed that the second test was significantly higher.[16] For the 12MWT, the distance achieved was reported to be repeatable in the only study in which patients were retested on the same day.[22] Equivocal results were reported for the repeatability of the ISWT by three studies;[18,21,23] one found that the distance was repeatable after familiarization,[18] but the other two reported poor repeatability even after familiarization (Table 2).[21,23] For the ESWT, exercise duration was reported to be repeatable in the two studies in which patients were retested on the same day.[23,24] One study reported that peak VO_2 and W_{max} were repeatable for the ICET (Table 3).[25] No repeatability data were found for the ECET, ITT, or ETT.

Reproducibility of the 6MWT was assessed in 12 studies,[17,18,26–35] six of which reported that distances achieved in the 6MWT demonstrated good reproducibility between the first and second tests (Table 1).[17,27,28,32,35] Of two studies presenting reproducibility results from three 6MWTs,[18,28] only one reported reproducibility between the second and third tests.[28] Two further studies reported results of tests after familiarization;[26,30] only one found 6MWT results to be reproducible.[30] The only study presenting ICC data between the first and second 6MWT performance showed high reproducibility (ICC =0.88), but also that there was a significant increase in distance in the second 6MWT.[33] Reproducibility of the 12MWT was assessed in seven studies.[22,36–41] Five of these presented the results of three or more tests, and reported that the 12MWT distance increased significantly from the first to the second test;[36,37,39–41] one of these studies reported that in a subset of patients who readily experienced exercise-induced hypoxia, the 12MWT distance did not significantly change from the first to the second to the third test.[36] Two additional studies, in which there had been prior familiarization with the test, reported no significant change in 12MWT distance between subsequent first and second tests.[22,38] For the ISWT, five studies presented data assessing reproducibility: from the first to the second test with prior familiarization,[42,43] from the first to the second test without prior familiarization,[44,45] and from the second to the third test without prior familiarization (Table 2).[18] Two further studies evaluating the ESWT reported reproducibility (either from the first to the second test with prior familiarization,[46] or from the second to the third test without familiarization).[47] Good reproducibility of the ICET from the first to the second test was reported by four studies both with[30] and without prior familiarization (Table 3).[25,48,49] Two further studies reported an increase in ICET duration from the first to the second test with no familiarization,[41,50] with one of these studies reporting progressive increases in ICET duration from the first through to the fourth test.[41] In the only study reporting data for the ECET, duration was found to have excellent reproducibility (ICC =0.85).[51] Reproducibility was also found to be excellent in the only study reporting such data for the ETT (no significant increases from the first to the second test, ICC =0.85),[52] but less so for the ITT (increased peak VO_2 from the first to the second test, statistical test not reported).[50]

Several studies were identified that compared the repeatability and/or reproducibility of two or more exercise tests. One study observed that repeatability for the 6MWT and the ISWT was comparable, but that the ISWT was more reproducible.[18] However, another study showed that the ISWT was more repeatable than the 6MWT.[21] One study reported that both the 6MWT and the ICET were reproducible.[30] The ESWT was reported to be more repeatable than the ISWT in one study, when measured in two sessions before and two sessions after pulmonary rehabilitation.[23] In another study, both 12MWT distance and ICET performance were found to increase significantly and progressively over four tests; the ICET was found to have no obvious advantages over the 12MWT when assessing exercise performance.[41] The final study to report reproducibility of more than one test reported that peak VO_2 increased from test 1 to test 2 in both the ICET and the ITT; however, the authors did not report the statistical tests used.[50] Three studies[18,28,30] were found that compared the reproducibility of two or more exercise tests. Of these, two reported that the 6MWT was found to have similar reproducibility to the ISWT[18] and the ICET.[30]

Studies comparing responses to interventions among exercise tests

In total, 23 studies were identified that compared responses of two or more exercise tests after one of the following interventions: pulmonary rehabilitation (16 studies),[23,47,53–66] administration of bronchodilators (six studies),[14,18,67–70] and lung-volume reduction surgery (one study)[71] (Table 4). Of the 16 studies that assessed pulmonary rehabilitation in patients with COPD, the most commonly assessed test was the 6MWT, which was reported by eleven studies.[53–55,57–59,61,63–66]

Two studies compared the response to the 6MWT and ITT after pulmonary rehabilitation; both reported significant increases in 6MWT distance and ITT performance (peak VO_2[53] and work-level completed).[57] However, the latter study did not observe a significant response in peak VO_2 during the ITT after pulmonary rehabilitation.[57] One further study assessed the 6MWT and ITT during nutritional supplementation and placebo, and reported that the 6MWT distance was sensitive to pulmonary rehabilitation (>MCID); but these authors did not present peak VO_2 or W_{max} data for the ITT.[54] Another study assessed the 6MWT, ETT, and ITT, and found that both the ETT and the ITT were sensitive to pulmonary rehabilitation, whereas the 6MWT was not (again, the authors did not report peak VO_2 or W_{max} data for the ITT).[58] Several further studies reported equivocal findings when comparing the 6MWT with the ICET[55,63–65] after pulmonary rehabilitation. All three studies comparing the ECET with the 6MWT found the ECET to be more responsive to pulmonary rehabilitation.[63,65,66] One study assessing responses to pulmonary rehabilitation reported similar sensitivities for the 6MWT and the ISWT, with both giving responses that exceeded the MCID.[61] All four studies assessing the sensitivity of the ISWT and the ESWT to pulmonary rehabilitation reported a significant improvement in performance for both tests;[23,47,60,62] however, in all four studies the response of the ESWT was greater and in two the ISWT response did not reach the MCID.[23,47] An additional study suggested that although both tests showed a significant response that was above the MCID, the ESWT was more responsive to pulmonary rehabilitation than the 6MWT.[59] Equivocal sensitivity was observed in response to pulmonary rehabilitation when using the 12MWT and the ICET.[56]

Of the six studies comparing the responses of two or more exercise tests to bronchodilator therapy,[14,18,67–70] one reported the 6MWT to be more responsive to pharmacological intervention than the 12MWT[67] and one reported the 6MWT to be more responsive to pharmacological intervention

Table 4 Between-test comparisons of sensitivity to interventions

Study	N	Result	Additional information
Pulmonary rehabilitation			
6-minute walk test			
Borghi-Silva et al[54]	8 (Car)	Significant increase in distance: Δ87 m (19.8%) (>MCID)	Training with concomitant carnitine versus placebo; ITT conducted but peak VO_2 and W_{max} NR
	8 (Pla)	Significant increase in distance: Δ34 m (7.3%) (>MCID)	
Borghi-Silva et al[53]	20	Significant increase in distance: Δ105 m (27.8%) (>MCID)	Training significantly improved peak VO_2 compared with control group
Carrieri-Kohlman et al[57]	51	Significant increase in distance: Δ52.1 m (11.7%) and Δ50.3 m (12.7%) (>MCID)	Improvements were similar in both tests whether patients were given a monitored or coached pulmonary rehabilitation program
Cooper[55]	7 (GPT)	Significant increase in distance: Δ32 m (6.9%) (>MCID)	Patients in IMT group showed significant improvement in pulmonary function test; increase in GPT group was not significant
	9 (IMT)	IMT, Δ23 m (4.8%)	
Cortopassi et al[58]	71	NSD in distance	ITT and ETT both responded significantly to PR, but 6MWT did not
Eaton et al[59]	20	Significant increase in distance: Δ47 m (17%) (>MCID)	Standardized mean differences after PR: 6MWT, 0.32 versus ESWT, 0.54 The authors note: "The ESWT has potential advantages in that it may be more responsive than the 6MWT" Values taken from text (do not match Eaton et al)[59]
Ngaage et al[61]	14	Significant increase in distance: Δ68.3 m (36.3%) (>MCID)	ISWT values taken from Ngage et al;[61] changes in both 6MWT and ISWT after PR were clinically significant
Ong et al[63]	37	Significant increase in distance: Δ36 m (9.0%) (>MCID)	6MWT did not correlate with ICET or ECET responses to PR; moderate correlation between ICET W_{max} and ECET duration (r=0.406; P=0.013)
Ries et al[64]	1,218	Significant increase in distance: Δ23.0 m (6.6%) (non-MCID)	Patients who had never previously received PR responded more favorably than those who had (Δ6MW distance: 31.0 versus 18.5 m; ICET W_{max}: 4.3 versus 2.4 W)
Van Helvoort et al[65]	18	Significant increase in distance: Δ63 m (14.8%) (>MCID)	6MWT, ICET, and ECET had greater validity in assessing responses to PR than cardiorespiratory measures
Van Ranst et al[66]	389	Significant increase in distance: Δ52 m (14.2%) (>MCID)	Clinically relevant response after PR for both tests; greater sensitivity in the ECET than in the 6MWT (not assessed statistically)
12-minute walk test			
Arnardóttir et al[56]	28 (IT)	Significant increase in distance: Δ75 m (9.0%)	Similar improvements in 12MWT distance and W_{max} whether training was IT or CT; peak VO_2 was higher in CT, but VO_2 values at isotime were lower in IT
	32 (CT)	Significant increase in distance: Δ94 m (10.8%)	
Incremental shuttle walk test			
Greening et al[60]	601	Significant increase in distance: Δ61 m (28.9%) (>MCID)	Clinically relevant response after PR for both tests; greater sensitivity in the ESWT than in the ISWT (not assessed statistically)
McKeough et al[23]	31	Significant increase in distance: Δ46 m (15%) (non-MCID)	Results are calculated from the better of two pre-PR test results versus the better of two post-PR test results Values taken from Table 2
Ngaage et al[61]	14	Significant increase in distance: Δ57.1 m (52.9%) (>MCID)	ISWT values taken from Ngage et al;[61] changes in both 6MWT and ISWT after PR were clinically significant
O'Farrell et al[62]	85	Significant increase in distance: Δ52.9 m (33.3%) (>MCID)	ESWT more responsive than ISWT to determine improvements in exercise capacity following PR

(*Continued*)

Table 4 (*Continued*)

Study	N	Result	Additional information
Revill et al[47]	44	Significant increase in distance: Δ36 m (20.2%) (non-MCID)	Clinically relevant response after PR in ESWT but not ISWT; greater sensitivity in ESWT than in ISWT (effect size 2.90 versus 0.41)
Endurance shuttle walk test			
Eaton et al[59]	20	Significant increase in distance: Δ302 m (92%) (>MCID); significant increase in duration: Δ270 s (88%) (>MCID)	Standardized mean differences after PR: 6MWT, 0.32 versus ESWT, 0.54 The authors note: "The ESWT has potential advantages in that it may be more responsive than the 6MWT" Values taken from text (do not match Eaton et al)[59]
Greening et al[60]	601	Significant increase in duration: Δ408 s (205.4%) (>MCID)	Clinically relevant response after PR for both tests; greater sensitivity in the ESWT than in the ISWT (not assessed statistically)
McKeough et al[23]	31	Significant increase in duration: Δ182 s (58%)	Results are calculated from the better of two pre-PR test results versus the better of two post-PR test results Values taken from McKeough et al[23]
O'Farrell et al[62]	85	Significant increase in distance: Δ271.0 m (106.7%) (>MCID); significant increase in duration: Δ293.1 s (no percentage change or baseline value provided) (>MCID)	ESWT more responsive than ISWT to determine improvements in exercise capacity following PR
Revill et al[47]	44	Significant increase in distance: Δ334 m (140%) (>MCID); significant increase in duration: Δ404 s (150%) (>MCID)	Clinically relevant response after PR in ESWT but not ISWT; greater sensitivity in ESWT than in ISWT (effect size 2.90 versus 0.41)
Incremental cycle ergometry test			
Arnardóttir et al[56]	28 (IT)	Significant increase in peak VO_2: Δ53 mL/min (5.4%); significant increase in W_{max}: Δ11 W (18.0%) (>MCID)	Similar improvements in 12MWT distance and W_{max} whether training was IT or CT; peak VO_2 was higher in CT, but VO_2 values at isotime were lower in IT
	32 (CT)	Significant increase in peak VO_2: Δ146 mL/min (15.0%); significant increase in W_{max}: Δ11 W (17.2%) (>MCID)	
Cooper[55]	7 (GPT)	NSD in peak VO_2 or W_{max}	Patients in IMT group showed significant improvement in pulmonary function test, increase in GPT group was not significant
	9 (IMT)	NSD in peak VO_2 or W_{max}	
Ong et al[63]	37	Significant increase in peak VO_2: Δ172 mL/min (20.0%); significant increase in W_{max}: 6 W (12.0%) (>MCID)	6MWT did not correlate with ICET or ECET responses to PR; moderate correlation between ICET W_{max} and ECET duration (r=0.406; P=0.013)
Ries et al[64]	1,218	Significant increase in W_{max}: Δ3.1 W (8.6%) (non-MCID)	Patients who had never previously received PR responded more favorably than those who had (Δ6MWT distance: 31.0 versus 18.5 m; ICET W_{max}: 4.3 versus 2.4 W)
Van Helvoort et al[65]	18	NSD in W_{max} and peak VO_2	6MWT, ICET, and ECET had greater validity in assessing responses to PR than cardiorespiratory measures
Endurance cycle ergometry test			
Ong et al[63]	37	Significant increase in duration: Δ322 s (73.5%)	6MWT did not correlate with ICET or ECET responses to PR; moderate correlation between ICET W_{max} and ECET duration (r=0.406; P=0.013)
Van Helvoort et al[65]	18	Significant increase in duration: Δ6.5 min (166%)	6MWT, ICET, and ECET had greater validity in assessing responses to PR than cardiorespiratory measures
Van Ranst et al[66]	389	Significant increase in duration: Δ241 s (84.6%)	Clinically relevant response after PR for both tests; greater sensitivity in the ECET than in the 6MWT (not assessed statistically)
Incremental treadmill test			
Borghi-Silva et al[54]	8 (Car)	Peak VO_2 and W_{max} NR	Training with concomitant carnitine versus placebo; ITT conducted but peak VO_2 and W_{max} NR
	8 (Pla)	Peak VO_2 and W_{max} NR	

Borghi-Silva et al[53]	20	Significant increase in peak VO_2: Δ2 mL/min/kg (14.3%)	Training significantly improved peak VO_2 compared with control group
Carrieri-Kohlman et al[57]	51	NSD in peak VO_2 in either group; significant increase in maximum work (exercise stage): Δ1.6 levels (34.0%) and Δ1.5 levels (34.1%)	Improvements were similar in both tests whether patients were given a monitored or coached pulmonary rehabilitation program
Cortopassi et al[58]	71	Significant increase in duration: Δ184 s (27.4%); significant increase in distance: Δ190 m (22.1%); VO_2 and W_{max} NR	ITT and ETT both responded significantly to PR, but 6MWT did not
Endurance treadmill test			
Cortopassi et al[58]	71	Significant increase in duration: Δ534 s (106.0%); VO_2 and W_{max} NR	ITT and ETT both responded significantly to PR, but 6MWT did not
Bronchodilator therapy			
6-minute walk test			
Cazzola et al[67]	22	Significant increase in distance: Δ53.6 m (baseline values NR) (>MCID)	"6-MWT seems to be a more appropriate instrument than 12-MWT for assessing the exercise response to a bronchodilator"
Eiser et al[18]	54 (6MWT)	Significant increase in distance: Δ37 m (8.7%) (>MCID)	Sensitivity index: 6MWT, 0.84; ISWT, 0.76; distance calculated by number of shuttles performed (1 shuttle =10 m) "There is no important difference in either the reproducibility or sensitivity of self-paced or externally paced walking tests"
Oga et al[68]	38	Significant increase in distance: Δ6 m (1.2%) (non-MCID)	Percentage improvement: 6MWT, 1%; ICET, 3%; ECET, 19% "Among the frequently used post-PR exercise tests, the most responsive index, as measured by the percentage change from baseline, is the endurance time. The correlation between the post-PR changes in these exercise indices is poor"
Pepin et al[14]	14	NSD in distance	"The endurance shuttle walk is more responsive to the acute effects of bronchodilation than the 6MWT"
12-minute walk test			
Cazzola et al[67]	22	Significant increase in distance: Δ59.9 m (baseline values NR) (>MCID)	"6-MWT seems to be a more appropriate instrument than 12-MWT for assessing the exercise response to a bronchodilator"
Incremental shuttle walk test			
Eiser et al[18]	50 (ISWT)	Significant increase in distance: Δ30 m (11.1%) (non-MCID)	Sensitivity index: 6MWT, 0.84; ISWT, 0.76; distance calculated by number of shuttles performed (1 shuttle =10 m) "There is no important difference in either the reproducibility or sensitivity of self-paced or externally paced walking tests"
Endurance shuttle walk test			
Pepin et al[14]	14	Significant increase in distance: Δ144 m (40.0%) (>MCID); duration NR	"The endurance shuttle walk is more responsive to the acute effects of bronchodilation than the 6MWT"
Pepin et al[69]	17	Significant increase in duration: 164 s (>MCID)	"Standardized response mean larger for walking than cycling (0.93 and 0.20, respectively)"; % change and baseline NR
Incremental cycle ergometry test			
Oga et al[68]	38	NSD in peak VO_2; significant increase in W_{max}: 3 W (3.4%)	Percentage improvement: 6MWT, 1%; ICET, 3%; ECET, 19% "Among the frequently used post-PR exercise tests, the most responsive index, as measured by the percentage change from baseline, is the endurance time. The correlation between the post-PR changes in these exercise indices is poor"

(*Continued*)

Table 4 (*Continued*)

Study	N	Result	Additional information
Endurance cycle ergometry test			
Oga et al[68]	38	Significant increase in duration: 34 s (18.0%)	Percentage improvement: 6MWT, 1%; ICET, 3%; ECET, 19% "Among the frequently used post-PR exercise tests, the most responsive index, as measured by the percentage change from baseline, is the endurance time. The correlation between the post-PR changes in these exercise indices is poor"
Pepin et al[69]	17	NSD in duration	"Standardized response mean larger for walking than cycling (0.93 and 0.20, respectively)"; % change and baseline NR
Zhang et al[70]	20	Significant increase in duration: Δ157 s (47.1%)	Values taken from Zhang et al[70] (percentage estimates in text are different); greater percentage change in ECET than in TT
Endurance treadmill test			
Zhang et al[70]	20	Significant increase in duration: Δ110 s (30.4%); W_{max} NR	Values taken from Table 4 (percentage estimates in text are different); greater percentage change in ECET than in TT
Lung-volume reduction surgery			
6-minute walk test			
Lederer et al[71]	23	Significant increase in distance: Δ38 m (9.1%) (>MCID)	Measurements taken 1 year after surgery; clinically relevant change in 6MWT; greater percentage change in ICET than in 6MWT
Incremental cycle ergometry test			
Lederer et al[71]	23	NSD in peak VO_2; significant increase in W_{max}: 6 W (15.4%)	Measurements taken 1 year after surgery; clinically relevant change in 6MWT; greater percentage change in ICET than in 6MWT

Abbreviations: 6MWT, 6-minute walk test; 12MWT, 12-minute walk test; Car, patients receiving carnitine supplements; CT, continuous training; ECET, endurance cycle ergometer test; ESWT, endurance shuttle walk test; ETT, endurance treadmill test; GPT, general physical training; ICET, incremental cycle ergometer test; IMT, inspiratory muscle training; ISWT, incremental shuttle walk test; IT, interval training; ITT, incremental treadmill test; MCID, minimal clinically important difference; NR, not reported; NSD, no significant difference; pla, patients receiving placebo; VO_2, oxygen consumption; PR, pulmonary rehabilitation; TT, treadmill test; VAS, visual analog scale; V_E, ventilation; W_{max}, maximum workload; W, watt.

than the ISWT.[18] The standardized (percentage) increase in response to the ECET was higher than that of either the 6MWT or the ICET in one study,[68] and higher than that of the ETT in another.[70] One study reported the response of the ESWT to bronchodilation to be greater than that of the 6MWT[14] and one study reported the response of the ESWT to bronchodilation to be greater than that of the ECET.[69]

Finally, one study assessed exercise test performance 1 year after lung-volume reduction surgery, and reported a 9.1% increase in 6MWT distance and a 15.4% increase in ICET W_{max}, but noted that ICET peak VO_2 did not increase significantly.[71]

Minimal clinically important differences in responses to interventions

MCIDs have been thus far ascertained for the 6MWT (26 meters),[12] the ISWT (48 meters),[13] the ESWT (45 seconds–85 seconds or 60 meters–115 meters [MCID calculated after bronchodilatory intervention]),[72] and the ICET (4 watts).[12] Of the eleven studies assessing the 6MWT in response to pulmonary rehabilitation, nine reported an increase in excess of the recognized MCID,[53–55,57,59,61,63,65,66] with another reporting a significant increase in distance of less than the MCID[64] (the remaining study reported no significant change in 6MWT distance after pulmonary rehabilitation[58]); these increases ranged in magnitude from 4.8% to 36.3%. Five studies reported a significant response of the ISWT to pulmonary rehabilitation ranging from 15.0% to 52.9%,[23,47,60–62] with three finding that the distance observed reached the MCID.[60–62] All five studies assessing the sensitivity of the ESWT to pulmonary rehabilitation reported that distance and/or duration increased in excess of the MCID (increases ranged from 88.0% to 205.4% when expressed as time [seconds];[23,47,59,60,62] and from 92.0% to 140.0% when expressed as distance [meters]).[47,59,62] Of five studies assessing the ICET before and after pulmonary rehabilitation,[55,56,63–65] two reported W_{max} responses in excess of the MCID,[56,63] and another reported significant changes in W_{max} that did not reach the MCID[64] (the increased ICET performance observed across these studies ranged from 5.4% to 20.0% for peak VO_2 and 8.6% to 18.0% for W_{max}). The two remaining studies did not observe a significant change in ICET peak VO_2 or W_{max} after pulmonary rehabilitation.[55,65] All three studies reporting the response of the ECET to pulmonary rehabilitation observed significant increases in duration (ranging from 73.5% to 166.0%).[63,65,66] Four studies assessed the ITT before and after pulmonary rehabilitation; two did not present data for either peak VO_2 or W_{max}.[54,58] Of the remaining two studies, one reported a significant increase in peak VO_2 of 14.3%,[53] while the other found no significant increase in peak VO_2 after pulmonary rehabilitation.[58] The only study assessing the ETT before and after pulmonary rehabilitation reported significant increases in duration (27.4%) and distance (22.1%).[58]

Of the four studies assessing the 6MWT before and after bronchodilator therapy, two reported improvements in distance that exceeded the MCID (one found an increase in 6MWT distance of 8.7%;[18] the other reported a higher absolute distance increase [53.6 meters], but did not present baseline values; therefore the percentage increase cannot be calculated[67]). The third study reported a small (1.2%) but significant increase in 6MWT distance,[68] while the remaining study found no significant difference in 6MWT distance.[14] One study found a significant increase in 12MWT distance after bronchodilation in excess of the MCID (59.9 meters), but did not provide baseline values.[67] One study reported a significant increase in ISWT performance after bronchodilators (30 meters) that did not reach the MCID.[18] Two studies reported significant increases in ESWT performance above the MCID, of 144 meters[14] and 164 seconds,[69] respectively. One study reported a small but significant improvement in ICET W_{max} of 3.4% after bronchodilator therapy, but noted that peak VO_2 did not increase significantly.[68] Three studies assessed ECET performance after bronchodilators; two found significant increases in duration of 18.0%[68] and 47.1%,[70] with the remaining study reporting no significant improvement in ECET duration.[69]

In the only study assessing lung-volume reduction surgery, a 6MWT improvement in excess of the MCID was seen; however, ICET improvements (whether peak VO_2 or W_{max}) did not reach MCID.[71]

Studies assessing within-test variations in protocol

Eighteen studies were identified that assessed minor variations in protocol within a specific exercise test (Table 5).[17,20,33,46,73–86] Variations (such as track environment or layout and the type of encouragement provided by the investigators to the patient) affected test outcomes (and consequently, their repeatability and reproducibility).

Discussion

A number of laboratory- and field-based exercise tests are used to assess the degree of functional impairment in patients with COPD. However, the choice of which test to use in clinical trials historically seems to have been made

Table 5 Protocol variations reported to affect performance of exercise capacity test

Test	Variation	Study
6MWT	Track layout	Sciurba et al[33]
		Casas et al[76]
		Bansal et al[17]
	Supplemental oxygen provided	Revill et al[46]
		Borghi-Silva et al[75]
		Davidson et al[77]
		Ozalevli et al[83]
	Corridor versus treadmill	Stevens et al[20]
	Indoor versus outdoor environment	de Almeida et al[78]
	Wheeled walker support provided	Honeyman et al[80]
	Verbal encouragement provided	Crisafulli et al[73]
12MWT	Corridor versus treadmill	Swerts et al[86]
ISWT	Verbal encouragement	Rosa et al[84]
	Supplemental oxygen provided	Sandland et al[85]
ESWT	Supplemental oxygen provided	Sandland et al[85]
ICET	Incremental workload gradient used	Benzo et al[74]
		Miyahara et al[82]
ECET	Supplemental oxygen provided	Dean et al[79]
ITT	Incremental workload gradient used	Hsia et al[81]

Abbreviations: 6MWT, 6-minute walk test; 12MWT, 12-minute walk test; ECET, endurance cycle ergometer test; ESWT, endurance shuttle walk test; ICET, incremental cycle ergometer test; ISWT, incremental shuttle walk test, ITT, incremental treadmill test.

on a practical basis (tests such as the 6MWT and 12MWT require little time, organization, or equipment), or without necessarily taking into account how representative the exercise modality used is to activities of daily living for these patients (eg, cycling tests) or the likely impact of the intervention on the outcome of the test. Equally there are few data describing the relative merits of these tests employed simultaneously to evaluate interventions such as rehabilitation and bronchodilator therapy. Results from this systematic review indicate that there is an extensive body of published literature regarding the performance of the eight exercise tests that are widely used.

Repeatability data were found for only the 6MWT, 12MWT, ISWT, ESWT, and ICET. As could be anticipated, we did not identify studies assessing the repeatability of the ECET, ITT, or ETT; these studies test to exhaustion and would be impractical for patients with COPD to perform repeatedly on the same day. Of those reported, the 6MWT was by far the most thoroughly assessed. This may reflect its simplicity and relevance to daily life.[87] However, a substantial proportion of the reported data does not explicitly support 6MWT repeatability. Some studies suggested that the ISWT was more repeatable[21] and reproducible[18] than the 6MWT, while another found that the ESWT was, in turn, more repeatable than the ISWT.[23] The comparative results of exercise tests are inconsistent. Furthermore, there are only very limited data to support the repeatability and/or reproducibility of all cycle tests and treadmill tests. Repeatability and reproducibility were generally improved with familiarization in all types of test. Although this review assesses the influence of protocol variations, it is possible that some studies in which variations were not the primary focus may not have been identified. However, it is clear that even in ostensibly identical tests (eg, two studies reporting the 6MWT), responses can be significantly affected by subtle variations in track layout or environment, or by encouragement from the researchers conducting the test.

When studies reporting the sensitivity of two or more exercise tests to therapeutic intervention were reviewed, there was no consistent evidence supporting the use of one test over any other. Nine of the eleven studies assessing the 6MWT after pulmonary rehabilitation,[53–55,57,59,61,63,65,66] two of the four studies assessing 6MWT after bronchodilators,[18,67] and the only study assessing the 6MWT after lung-volume reduction surgery[71] all reported distance improvements greater than the MCID. Additionally, three of the five studies assessing ISWT[60–62] and all five of the studies assessing ESWT after pulmonary rehabilitation[23,47,59,60,62] reported distance improvements greater than the MCID. Performance improvements were also observed to be in excess of the MCID in the only two studies assessing the ESWT following bronchodilator therapy;[14,69] this was therefore the only exercise test reported by multiple papers that consistently responded to therapeutic interventions to a clinically relevant degree. It must be considered that these exercise tests have differing physiological demands, and it may be that the benefits of each intervention are measured differently by each test. However, limited data for bronchodilator therapy and lung-volume reduction surgery make it difficult to identify any obvious differences in the responses of tests to these interventions.

The review has several limitations that must be acknowledged. As well as identifying whether or not test responses have exceeded the MCID, we have also reported percentage changes in exercise test performance whenever possible to enable a very crude comparison of test outcomes recorded in different units. However, we are aware that the validity of this comparison relies on a direct relationship between these test outcomes, which is unlikely to be true: a large percentage change in one test result may not be equivalent to a large percentage change in another.

In an attempt to assess the validity of exercise tests in patients with COPD as comprehensively as possible, we have collated data from studies that have used various definitions of COPD, which include distinct subcategories such as emphysema and chronic bronchitis, and that have employed diverse methods of assessing diagnosis and severity. Moreover, the study designs included are very diverse. These issues make meaningful meta-analysis difficult. We have, however, tried to provide the ranges of responses (both absolute and percentage changes) when possible, to provide an indication of the magnitude of exercise test responses. A further consequence of the comprehensive nature of this review is that the sample sizes of the identified studies vary widely. For this reason, we have included study sample sizes in our tables.

Decisions regarding which test to use are also influenced by the practicalities of routine clinical practice. It is reasonable to assume that walking is more representative of daily life than cycling for patients with COPD. Furthermore, given the equivocal evidence for the use of the ESWT over the ISWT, clinicians may wish to consider that the ESWT requires a prior "workload setting" ISWT to be performed by the patient, requiring additional time and resource considerations. Any test of exercise capacity should be highly repeatable and reproducible and also should be able to detect changes in performance after interventions aiming to improve exercise capacity.

Conclusion

This review of the published literature has found good evidence to support the repeatability and reproducibility of all tests, particularly the 6MWT, as long as a prior familiarization is conducted. There is consistent evidence to suggest that the ESWT is highly sensitive to therapeutic intervention. Sensitivity data that are available for other tests are largely inconsistent, and the 6MWT and ISWT appear to be less sensitive to intervention than the ESWT and ICET. These factors, allied to practical aspects, must be considered when planning interventional trials.

Author contributions

All authors contributed to the conception and design of the study, the analysis and interpretation of data, and revision of the manuscript. All authors approved the final version of the manuscript for publication.

Acknowledgments

The study was funded by GlaxoSmithKline UK. Writing support was provided by Martin Bell of Oxford PharmaGenesis™ Ltd, funded by GlaxoSmithKline UK. Georgina Meakin and Iain Fotheringham are employees of Oxford PharmaGenesis Ltd, which has received funding from GlaxoSmithKline UK. Yogesh Suresh Punekar, John Riley, and Sarah Cockle are current employees of GlaxoSmithKline, Uxbridge, UK. Sally J Singh was involved with the development of the incremental shuttle walk test, and has served on advisory boards for GlaxoSmithKline. Sally J Singh was part funded by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care East Midlands. 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 National Health Service, the NIHR, or the Department of Health.

Disclosure

Other than the funding outlined in the "Acknowledgments" section, the authors declare no conflicts of interest in this work.

References

1. Lopez AD, Shibuya K, Rao C, et al. Chronic obstructive pulmonary disease: current burden and future projections. *Eur Respir J*. 2006;27(2):397–412.
2. Halbert RJ, Natoli JL, Gano A, Badamgarav E, Buist AS, Mannino DM. Global burden of COPD: systematic review and meta-analysis. *Eur Respir J*. 2006;28(3):523–532.
3. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med*. 2006;3(11):e442.
4. American Thoracic Society/European Respiratory Society Task Force. *Standards for the Diagnosis and Management of Patients with COPD* [Internet]. Version 1.2. New York: American Thoracic Society; 2004 [updated September 8, 2005]. Available from: http://www.thoracic.org/go/copd-guidelines/. Accessed May 19, 2014.
5. Global Initiative for Chronic Obstructive Lung Disease (GOLD). *Global Strategy for the Diagnosis, Management and Prevention of COPD* [Internet]. GOLD; 2014. Available from: http://www.goldcopd.org/uploads/users/files/GOLD_Report_2014_Jun11.pdf. Accessed September 12, 2014.
6. Corhay JL, Dang DN, Van Cauwenberge H, Louis R. Pulmonary rehabilitation and COPD: providing patients a good environment for optimizing therapy. *Int J Chron Obstruct Pulmon Dis*. 2014;9:27–39.
7. Jones P, Miravitlles M, van der Molen T, Kulich K. Beyond FEV(1) in COPD: a review of patient-reported outcomes and their measurement. *Int J Chron Obstruct Pulmon Dis*. 2012;7:697–709.
8. Cooper CB. The connection between chronic obstructive pulmonary disease symptoms and hyperinflation and its impact on exercise and function. *Am J Med*. 2006;119(10 Suppl 1):21–31.
9. Jones PW. Issues concerning health-related quality of life in COPD. *Chest*. 1995;107(Suppl 5):187S–193S.
10. Mahler DA, Harver A. A factor analysis of dyspnea ratings, respiratory muscle strength, and lung function in patients with chronic obstructive pulmonary disease. *Am Rev Respir Dis*. 1992;145(2 Pt 1):467–470.
11. National Institute for Health and Clinical Excellence (NICE). Management of chronic obstructive pulmonary disease in adults in primary and secondary care (partial update). NICE guideline CG101. London: NICE; 2010. Available from: http://guidance.nice.org.uk/cg101. Accessed May 19, 2014.
12. Puhan MA, Chandra D, Mosenifar Z, et al. The minimal important difference of exercise tests in severe COPD. *Eur Respir J*. 2011;37(4):784–790.

13. Singh SJ, Jones PW, Evans R, Morgan MD. Minimum clinically important improvement for the incremental shuttle walking test. *Thorax*. 2008;63(9):775–777.
14. Pepin V, Brodeur J, Lacasse Y, et al. Six-minute walking versus shuttle walking: responsiveness to bronchodilation in chronic obstructive pulmonary disease. *Thorax*. 2007;62(4):291–298.
15. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ*. 2009;339:b2535.
16. Andersson M, Moberg L, Svantesson U, Sundbom A, Johansson H, Emtner M. Measuring walking speed in COPD: Test-retest reliability of the 30-metre walk test and comparison with the 6-minute walk test. *Prim Care Respir J*. 2011;20(4):434–440.
17. Bansal V, Hill K, Dolmage TE, Brooks D, Woon LJ, Goldstein RS. Modifying track layout from straight to circular has a modest effect on the 6-min walk distance. *Chest*. 2008;133(5):1155–1160.
18. Eiser N, Willsher D, Dore CJ. Reliability, repeatability and sensitivity to change of externally and self-paced walking tests in COPD patients. *Respir Med*. 2003;97(4):407–414.
19. Jenkins S, Cecins NM. Six-minute walk test in pulmonary rehabilitation: do all patients need a practice test? *Respirology*. 2010;15(8): 1192–1196.
20. Stevens D, Elpern E, Sharma K, Szidon P, Ankin M, Kesten S. Comparison of hallway and treadmill six-minute walk tests. *Am J Respir Crit Care Med*. 1999;160(5 Pt 1):1540–1543.
21. Vagaggini B, Taccola M, Severino S, et al. Shuttle walking test and 6-minute walking test induce a similar cardiorespiratory performance in patients recovering from an acute exacerbation of chronic obstructive pulmonary disease. *Respiration*. 2003;70(6):579–584.
22. O'Reilly JF, Shaylor JM, Fromings KM, Harrison BD. The use of the 12 minute walking test in assessing the effect of oral steroid therapy in patients with chronic airways obstruction. *Br J Dis Chest*. 1982;76(4):374–382.
23. 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(1):35–39.
24. Revill SM, Williams J, Sewell L, Collier R, Singh SJ. Within-day repeatability of the endurance shuttle walk test. *Physiotherapy*. 2009;95(2):140–143.
25. Brown SE, Fischer CE, Stansbury DW, Light RW. Reproducibility of VO2max in patients with chronic air-flow obstruction. *Am Rev Respir Dis*. 1985;131(3):435–438.
26. Behnke M, Wewel AR, Kirsten D, Jorres RA, Magnussen H. Exercise training raises daily activity stronger than predicted from exercise capacity in patients with COPD. *Respir Med*. 2005;99(6):711–717.
27. Brooks D, Solway S, Weinacht K, Wang D, Thomas S. Comparison between an indoor and an outdoor 6-minute walk test among individuals with chronic obstructive pulmonary disease. *Arch Phys Med Rehabil*. 2003;84(6):873–876.
28. Chatterjee AB, Rissmiller RW, Meade K, et al. Reproducibility of the 6-minute walk test for ambulatory oxygen prescription. *Respiration*. 2010;79(2):121–127.
29. Kozu R, Jenkins S, Senjyu H, Mukae H, Sakamoto N, Kohno S. Peak power estimated from 6-minute walk distance in Asian patients with idiopathic pulmonary fibrosis and chronic obstructive pulmonary disease. *Respirology*. 2010;15(4):706–713.
30. Poulain M, Durand F, Palomba B, et al. 6-minute walk testing is more sensitive than maximal incremental cycle testing for detecting oxygen desaturation in patients with COPD. *Chest*. 2003;123(5): 1401–1407.
31. Rejeski WJ, Foley KO, Woodard CM, Zaccaro DJ, Berry MJ. Evaluating and understanding performance testing in COPD patients. *J Cardiopulm Rehabil*. 2000;20(2):79–88.
32. Roomi J, Johnson MM, Waters K, Yohannes A, Helm A, Connolly MJ. Respiratory rehabilitation, exercise capacity and quality of life in chronic airways disease in old age. *Age Ageing*. 1996;25(1):12–16.
33. Sciurba F, Criner GJ, Lee SM, et al. Six-minute walk distance in chronic obstructive pulmonary disease: reproducibility and effect of walking course layout and length. *Am J Respir Crit Care Med*. 2003;167(11):1522–1527.
34. Spencer LM, Alison JA, McKeough ZJ. Six-minute walk test as an outcome measure: are two six-minute walk tests necessary immediately after pulmonary rehabilitation and at three-month follow-up? *Am J Phys Med Rehabil*. 2008;87(3):224–228.
35. Troosters T, Vilaro J, Rabinovich R, et al. Physiological responses to the 6-min walk test in patients with chronic obstructive pulmonary disease. *Eur Respir J*. 2002;20(3):564–569.
36. Arnardóttir RH, Sorensen S, Ringqvist I, Larsson K. No increase in walking distance on repeated tests in COPD patients with exercise-induced hypoxaemia. *Adv Physiother*. 2007;9(4):161–168.
37. Beaumont A, Cockcroft A, Guz A. A self paced treadmill walking test for breathless patients. *Thorax*. 1985;40(6):459–464.
38. Berger R, Smith D. Effect on inhaled metaproterenol on exercise performance in patients with stable 'fixed' airway obstruction. *Am Rev Respir Dis*. 1988;138(3):624–629.
39. McGavin CR, Gupta SP, McHardy GJ. Twelve minute walking test for assessing disability in chronic bronchitis. *BMJ*. 1976;1(6013): 822–823.
40. Mungall IP, Hainsworth R. Assessment of respiratory function in patients with chronic obstructive airways disease. *Thorax*. 1979;34(2): 254–258.
41. Swinburn CR, Wakefield JM, Jones PW. Performance, ventilation, and oxygen consumption in three different types of exercise test in patients with chronic obstructive lung disease. *Thorax*. 1985;40(8):581–586.
42. Campo LA, Chilingaryan G, Berg K, Paradis B, Mazer B. Validity and reliability of the modified shuttle walk test in patients with chronic obstructive pulmonary disease. *Arch Phys Med Rehabil*. 2006;87(7): 918–922.
43. Perrault H, Baril J, Henophy S, Rycroft A, Bourbeau J, Maltais F. Paced-walk and step tests to assess exertional dyspnea in COPD. *COPD*. 2009;6(5):330–339.
44. Arnardóttir RH, Emtner M, Hedenstrom H, Larsson K, Boman G. Peak exercise capacity estimated from incremental shuttle walking test in patients with COPD: a methodological study. *Respir Res*. 2006;7(127).
45. Singh SJ, Morgan MD, Scott S, Walters D, Hardman AE. Development of a shuttle walking test of disability in patients with chronic airways obstruction. *Thorax*. 1992;47(12):1019–1024.
46. Revill SM, Noor MZ, Butcher G, Ward MJ. The endurance shuttle walk test: An alternative to the six-minute walk test for the assessment of ambulatory oxygen. *Chron Respir Dis*. 2010;7(4):239–245.
47. Revill SM, Morgan MD, Singh SJ, Williams J, Hardman AE. The endurance shuttle walk: a new field test for the assessment of endurance capacity in chronic obstructive pulmonary disease. *Thorax*. 1999;54(3):213–222.
48. Covey MK, Larson JL, Alex CG, Wirtz S, Langbein WE. Test-retest reliability of symptom-limited cycle ergometer tests in patients with chronic obstructive pulmonary disease. *Nurs Res*. 1999;48(1):9–19.
49. Cox NJ, Hendriks JC, Binkhorst RA, Folgering HT, van Herwaarden CL. Reproducibility of incremental maximal cycle ergometer tests in patients with mild to moderate obstructive lung diseases. *Lung*. 1989;167(2):129–133.
50. Mathur RS, Revill SM, Vara DD, Walton R, Morgan MD. Comparison of peak oxygen consumption during cycle and treadmill exercise in severe chronic obstructive pulmonary disease. *Thorax*. 1995;50(8): 829–833.
51. van't Hul A, Gosselink R, Kwakkel G. Constant-load cycle endurance performance test-retest reliability and validity in patients with COPD. *J Cardiopulm Rehabil*. 2003;23(2):143–150.
52. Cooper CB, Abrazado M, Legg D, Kesten S. Development and implementation of treadmill exercise testing protocols in COPD. *Int J Chron Obstruct Pulmon Dis*. 2010;5:375–385.

53. Borghi-Silva A, Arena R, Castello V, et al. Aerobic exercise training improves autonomic nervous control in patients with COPD. *Respir Med*. 2009;103(10):1503–1510.
54. Borghi-Silva A, Baldissera V, Sampaio LM, et al. L-carnitine as an ergogenic aid for patients with chronic obstructive pulmonary disease submitted to whole-body and respiratory muscle training programs. *Braz J Med Biol Res*. 2006;39(4):465–474.
55. Cooper CB. Desensitization to dyspnea in COPD with specificity for exercise training mode. *Int J Chron Obstruct Pulmon Dis*. 2009;4:33–43.
56. Arnardóttir RH, Boman G, Larsson K, Hedenström H, Emtner M. Interval training compared with continuous training in patients with COPD. *Respir Med*. 2007;101(6):1196–1204.
57. Carrieri-Kohlman V, Gormley JM, Douglas MK, Paul SM, Stulbarg MS. Exercise training decreases dyspnea and the distress and anxiety associated with it. Monitoring alone may be as effective as coaching. *Chest*. 1996;110(6):1526–1535.
58. Cortopassi F, Castro AA, Porto EF, et al. Comprehensive exercise training improves ventilatory muscle function and reduces dyspnea perception in patients with COPD. *Monaldi Arch Chest Dis*. 2009;71(3):106–112.
59. Eaton T, Young P, Nicol K, Kolbe J. The endurance shuttle walking test: A responsive measure in pulmonary rehabilitation for COPD patients. *Chron Respir Dis*. 2006;3(1):3–9.
60. Greening NJ, Evans RA, Williams JE, Green RH, Singh SJ, Steiner MC. Does body mass index influence the outcomes of a waking-based pulmonary rehabilitation programme in COPD? *Chron Respir Dis*. 2012;9(2):99–106.
61. Ngaage DL, Hasney K, Cowen ME. The functional impact of an individualized, graded, outpatient pulmonary rehabilitation in end-stage chronic obstructive pulmonary disease. *Heart Lung*. 2004;33(6):381–389.
62. O'Farrell R, Gargoum F, O'Connor B, et al. Comparison of incremental and endurance shuttle walk testing after pulmonary rehabilitation. *Am J Resp Crit Care Med*. 2011;183:Abs A2024.
63. Ong KC, Chong WF, Soh C, Earnest A. Comparison of different exercise tests in assessing outcomes of pulmonary rehabilitation. *Respir Care*. 2004;49(12):1498–1503.
64. Ries AL, Make BJ, Lee SM, et al. The effects of pulmonary rehabilitation in the National Emphysema Treatment Trial. *Chest*. 2005;128(6): 3799–3809.
65. Van Helvoort HA, De Boer RC, Van De Broek L, Dekhuijzen R, Heijdra YF. Exercises commonly used in rehabilitation of patients with chronic obstructive pulmonary disease: cardiopulmonary responses and effect over time. *Arch Phys Med Rehabil*. 2011;92(1):111–117.
66. Van Ranst D, Otten H, Meijer JW, van't Hul AJ. Outcome of pulmonary rehabilitation in COPD patients with severely impaired health status. *Int J Chron Obstruct Pulmon Dis*. 2011;6(1):647–657.
67. Cazzola M, Biscione GL, Pasqua F, et al. Use of 6-min and 12-min walking test for assessing the efficacy of formoterol in COPD. *Respir Med*. 2008;102(10):1425–1430.
68. Oga T, Nishimura K, Tsukino M, Hajiro T, Ikeda A, Izumi T. The effects of oxitropium bromide on exercise performance in patients with stable chronic obstructive pulmonary disease: a comparison of three different exercise tests. *Am J Respir Crit Care Med*. 2000;161(6):1897–1901.
69. Pepin V, Saey D, Whittom F, LeBlanc P, Maltais F. Walking versus cycling: sensitivity to bronchodilation in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2005;172(12):1517–1522.
70. Zhang X, Waterman LA, Ward J, Baird JC, Mahler DA. Advantages of endurance treadmill walking compared with cycling to assess bronchodilator therapy. *Chest*. 2010;137(6):1354–1361.
71. Lederer DJ, Thomashow BM, Ginsburg ME, et al. Lung-volume reduction surgery for pulmonary emphysema: Improvement in body mass index, airflow obstruction, dyspnea, and exercise capacity index after 1 year. *J Thorac Cardiovasc Surg*. 2007;133(6):1434–1438.
72. Pepin V, Laviolette L, Brouillard C, et al. Significance of changes in endurance shuttle walking performance. *Thorax*. 2011;66(2):115–120.
73. Crisafulli E, Lorenzi MC, Gherardini G, et al. Test del cammino dei 6 minuti in pazienti con bronchopneupatia cronica ostruttiva: confronto cone senza incoraggiamento [Six-minute walk test in patients with chronic obstructive pulmonary disease: comparison with and without active coaching]. *Rassegna di Patologia dell'Apparato Respiratorio*. 2007;22(3):186–192.
74. Benzo RP, Paramesh S, Patel SA, Slivka WA, Sciurba FC. Optimal protocol selection for cardiopulmonary exercise testing in severe COPD. *Chest*. 2007;132(5):1500–1505.
75. Borghi-Silva A, Mendes RG, Toledo AC, et al. Adjuncts to physical training of patients with severe COPD: oxygen or noninvasive ventilation? *Respir Care*. 2010;55(7):885–894.
76. Casas A, Vilaro J, Rabinovich R, et al. Encouraged 6-min walking test indicates maximum sustainable exercise in COPD patients. *Chest*. 2005;128(1):55–61.
77. Davidson AC, Leach R, George RJ, Geddes DM. Supplemental oxygen and exercise ability in chronic obstructive airways disease. *Thorax*. 1988;43(12):965–971.
78. de Almeida FG, Victor EG, Rizzo JA. Hallway versus treadmill 6-minute-walk tests in patients with chronic obstructive pulmonary disease. *Respir Care*. 2009;54(12):1712–1716.
79. Dean NC, Brown JK, Himelman RB, Doherty JJ, Gold WM, Stulbarg MS. Oxygen may improve dyspnea and endurance in patients with chronic obstructive pulmonary disease and only mild hypoxemia. *Am Rev Respir Dis*. 1992;146(4):941–945.
80. Honeyman P, Barr P, Stubbing DG. Effect of a walking aid on disability, oxygenation, and breathlessness in patients with chronic airflow limitation. *J Cardiopulm Rehabil*. 1996;16(1):63–67.
81. Hsia D, Casaburi R, Pradhan A, Torres E, Porszasz J. Physiological responses to linear treadmill and cycle ergometer exercise in COPD. *Eur Resp J*. 2009;34(3):605–615.
82. Miyahara N, Eda R, Takeyama H, et al. Cardiorespiratory responses during cycle ergometer exercise with different ramp slope increments in patients with chronic obstructive pulmonary disease. *Intern Med*. 2000;39(1):15–19.
83. Ozalevli S, Ozden A, Gocen Z, Cimrin AH. Comparison of six-minute walking tests conducted with and without supplemental oxygen in patients with chronic obstructive pulmonary disease and exercise-induced oxygen desaturation. *Ann Saudi Med*. 2007;27(2):94–100.
84. Rosa FW, Camelier A, Mayer A, Jardim JR. Evaluating physical capacity in patients with chronic obstructive pulmonary disease: Comparing the shuttle walk test with the encouraged 6-minute walk test. *J Bras Pneumol*. 2006;32(2):106–113.
85. Sandland CJ, Morgan MD, Singh SJ. Detecting oxygen desaturation in patients with COPD: Incremental versus endurance shuttle walking. *Respir Med*. 2008;102(8):1148–1152.
86. Swerts PM, Mostert R, Wouters EF. Comparison of corridor and treadmill walking in patients with severe chronic obstructive pulmonary disease. *Phys Ther*. 1990;70(7):439–442.
87. Morgan MD, Singh SJ. Assessing the exercise response to a bronchodilator in COPD: time to get off your bike? *Thorax*. 2007;62(4): 281–283.

Supplementary materials
Complete Ovid search strings

Table S1 Embase™ search strings, search conducted January 22, 2013

Search	Search strings	Results
1	chronic obstructive pulmonary disease.mp. or exp chronic obstructive lung disease/	68,878
2	chronic obstructive lung disease.mp.	65,527
3	chronic obstructive respiratory disease.mp.	119
4	copd.mp.	33,438
5	chronic obstructive airway disease.mp.	378
6	chronic bronchitis.mp. or exp chronic bronchitis/	16,146
7	lung emphysema.mp. or exp lung emphysema/	19,606
8	chronic obstructive airway disease.mp.	378
9	chronic airflow obstruction.mp.	627
10	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9	103,379
11	exp cardiopulmonary exercise test/or exp exercise tolerance/or exp aerobic exercise/ or exp exercise/or exp exercise test/	228,181
12	(exercise or exercise capacity or exercise tolerance).mp.	302,038
13	(clinically important difference or mcid or clinically meaningful).mp.	6,903
14	11 or 12 or 13	311,378
15	('6 minute walk' or '6m walk' or '6 m walk' or 'six minute walk' or '6MWD').mp.	4,644
16	('12 minute walk' or '12m walk' or '12 m walk' or 'twelve minute walk' or '12MWD').mp.	90
17	('shuttle walk' or 'shuttle walking' or 'iswt' or '10 metre walking' or '10 meter walking' or 'ESWT').mp.	1,010
18	cycle ergometer.mp. or exp bicycle ergometer/	5,540
19	bicycle ergometry.mp. or exp bicycle ergometry/	6,840
20	(cycle ergometry or cycle ergometric or bicycle ergometric).mp.	1,721
21	treadmill.mp. or exp treadmill/	29,379
22	(short physical performance battery or sppb).mp.	367
23	15 or 16 or 17 or 18 or 19 or 20 or 21 or 22	45,996
24	10 and 14 and 23	1,626
25	(test or assessment or capacity).mp.	3,225,394
24	10 and 14 and 23	1,408
25	(test or assessment or capacity).mp.	1,099

Table S2 Medline® search strings, search conducted January 22, 2013

Search	Search string	Results
1	chronic obstructive pulmonary disease.mp. or exp pulmonary disease, chronic obstructive/	31,239
2	chronic obstructive lung disease.mp.	2,686
3	exp lung diseases, obstructive/or chronic obstructive respiratory disease.mp.	150,257
4	copd.mp.	21,567
5	chronic obstructive airway disease.mp.	239
6	chronic bronchitis.mp. or exp bronchitis, chronic/	8,867
7	lung emphysema.mp.	298
8	chronic obstructive airway disease.mp.	239
9	chronic airflow obstruction.mp.	502
10	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9	161,227
11	exp exercise tolerance/or exp exercise test/ or exercise.mp. or exp exercise/	239,508
12	(exercise capacity or exercise tolerance).mp.	16,495
13	(clinically important difference or mcid or clinically meaningful).mp.	4,675
14	11 or 12 or 13	243,911
15	('6 minute walk' or '6m walk' or '6 m walk' or 'six minute walk' or '6mwd').mp.	2,563
16	('12 minute walk' or '12m walk' or '12 m walk' or 'twelve minute walk' or '12mwd').mp.	62
17	('shuttle walk' or 'shuttle walking' or 'iswt' or '10 metre walking' or '10 meter walking' or 'eswt').mp.	644
18	$cycle ergomet*.mp.	4,838
19	treadmill.mp.	20,912
20	(short physical performance battery or sppb).mp.	247
21	15 or 16 or 17 or 18 or 19 or 20	28,642
22	10 and 14 and 21	1,278
23	(test or assessment or capacity).mp.	1,939,771
24	10 and 14 and 21 and 23	1,065
25	limit 24 to (English language and humans)	928

Table S3 The Cochrane Library search strings, search conducted January 22, 2013

Search	Search string	Results
1	MeSH descriptor: [Pulmonary Disease, Chronic Obstructive] explode all trees	1,845
2	chronic obstructive pulmonary disease	4,929
3	MeSH descriptor: [Lung Diseases, Obstructive] explode all trees	12,759
4	chronic obstructive respiratory disease	3,340
5	copd	6,244
6	chronic obstructive airway disease	4,348
7	MeSH descriptor: [Bronchitis, Chronic] explode all trees	77
8	chronic bronchitis	1,969
9	lung emphysema	517
10	chronic obstructive airway disease	4,348
11	chronic airflow obstruction	416
12	#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11	18,973
13	MeSH descriptor: [Exercise Tolerance] explode all trees	1,431
14	MeSH descriptor: [Exercise Test] explode all trees	5,816
15	MeSH descriptor: [Exercise] explode all trees	11,528
16	exercise or exercise capacity or exercise tolerance	38,220
17	clinically important difference or mcid or clinically meaningful	7,450
18	#13 or #14 or #15 or #16 or #17	45,480
19	‘6 minute walk’ or ‘6m walk’ or ‘6 m walk’ or ‘six minute walk’ or ‘6mwd’	3,377
20	‘12 minute walk’ or ‘12m walk’ or ‘12 m walk’ or ‘twelve minute walk’ or ‘12mwd’	2,233
21	‘shuttle walk’ or ‘shuttle walking’ or ‘iswt’ or ‘10 metre walking’ or ‘10 meter walking’ or ‘eswt’	716
22	*cycle ergomet*	2,392
23	treadmill	3,759
24	short physical performance batter or sppb	41
25	#19 or #20 or #21 or #22 or #23 or #24	9,217
26	test or assessment or capacity	211,756
27	#12 and #18 and #25 and #26	779