REVIEW ARTICLE

Oman Medical Journal [2023], Vol. 38, No. 1: e455

Reliability and Responsiveness of Endurance Shuttle Walk Test to Estimate Functional Exercise Capacity in Patients with Chronic Obstructive Pulmonary Disease: A Systematic Review and Meta-analysis

Shagufa Amber, Aqsa Mujaddadi and Jamal Ali Moiz*

Centre for Physiotherapy and Rehabilitation Sciences, Jamia Millia Islamia (A Central University), New Delhi, India

ARTICLE INFO

Article history:
Received: 18 June 2021
Accepted: 27 February 2022

Online:
DOI 10.5001/omj.2023.21

Keywords:
Exercise Tolerance; Walk Test; Pulmonary Disease, Chronic Obstructive.

ABSTRACT

The endurance shuttle walk test (ESWT) is a simple, acceptable, field-based test first established in 1999 to measure endurance exercise capacity in patients with chronic obstructive pulmonary disease (COPD). The aim of this systematic review was to examine the reliability and responsiveness of ESWT in COPD. Of the 791 articles identified through electronic databases, 17 were included in this review. Qualitative and quantitative analyses were conducted according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses, and as per Consensus-Based Standards for the Selection of Health Status Measurements Instruments, the quality of the studies was graded as low for reliability and moderate for responsiveness. Qualitative analysis indicated inadequate evidence for the reliability of the ESWT in patients with COPD. The meta-analysis found strong evidence that ESWT was responsive to change following pulmonary rehabilitation with an estimated mean difference (ESWT time, seconds) 303.19 s (95% CI: 175.63–430.75; $p < 0.001$), ambulatory oxygen with a mean difference (ESWT time, seconds) 129.04 s (95% CI: 47.98–210.09; $p = 0.002$), and (ESWT mean distance, meters) 80.71 m (95% CI: 38.66–122.76; $p < 0.001$). The ESWT was also responsive to bronchodilation with a mean difference of 168.62 m (95% CI: 117.03–220.21; $p < 0.001$). Our findings suggest the strong potential of ESWT as a responsive test in COPD, but to draw a definitive conclusion regarding the reliability of the ESWT, further research is needed in this population.

The endurance shuttle walk test (ESWT) is a test of endurance capacity, first described in 1999 in patients with chronic obstructive pulmonary disease (COPD).[1] The measurement of endurance capacity using the ESWT is a simple, acceptable field-based test, with nominal cost or resource implications. It is valid and highly responsive with minimal learning effects.[2] The endurance time is used in the evaluation of exercise tolerance in COPD.[3] The ESWT was designed to complement the incremental shuttle walk test (ISWT) and uses the same 10-meter shuttle course.[4] The pace of the ESWT is traditionally calculated at a predefined percentage of peak performance on the ISWT around 70–85% estimated VO_2 peak.[5] However, recent data have indicated that the speed can also be accurately derived using 85% of walking speed on the ISWT, making it easier.[6] The test is terminated when the subject is limited by dyspnea or a heart rate > 85% predicted maximum, or when the subject is unable to maintain the required speed and hence, fails to complete a shuttle for a second consecutive time.[4] The primary outcome is the distance covered (meters, m) or the time required (seconds, s) to complete the test.

The leading cause of worldwide mortality and morbidity is attributed to COPD, ranked eighth in causing disability and disease burden in 2015 by disability-adjusted life years.[7–9] Breathlessness and reduced exercise capacity are characteristic symptoms of COPD.[10,11] Reduction in exercise capacity can result in reduced ability to perform activities of daily living (ADL) and further, the resultant sedentary lifestyle and inactivity can exacerbate exercise impairment (the COPD "vicious circle").[12] The systemic effects of COPD

*Corresponding author: ✉jmoiz@jmi.ac.in

impairs exercise tolerance, peripheral muscle endurance, and QOL.[13–15]

The gold standard method, the cardiopulmonary exercise test, has been used to assess exercise capacity in COPD by using a cycle ergometer to measure the indexes of pulmonary and cardiac performance, as the VO_2 maximum.[16] However, implementation of cardiopulmonary exercise test requires technical expertise and may not be readily available in every testing set-up.[16] The estimation of exercise capacity in patients with chronic cardiopulmonary diseases by field walking tests has been simpler and better in representing the demands of ADL.[17] The most common method to calculate endurance utilizing the field test, is the six-minute walk test (6MWT), ISWT,[18] and ESWT.[19] The 6MWT has certain limitations of being time-based and self-paced, while the ISWT is externally paced and controlled by a series of pre-recorded signals.[18] The ESWT on the other hand is a constant-load exercise test that measures the ability of the participant to sustain a given submaximal exercise capacity. For practical purposes, ESWT may be considered field-based tests that can provide a true measure of endurance capacity over ISWT as patients are unaware of any time limit and discouraged from estimating how long they are sustaining exercise intensity relative to the individual's submaximal exercise capacity.

The measurement property of any test is vital in the selection and administration of that specific test in rehabilitation settings.[20] The reliability and responsiveness of ESWT have been examined in studies in patients with COPD.[21] The difference between tests repeated on the same day was generally small and non-significant statistically.[6] There have been no reports of adverse events associated with performing the ESWT in clinical practice or in the context of clinical trials.[5] There was an insignificant difference in test-retest reliability and repeatability in the two studies.[21,22] The validity of the ESWT has not yet been established nor compared with laboratory-based exercise tests.[5] The responsiveness of ESWT in COPD has been reported in studies following pulmonary rehabilitation (PR), bronchodilation (BD), and ambulatory oxygen therapy (AO). The responsiveness of ESWT was moderate to high, with a standardized response mean (SRM) ranging from 0.52 to 1.27.[1,2,23–27] However, to date, there is no systematic review and meta-analysis that can qualitatively and quantitatively summarize the findings obtained from these studies. Therefore, the objective of this systematic review was to examine the reliability and responsiveness of ESWT in patients with COPD.

METHODS

Protocol and registration

This systematic review was registered in the International Prospective Register of Systematic Reviews (CRD42020217847) on 28 November 2020. The PRISMA (Parameters of the Preferred Reporting Items for Systematic reviews and Meta-Analyses),[28] and COSMIN (Consensus-Based Standards for the Selection of Health Status Measurements Instruments) guidelines were applied.[29]

Eligibility criteria

Studies describing the reliability and responsiveness of ESWT in COPD patients regardless of age, gender, and disease severity were included. Studies on repeatability and test-retest reliability were selected. The studies evaluating ESWT's responsiveness to PR, BD, and AO were included. The studies not in the English language, unavailable as full-text articles, and in populations other than COPD were excluded. After reviewing the titles and abstracts from relevant studies, full-text articles were retrieved.

Information sources and search strategy

The literature search was performed through the following electronic databases: Web of Science, PubMed, and Cochrane (via CENTRAL). The keywords used included, 'endurance shuttle walk test', 'chronic obstructive pulmonary disease', 'ESWT', 'COPD', 'responsiveness', and 'reliability'. The systematic search of articles was refined in three databases through Boolean operators 'AND' and 'OR'.[30] The search strategy of the Web of Science included 'endurance shuttle walk test' OR 'ESWT' OR 'endurance shuttle walking test' AND 'COPD' OR 'chronic obstructive pulmonary disease' AND 'responsiveness' AND 'reliability'.

Measurement properties

Reliability was defined as the degree by which the measure is free of random error and is consistent.[31]

In this systematic review, studies of test-retest reliability or of measurement of error were considered. Responsiveness is defined as the ability of a measuring instrument to detect a minimal change in the score when an actual change in status occurs over time.[32] Interpretability was defined as the degree of change (i.e., minimal clinically important difference (MCID)).[30]

Study selection

A literature search was conducted, and the studies were imported on Mendeley Desktop, reference manager. The articles from all the databases (Web of Science, PubMed, and Cochrane Library) were retrieved. The reviewer screened titles and abstracts of the selected articles for duplicates and adherence to eligibility criteria. The potentially relevant studies were scanned from reference lists of identified studies. Further, two authors retrieved full-text articles and individually evaluated them. In case of disagreement at any stage, it was resolved through discussion with the third reviewer.

Data extraction process

The two authors (S.A, A.M) extracted and tabulated data from each selected article under categories of study characteristics (sample size, age, gender, disease severity, and instrument administration), measurement properties (reliability, responsiveness, and MCID), and the main findings observed. The review was carried out in accordance with the PRISMA statement.[28] The mean difference (MD) and SE, were extracted for a meta-analysis of ESWT responsiveness to PR, BD, and AO therapy in COPD. Any disagreement was resolved through discussion with the third reviewer (J.M).

Quality assessment

The COSMIN guidelines were used for assessing the risk of bias (RoB) of the selected studies,[29] and the data were extracted for the reliability and responsiveness properties. COSMIN consists of a set of items for the evaluation of measurement property (reliability and responsiveness). The two authors (S.A, A.M) independently assessed and reviewed the selected articles and the disagreement was resolved by consensus. The RoB was done by the rating score system of four points as very good, adequate, doubtful, and inadequate.[33] Studies in

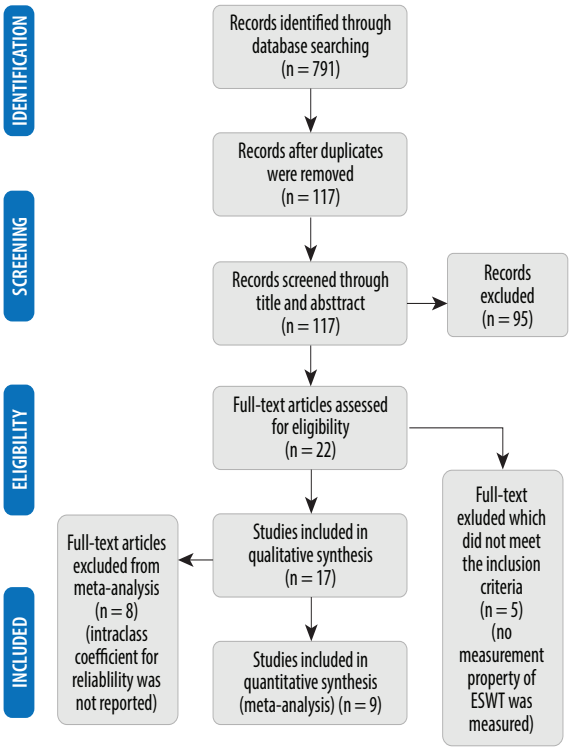

Figure 1: Parameters of the Preferred Reporting Items for Systematic reviews and Meta-Analyses flow chart-study selection process.

which more than one measurement property was analyzed, quality assessment was performed for each measurement property as per the COSMIN recommendations.[33] The overall quality is obtained as the worst count of each set of items for each measurement property. According to the recommendations,[30] the sample size was considered very good if n ≥ 30, adequate if n = 20–29, doubtful if n = 10–19, and inadequate if n ≤ 10.[34] The COSMIN has a set of questions for each reliability and responsiveness. In the case of interpretability or generalizability, there was no such scoring system in COSMIN so we extracted data characteristics for generalizability and minimal important change for interpretability.

Quantitative assessment

Review Manager 5.4 software was used for meta-analysis. The data were standardized by converting pre- and post-scores to MD and SE to enable meta-analysis of ESWT (time and distance) responsiveness to PR, BD, and AO therapy in COPD. The variability in study (i.e., heterogeneity) was reflected by the I^2 in meta-analysis. I^2 value > 75% depicts high heterogeneity,

Table 1: Demographic and clinical characteristics of the included study population.

Measurement property	First author, year of publication	Population			Disease severity	Instrument administration		
		n	Mean age	Gender (female), %		Setting	Country	Language
Reliability	McKeough et al,[41] 2018	66	70	46	Moderate to severe	General hospital	Australia	English
	Ngaiet al,[42] 2017	22	71	50	Moderate	Clinical setting	Australia	English
	Borel et al,[43] 2014	97	63	40	Stable	City hospital	USA, Canada	English
	Hill et al,[6] 2012	24	67	36	Mild to severe	Clinical setting	Australia	English
	McKeough et al,[22] 2011	53	72	35	Moderate	General hospital	Australia	English
	Revill et al,[37] 2010	23	67	43.4	COPD with exertional desaturation	Community hospital	Nottingh-amshire	English
	Revill et al,[21] 2009	44	68	25	severe	Outpatient department	UK	English
	Revill et al,[1] 1999	11	66	47.6	Moderate to severe	Hospital	UK	English
Pooled result		340						
Responsiveness	Zatloukal et al,[40] 2019	531	69.4	43	Stable	Hospital	UK	English
	Altenburg et al,[39] 2015	55	62	42	Stage II	hospital	Netherlands	English
	Borel et al,[43] 2014	255	63	45	Stable	City hospital	USA, Canada	English
	Pepin et al,[38] 2011	210	68	36	Stable	Hospital	Canada, UK	English
	Leung et al,[24] 2010	32	71	30.5	Stage II–IV	General hospital	Australia	English
	Revill et al,[37] 2010	23	67	43.4	COPD with exertional desaturation	Community hospital	Nottingh-amshire	English
	Brouillard et al,[26] 2008	20	65	30	-	Hospital	Canada	English
	Sandland et al,[27] 2008	41	71	29	Severe	Hospital	UK	English
	Pepin et al,[25] 2007	14	64	-	Stage II–III	General hospital	Canada	English
	Eaton et al,[2] 2006	20	71	45	-	City hospital	New Zealand	English
	Pepin et al,[23] 2005	17	65	-	-	General hospital	Canada	English
	Revill et al,[1] 1999	21	66	47.6	Moderate to severe	Hospital	UK	English
Pooled result		1239						

COPD: chronic obstructive pulmonary disease.

50–75% depicts moderate heterogeneity, and < 25% reflects low heterogeneity.[35] The overall quality of evidence of pooled results for systematic review was assessed by the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach.[36]

RESULTS

Literature search

PRISMA 2009 flow diagram in Figure 1 depicts the studies included in this systematic review. Of 791 articles identified through electronic searches,

Table 2: Studies that assessed reliability of the ESWT.

First author, year of publication	Population	Measurement property	Main results
McKeough et al,[41] 2018	Patients with COPD (exercise-induced oxygen desaturation) (n = 66)	Reliability (two ESWT was performed as baseline measure)	The mean differences (coefficient of repeatability) for the ESWTs were 19 seconds (142 seconds) ($p < 0.05$). No participant characteristic predicted the absence of improvement on the second ESWT (AUC ranged from 0.43 to 0.52, all $p > 0.3$).
Ngai et al,[42] 2017	Patients with COPD (moderate) (n = 22)	Reliability (test 1 and test 2 were performed on the same day, 30 minutes apart; test 3 within a week of test 2; and test 4 one week after test 3)	ESWT is repeatable in people with moderate COPD. (The learning effect was not evident). The mean durations of test 1 to test 4 were 368 ± 203 s, 371 ± 182 s, 386 ± 213 s, and 367 ± 223 s, respectively, with no time effect (effect size = 0.18, $p = 0.79$).
Borel et al,[43] 2014	Patients with Stable COPD (n = 97)	Reliability test 1 (ESWT) and test 2 (ESWT) in one week interval (repeatability and reproducibility)	The ICC value at 95% CI was 0.96 (0.95–0.97) for endurance time and 0.95 (0.94–0.96) for endurance distance. The mean difference was -6.7 ± 72.2 s and -7.3 ± 113.1 m.
Hill et al,[6] 2012	Patients with COPD (mild to severe) (n = 24)	Test-retest reliability (two ESWT test conducted within the same day)	Test 1 versus test 2 mean ± SD difference 50 ± 83 s for 18 subjects.
McKeough et al,[22] 2011	Patients with COPD (moderate) (n = 53)	Reliability (test 1 vs. test 2 (pre-exercise training), test 3 vs. test 4 (post-exercise training))	There was no significant difference in duration from first to second test (Δ-2 s [5.8%], ($p = 0.95$)) There was no significant difference in duration from third to fourth test (Δ44 s [8.7%], ($p = 0.07$))
Revill et al,[37] 2010	Patients with COPD (exertional desaturation) (n = 23)	Test-retest reliability (two ESWT test conducted one first and third day, while on air and ambulatory oxygen (reproducibility))	There was no significant difference in duration from first to second test after familiarization (tests performed with supplemental oxygen. The repeatability of the ESWT on oxygen was good with a mean difference of 0.91 m (3.6 s). The LOA (95%) was wide with a small number of individuals showing a larger variation in response. For the 11 patients that repeated the oxygen walk, the mean increase was still significant compared to the performance on air.
Revill et al,[21] 2009	Patients with COPD (severe) (n = 44)	Test-retest reliability (two ESWT test conducted within the same day)	Test 1 versus test 2 mean difference was 12 s (95%CI: -3–28 s). There was no significant difference in duration (Δ12 s [6.2%]). ESWT at 85% on best ISWT, had same Borg dyspnea score at the end of both tests in 77% of sample, Bland–Altman showed LOA from −88 to 112 seconds.
Revill et al,[1] 1999	Patients with COPD (moderate to severe) (n = 44) (11 participated in repeatability)	Test-retest Reliability (ESWT was conducted on first, second, and third days)	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%]). There was a strong relationship between tests 2 and 3 (r = 0.995) with no significant differences between these two tests. The LOA (2SD) between tests 2 and 3 was +15 (42) s ($p > 0.05$).

AUC: area under the curve; ISWT: incremental shuttle walk test; ESWT: endurance shuttle walk test; LOA: limits of agreement; COPD: chronic obstructive pulmonary disease; ICC: intraclass correlation coefficient NSD: normalized standard deviation.

including (n = 435, Web of Science), (n = 239, PubMed), and (n = 117, Cochrane database), 17 met the inclusion criteria, with three reporting on more than one measurement property. After duplicate removal, the remaining articles (n = 117) were screened for eligibility through the title and the abstract; 22 were retrieved full-text, among which five were removed for not meeting the eligibility criteria.

Eight articles included data on the reliability of ESWT in patients with COPD, eight on responsiveness, and four on interpretability. Among 17 studies, nine articles[1,24–27,37–40] were analyzed quantitatively. The remaining studies were analyzed qualitatively as they did not report intraclass correlation coefficient (ICC)[1,6,21,22,37,41,42] or mean change in distance or time.[2,23,43] Only one study reported

Table 3: Studies that assessed responsiveness and MCID of the ESWT.

First author, year of publication	Population	Measurement property	Main results
Zatloukal et al,[40] 2019	Patients with stable COPD (n = 531)	Six weeks of PR and the mean change of ESWT was 342.0 s (95% CI: 312.4–371.6).	MCID of ESWT in COPD after a 6-week PR was between 174 s and 279 s. By the distribution method (0.5 SD) MCID of 173.7 seconds, the global rating of change scale 279.2 s (95% CI: 244.9–313.5) and the ROC method 207 s. The mean change was 341.6 s (347.3) with ES = 2.87 and SRM = 0.98.
Altenburg et al,[39] 2015	Patients with COPD (GOLD stage IV) (n = 55)	Six weeks of PR with or without noninvasive positive pressure ventilation	MCID values of ESWT from different anchors ranged 186–199 s, 76–82%, and 154–164 m. In the distribution-based method, the MCID was 144 s, 61%, and 137 m. The mean change was 121 s (290) and 153 m (274). The ES = 0.18 and SRM = 0.41.
Borel et al,[43] 2014	Patients with stable COPD (n = 255)	Eight weeks of study to bronchodilation with two ESWT (baseline) with tiotropium (one week apart), one after a single dose and one after four weeks of either fluticasone propionate/ salmeterol combination or placebo in addition to tiotropium.	MCID values ranging from 56 to 61 s and 70 to 82 m in endurance time and endurance distance, respectively.
Pepin et al,[38] 2011	Patients with COPD (n = 201)	PR (n = 132 for seven weeks and two ESWT at beginning and end) and bronchodilation (n = 69, two ESWT)	MCID following PR was not estimated but by bronchodilation is 45–85 s (or 60–115 m) was likely at 95% CI and in walking distance (r = 0.53, $p < 0.001$) and endurance time (r = 0.55, $p < 0.001$). The mean change in the PR group was 484.3 s (374.5) with ES of 2.2. The mean change in the bronchodilation group was 90.62 s (120) with SRM of 0.75.
Leung et al,[24] 2010	Patients with COPD (GOLD stage I–IV) (n = 32)	Responsiveness to walking (PR) (eight weeks)	The endurance walking time of the walking training group was (mean change = 439 ± 346 s) (95% CI: 70–483) more than the cycle training group (mean change = 160 ± 204 s). The ES and SRM for the walking group were 2.23 and 1.27, respectively.
Revill et al,[37] 2010	Patients with COPD (exertional desaturation) (n = 23)	Responsiveness to PR (oxygen therapy) (six weeks)	The mean difference (95% CI) between ESWT distances was 0.91 m (47–49) and between endurance time was 3.6 s (63–56). Compared to the air walk the mean increases on oxygen were 80.5 m and 79.5 m and in walking time were 95 s and 98 s (days two and three, respectively). The mean change in ESWT distance equated to an increase of 33 (46%) while breathing oxygen. Seventeen patients (74%) had 10% improvement in walking distance with oxygen. The SRM was 0.73.
Brouillard et al,[26] 2008	Patients with COPD (n = 20)	Responsiveness to salmeterol	There was a significant improvement in ESWT results (difference in endurance time salmeterol-placebo: 117 ± 208 s; $p = 0.02$) and walking distance (difference in walking distance salmeterol-placebo: 160 ± 277 m; $p = 0.02$) with salmeterol inhalation. The SRM was 0.56.
Sandland et al,[27] 2008	Patients with COPD (severe hypoxemic) (n = 41)	Responsiveness to AO therapy (seven weeks)	There was no significant difference in the ESWT (112.0 ± 217.1 m) ($p < 0.05$) and SRM was 0.52 with a moderate ESWT (0.69).
Pepin et al,[25] 2007	Patients with COPD (GOLD stage II–III) (n = 14)	Responsiveness to bronchodilation	There was a significant improvement in the distance walked on the ESWT: (mean change in walking distance was 144 ± 219 m) The EWST was more responsive than the 6MWT for detecting changes in exercise performance following bronchodilation. The ES was moderate = 0.78 and the SRM was 0.66.

Table 3: Studies that assessed responsiveness and MCID of the ESWT.

-continued

First author, year of publication	Population	Measurement property	Main results
Eaton et al,[2] 2006	Patients with COPD (n = 20)	Responsiveness to PR (seven weeks)	There was a significant improvement in ESWT of 92% (302m, 95% CI: 104–501) in ESWT distance. The ES was 0.54 (SRM = 0.78).
Pepin et al,[23] 2005	Patients with COPD (n = 17)	Responsiveness to bronchodilation	There was a significant improvement in walking endurance time (endurance time ipratropium bromide placebo: 164 ± 177 s, $p < 0.01$). A 17% fall in quadriceps twitch force was observed after cycling, whereas no significant change was seen after walking. (SRM = 0.93).
Revill et al,[1] 1999	Patients with COPD (moderate to severe) (n = 44) (21 participated in the responsiveness measure)	Responsiveness to PR (seven weeks)	The mean percentage improvement in endurance capacity was 160 (110%) There was a large ES for the ESWT (2.9).

PR: pulmonary rehabilitation; SRM: standardized response mean (mean change/SD of change); ES: effect size; ISWT: incremental shuttle walk test; ESWT: endurance shuttle walk test; GOLD: Global Initiative for Chronic Obstructive Lung Disease; 6MWT: six-minute walk test; MCID: minimal clinical important difference; AO: ambulatory oxygen; ROC: receiver-operating characteristic.

the ICC which cannot be pooled for conducting a meta-analysis.[43]

Generalizability

The COSMIN criteria for generalizability in Table 1 included mean age, distribution of gender, description of treatment, disease characteristics, country, study settings, language, method of patient selection, and percentage of responses missing.[44] All 17 articles met the criteria of generalizability; however, two studies did not mention the gender distribution,[23,25]and three studies did not mention the disease severity.[2,23,26]

Reliability

The eight included studies[1,6,21,22,37,41–43,] depict the test-retest reliability and repeatability of ESWT in COPD [Table 2]. ICCs were reported in one study[43] as 0.96 (95% CI: 0.95–0.97) for endurance time and 0.95 (95% CI: 0.94–0.96) for endurance distance to verify the agreement between test and retest. The correlation coefficient was 0.92 and 0.90 for endurance time and distance, respectively.[43] In a study including patients with moderate to severe COPD with exercise-induced oxygen desaturation, the coefficient of repeatability of ESWT was 19 s.[41] In another study, the coefficient of repeatability was narrow compared to the ISWT.[6] There was an insignificant difference in test-retest reliability and repeatability among the two studies.[21,22] There was good repeatability of ESWT to AO in COPD patients with exertional desaturation with wider limits of agreement, and MD of 0.91 m.[37] Ngai et al,[42] reported that ESWT is repeatable in moderate COPD without a learning effect. In one study, a significant difference was observed in the first two tests while no statistically significant difference was observed in tests 2 and 3.[1]

Responsiveness

Eight studies described the responsiveness of ESWT in patients with COPD summarized [Table 3]. Two studies,[1,2] reported the responsiveness to PR with effect size (ES) moderate and large, respectively. Two studies showed large SRM (1.27) sensitivity to change,[24] and 0.93,[23] whereas the remaining studies showed a moderate sensitivity index.[25–27]

Four studies[38–40,43] assessed the interpretability of the ESWT [Table 3]. There were no predictive equations for the distance of ESWT in either of the studies included. Borel et al,[43] depicted MCID values ranging from 56–61 s and 70–82 m in response to BD. The two studies depicted responses to the PR with MCID values of 144 s and 137 m,[39] and 174–279 s.[40] Pepin et al,[38] reported MCID values 45–85 s and 60–115 m response to BD.

Quality assessment

Table 4 depicts the methodological quality assessment of the included studies. The quality assessment is graded in accordance with the COSMIN checklist pertaining to reliability and

Table 4: Methodological quality of the included studies.

First author, year of publication	Reliability (quality)	Rating	Responsiveness (quality)	Rating
Zatloukal et al,[40] 2019			Very good	?
McKeough et al,[41] 2018	Inadequate	?		
Ngai et al,[42] 2017	Inadequate	?		
Altenburg et al,[39] 2015			Adequate	?
Borel et al,[43] 2014	Very good	+	Very good	?
Hill et al,[6] 2012	Inadequate	?		
McKeough et al,[22] 2011	Inadequate	?		
Pepin et al,[38] 2011			Adequate	+
Leung et al,[24] 2010			Adequate	+
Revill et al,[37] 2010	Inadequate	?	Doubtful	+
Revill et al,[21] 2009	Inadequate	?		
Brouillard et al,[26] 2008			Adequate	+
Sandland et al, [27] 2008			Doubtful	?
Pepin et al,[25] 2007			Inadequate	+
Eaton et al,[2] 2006			Adequate	+
Pepin et al,[23] 2005			Doubtful	+
Revill et al,[1] 1999	Doubtful	?	Adequate	+

(+): sufficient; (?): indeterminate.

Table 5: Summary of findings.

Measurement property	Summary or pooled result	Overall rating	Quality of evidence
Responsiveness	PR – Mean difference (time) = 303.19 s (95% CI 175.63–430.75; $p < 0.001$). Sample size = 756 BD – Mean difference (distance)= 168.62 m (95% CI 117.03–220.21; $p < 0.001$). Sample size = 103 AO – Mean difference (time)= 129.04 s (95% CI: 47.98–210.09; $p = 0.002$). Sample size- 64 AO – Mean difference (distance) = 80.71 m (95% CI: 38.66–122.76; $p < 0.001$). Sample size = 64	+	Moderate

PR: pulmonary rehabilitation; BD: bronchodilation; AO: ambulatory oxygen; (+): sufficient.

responsiveness.[29] The articles were individually and independently judged, further resolving the inconsistencies through discussion. The COSMIN checklist consists of a set of questions for all the measurement properties. The questions evaluated the methodological quality of each study as very good, adequate, inadequate, or doubtful.[30] Rating was done for each measurement property and graded as sufficient (+), insufficient (-), and indeterminate (?) as per COSMIN recommendations in Table 4. In reliability studies, ICC ≥ 0.70 is sufficient, ICC < 0.70 (insufficient), and indeterminate if ICC was not reported. In studies of responsiveness, the result being in accordance with the hypothesis or the AUC ≥ 0.70 is sufficient, and if the result is not as per the hypothesis or AUC < 0.70, it is rated insufficient. If the hypothesis is not defined in the studies, it is rated as indeterminate. The quantitative or pooled data of responsiveness studies after meta-analysis are shown in Table 5. The data were pooled based on the MD obtained.

Quantitative assessment

The whole measurement properties were rated as sufficient, insufficient, and indeterminate through the COSMIN 75% rule. The GRADE approach was applied to the pooled data of responsiveness as high, moderate, low, or very low evidence. The two authors (S.A and A.M) assessed the article and resolved doubts through discussion.

Meta-analysis of responsiveness

The MD (time and distance) was significant for PR ($p < 0.001$), BD ($p < 0.001$), and AO with ESWT time

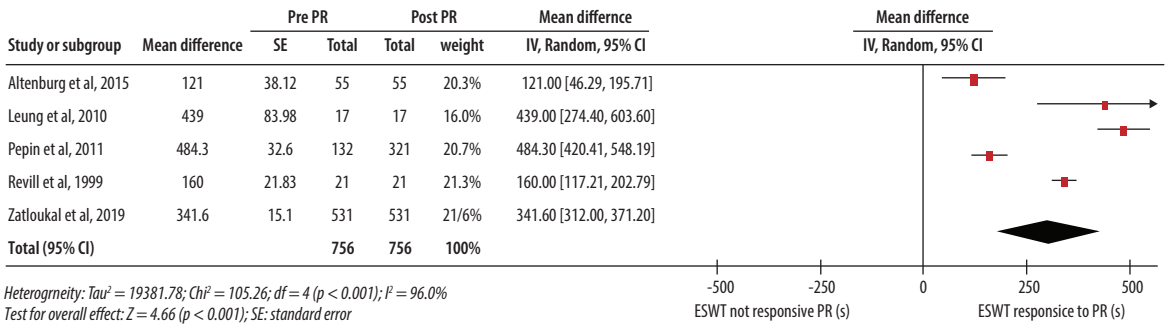

Figure 2: Responsiveness of endurance shuttle walk test (ESWT) following pulmonary rehabilitation (PR).

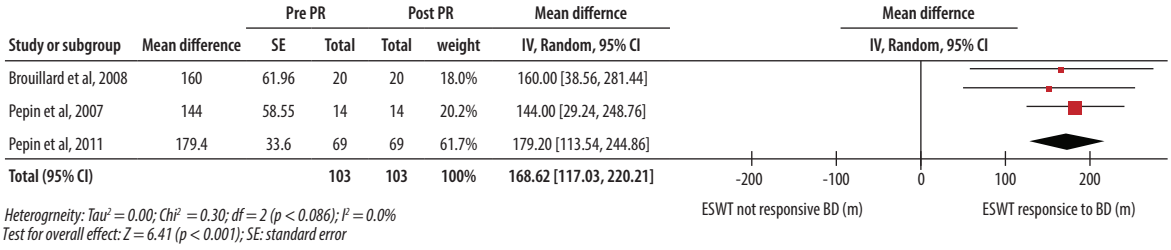

Figure 3: Responsiveness of endurance shuttle walk test (ESWT) following bronchodilator (BD).

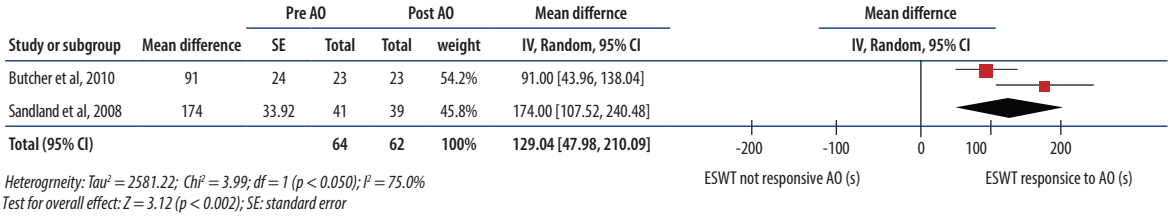

Figure 4: Responsiveness of endurance shuttle walk test (ESWT) following ambulatory oxygen (AO).

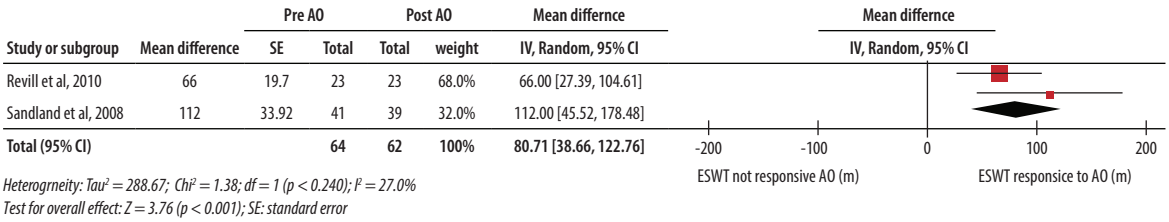

Figure 5: Responsiveness of endurance shuttle walk test (ESWT) following ambulatory oxygen (AO).

(p = 0.002) and ESWT distance (p < 0.001). The results indicated high heterogeneity (I^2 = 96.0% [PR] and I^2 = 75.0% [AO]). The heterogeneity in the studies could be due to clinical factors such as gender, age, and disease severity. It can also be due to a smaller number of studies and the quality of the selected study.

Responsiveness to PR

The five studies reporting the mean change in the ESWT (time) to PR were included in the meta-analysis.[1,24,38–40] The ES ranged from 0.18 to 2.90. Figure 2 shows the MD in ESWT time following PR (mean 303.19 s, 756 participants, 95% CI: 175.63–430.75 s; p < 0.001) in COPD patients.

Responsiveness to BD

Three studies including 103 participants reporting the mean change in the ESWT (distance) to BD were included in the meta-analysis with SRM ranging from 0.56 to 0.93.[25,26,38] The MD in ESWT distance following BD was 168.62 m (95% CI: 117.03–220.21m; p < 0.001) in COPD patients [Figure 3].

Responsiveness to AO

Two studies reporting the mean change in the ESWT (time) to AO were included in the meta-analysis with an ES of 0.69 and 1.08.[27,37] The MD in ESWT time (mean = 129.04 s, 64 participants, 95% CI: 47.98–210.09 s; $p = 0.002$) [Figure 4], and ESWT distance (mean = 80.71 m, 64 participants, 95% CI: 38.66–122.76 m; $p < 0.001$) [Figure 5].

DISCUSSION

This systematic review and meta-analysis evaluated the reliability and responsiveness of ESWT in patients with COPD. It reported evidence that ESWT is a reliable and responsive test to evaluate functional exercise capacity in patients with COPD. The qualitative evidence suggested low and moderate quality evidence for reliability and responsiveness respectively. The meta-analysis of responsiveness depicted ESWT as a responsive test following PR, BD, and AO in COPD patients. However, the results must be extrapolated in the light of caution for PR due to the high heterogeneity obtained in the included studies.

The reliability of ESWT could not be analyzed quantitatively as there was only one study evaluating ESWT performance with ICC of 0.96 and 0.95 for endurance time and distance, respectively. Pearson's correlations also confirmed the reproducibility of ESWT performance following eight weeks of BD in COPD patients.[43] There was also a strong linear relationship between endurance time (correlation coefficients (r) = 0.92) and distances (r = 0.90) obtained at ESWT 1 and ESWT 2. ESWT has good repeatability and reproducibility with prior familiarization.[45] ESWT was also reported to be sensitive to therapeutic intervention.[45] The repeatability of the ISWT and ESWT in COPD patients with exercise-induced dyspnea showed a small difference between the first and second tests.[41] There was a statistically significant difference between the end-test dyspnea score for the ISWT and ESWT with a repeat test.[41] It has also been proven that ESWT has favorable within-day repeatability as compared with 6MWT which reported mean increase of 0–17% within one day.[21,46] The learning effect was not reported even when ESWT was repeated within the same week or one week apart. Being externally paced, it potentially improves test-retest repeatability.[1] However, external pacing does not always adequately remove the learning effect as demonstrated in ISWT.[19,22] Revill et al,[1] and McKeough et al,[22] reported insignificant changes in test-retest reliability to ESWT. The measurements of oxygen saturation (SpO_2), HR, and modified Borg dyspnea scale were repeated well during the test.[6,22] One investigation even observed that ESWT was more repeatable than ISWT.[22] The reliability of ESWT has not been studied in any other chronic respiratory diseases.[19]

The responsiveness of ESWT following PR[1,2,24,38–40] showed a low to large ES (0.18–2.9) following BD,[23,26] SRM was 0.56 to 0.93, and larger ES (1.08 and 0.69) was reported following AO.[27,37] The studies[1,22,47,48] assessing the response of ESWT and ISWT following PR resulted in a significant improvement following both tests.[1,22,47,48] The response to ESWT was greater in all four studies whereas in two the response to ISWT did not reach its MCID.[1,22] Another investigation,[2] indicated that both 6MWT and ESWT reported a significant response above the MCID, the ESWT was reportedly more responsive to PR than 6MWT[2] and ISWT.[49] ESWT was even found to be more responsive following BD compared to 6MWT and endurance cycle ergometer tests.[23,25] The improvements observed in the performances of two studies assessing ESWT were exceeding MCID, following BD therapy.[23,25]

The supplemental oxygen might have an impact on the performance of ESWT.[27,37,50,51] The studies reflected a difference regarding whether the cylinder was carried by the operator or the participant. The performance changes with oxygen compared with air ranged from 70–174 s,[27,50] or 32–76% of the walking time.[51] Large improvements in ESWT distance were reported with the application of supplemental AO in a study (mean increase = 275 m, 95% CI: 197–352). It was a non-blinded study with the oxygen cylinder being carried by the operator and all participants were known as oxygen "responders" (increase in ESWT of ≥10% on oxygen).[52] Conclusively, the degree of increase in ESWT performance with oxygen shall not be inferred in all the patients with COPD. Three studies reported the potential use of SpO_2 in recording exertional desaturation during ESWT in either AO assessment or breathing air.[27,37,50] In comparison, a study of 6MWT reported greater desaturation with the ESWT in COPD.[37] MCID scores reflect changes in clinical intervention which are meaningful for the patient.[53] The MCID was estimated either with the

anchor (45–85 s, 186–199 s, 174–279 s, and 56–61 s),[38–40,43] or distribution-based method (203 m, 137 m, 173.7 m, and 81 m).[38–40,43] The distribution-based considers measurement error but lacks clinical explanation in different samples. The anchor-based approach offers the clinical significance of MCID, through the external changes to an anchor, but does not consider measurement error.[54] There were no studies that evaluated the correlation of ESWT and hospitalization or survival in COPD.[19] There is no accord about the most appropriate test to be used in patients with COPD as all tests have distinct primary outcomes (such as endurance time and distance). In turn, it reflects various physiological parameters. Consequently, it becomes challenging to compare results across analyses from the available literature.

Most of the studies evaluated were rated as indeterminate for reliability while sufficient evidence was reported for responsiveness. The summarized or pooled result depicted sufficient and moderate overall quality of evidence for responsiveness. It is to be noted that the COSMIN checklist allows sufficient flexibility in quality interpretation making the results of the study's quality absolute.[44] The meta-analysis of ESWT responsiveness was significant following PR, BD, and AO in patients with COPD.

Strengths and limitations

This is the first study to systematically review and meta-analyze the reliability and responsiveness of ESWT in patients with COPD. This review provides evidence for practitioners to use this test in their routine clinical practice as walking is more representative of ADL than cycling in COPD patients. The potential of ESWT in providing a more responsive change to PR, BD, and AO is crucial for clinicians and patients as it is a fact of interest in assessing the outcome measures. Furthermore, this study provides equivocal evidence for the use of the ESWT over the ISWT in determining endurance capacity. The database search engines were limited so few relevant articles might have been missed. The sample size was limited which might have compromised the results of the meta-analysis. Most of the reliability studies did not report the ICC.

CONCLUSION

This review found that ESWT is suitable before and after PR, BD, and AO in COPD patients. The quantitative analysis must be deduced with caution in clinical settings and in research due to the high heterogeneity obtained in the included studies. ESWT lacks enough studies reporting ICC to reach a definitive conclusion as a reliable tool in patients with COPD. Further research examining the reliability of ESWT is required by reporting ICC values.

Disclosure

The authors declared no conflicts of interest. No funding was received for this study.

REFERENCES

1. 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. ThoraxMar 1999;54(3):213-222.
2. 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.
3. Puente-Maestu L, Villar F, de Miguel J, Stringer WW, Sanz P, Sanz ML, et al. Clinical relevance of constant power exercise duration changes in COPD. Eur Respir J 2009 Aug;34(2):340-345.
4. 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 Dec;47(12):1019-1024.
5. Singh SJ, Puhan MA, Andrianopoulos V, Hernandes NA, Mitchell KE, Hill CJ, et al. An official systematic review of the European respiratory society/American thoracic society: measurement properties of field walking tests in chronic respiratory disease. Eur Respir J 2014 Dec;44(6):1447-1478.
6. Hill K, Dolmage TE, Woon L, Coutts D, Goldstein R, Brooks D. Comparing peak and submaximal cardiorespiratory responses during field walking tests with incremental cycle ergometry in COPD. Respirology 2012 Feb;17(2):278-284.
7. Vos T, Allen C, Arora M, Barber RM, Bhutta ZA, Brown A, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990 – 2015: a systematic analysis for the global burden of disease study 2015. Lancet 2016;388(10053):1545-1602.
8. Wang H, Naghavi M, Allen C, Barber RM, Bhutta ZA, Carter A, et al. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980 – 2015: a systematic analysis for the global burden of disease study 2015. Lancet 2016 Oct 8;388(10053):1459-1544.
9. Shukla RK, Kant S, Bhattacharya S, Mittal B. Association of cytokine gene polymorphisms in patients with chronic obstructive pulmonary disease. Oman Med J 2012 Jul;27(4):285-290.
10. Pauwels R, Buist S, Calverley P, Jenkins C, Hurd S. Global strategy for the diagnosis, management and prevention of chronic obstructive pulmonary disease. NHBLI/WHO global initiative for chronic obstructive lung disease (GOLD) workshop summary. Am J Respir Crit Care Med 2001 Apr;163(5):1256-1276.
11. Orooj M, Moiz JA, Mujaddadi A, Ali MS, Talwar D. Effect of pulmonary rehabilitation in patients with asthma COPD overlap syndrome: a randomized control trial. Oman Med J 2020 Jun;35(3):e136.
12. 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.

13. Vilaro J, Rabinovich R, Gonzalez-deSuso JM, Troosters T, Rodríguez D, Barberà JA, et al. Clinical assessment of peripheral muscle function in patients with chronic obstructive pulmonary disease. American Journal of Physical Medicine &Rehabilitation 2009;(January):39-46.
14. Arne M, Janson C, Janson S, Boman G, Lindqvist U, Berne C, et al. Physical activity and quality of life in subjects with chronic disease: chronic obstructive pulmonary disease compared with rheumatoid arthritis and diabetes mellitus. Scand J Prim Health Care 2009;27(3):141-147.
15. Viegi G, Pistelli F, Sherrill DL, Maio S, Baldacci S, Carrozzi L. Definition, epidemiology and natural history of COPD. Eur Respir J 2007 Nov;30(5):993-1013.
16. Moloney ED, Clayton N, Mukherjee DK, Gallagher CG, Egan JJ. The shuttle walk exercise test in idiopathic pulmonary fibrosis. Respir Med 2003 Jun;97(6):682-687.
17. Ushiki A, Nozawa S, Yasuo M, Urushihata K, Yamamoto H, Hanaoka M, et al. Associations between the distance covered in the incremental shuttle walk test and lung function and health status in patients with chronic obstructive pulmonary disease. Respir Investig 2017 Jan;55(1):33-38.
18. Ishrat R, Mujaddadi A, Ali MS, Talwar D, Hussain ME. Reliability and responsiveness of incremental shuttle walk test to estimate exercise capacity in patients with pulmonary arterial hypertension. Comp ExercPhysiol 2020;16(3):179-185.
19. Holland AE, Spruit MA, Troosters T, Puhan MA, Pepin V, Saey D, et al. An official European Respiratory Society / American Thoracic Society technical standard: field walking tests in chronic respiratory disease. European Respiratory Journal 2014;44(6):1428-1446.
20. Parreira VF, Janaudis-Ferreira T, Evans RA, Mathur S, Goldstein RS, Brooks D. Measurement properties of the incremental shuttle walk test. a systematic review. Chest 2014 Jun;145(6):1357-1369.
21. 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.
22. 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 Jan;90(1):35-39.
23. 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 Dec;172(12):1517-1522.
24. Leung RW, Alison JA, McKeough ZJ, Peters MJ. Ground walk training improves functional exercise capacity more than cycle training in people with chronic obstructive pulmonary disease (COPD): a randomised trial. J Physiother 2010;56(2):105-112.
25. Pepin V, Brodeur J, Lacasse Y, Milot J, Leblanc P, Whittom F, et al. Six-minute walking versus shuttle walking: responsiveness to bronchodilation in chronic obstructive pulmonary disease. Thorax 2007 Apr;62(4):291-298.
26. Brouillard C, Pepin V, Milot J, Lacasse Y, Maltais F.Endurance shuttle walking test: responsiveness to salmeterol in COPD. European Respiratory Journal 2008;31(3):579-584.
27. Sandland CJ, Morgan MDL, Singh SJ. Detecting oxygen desaturation in patients with COPD: incremental versus endurance shuttle walking. Respiratory Medicine 2008;102(8):1148-1152.
28. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Systematic Reviews 2015 Dec;4(1):1-9.
29. Mokkink LB, Prinsen CA, Patrick D, Alonso J, Bouter LM, de Vet HC, et al. COSMIN study design checklist for patient-reported outcome measurement instruments. Amsterdam, The Netherlands 2019:1-32.
30. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res 2012 May;21(4):651-657.
31. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. Stat Med 1998 Jan;17(1):101-110.
32. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. J Clin Epidemiol 2000 May;53(5):459-468.
33. Mokkink LB, Prinsen C, Patrick DL, Alonso J, Bouter L, de Vet HC, et al. COSMIN methodology for systematic reviews of patient - reported outcome measures (PROMs). User Manual 2018;78(1).
34. Jones BE, Herman DD, Dela Cruz CS, Waterer GW, Metlay JP, Ruminjo JK, et al. Summary for clinicians: clinical practice guideline for the diagnosis and treatment of community-acquired pneumonia. Annals of the American Thoracic Society 2020 Feb;17(2):133-138.
35. Zlowodzki M, Poolman RW, Kerkhoffs GM, Tornetta PIII, Bhandari M; International Evidence-Based Orthopedic Surgery Working Group. How to interpret a meta-analysis and judge its value as a guide for clinical practice. Acta Orthop 2007 Oct;78(5):598-609.
36. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. Chinese J Evidence-Based Med 2009;9(1):8-11.
37. 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.
38. Pepin V, Laviolette L, Brouillard C, Sewell L, Singh SJ, Revill SM, et al. Significance of changes in endurance shuttle walking performance. Thorax 2011 Feb;66(2):115-120.
39. Altenburg WA, Duiverman ML, Ten Hacken NH, Kerstjens HA, de Greef MH, Wijkstra PJ, et al. Changes in the endurance shuttle walk test in COPD patients with chronic respiratory failure after pulmonary rehabilitation: the minimal important difference obtained with anchor- and distribution-based method. Respir Res 2015 Feb;16(1):27.
40. Zatloukal J, Ward S, Houchen-Wolloff L, Harvey-Dunstan T, Singh S. The minimal important difference for the endurance shuttle walk test in individuals with chronic obstructive pulmonary disease following a course of pulmonary rehabilitation. Chron Respir Dis 2019;16.
41. McKeough Z, Leung R, Neo JH, Jenkins S, Holland A, Hill K, et al. Shuttle walk tests in people with COPD who demonstrate exercise-induced oxygen desaturation: An analysis of test repeatability and cardiorespiratory responses. Chron Respir Dis 2018 May;15(2):131-137.
42. Ngai SP, Spencer LM, Jones AY, Alison JA. Repeatability of the endurance shuttle walk test in people with chronic obstructive pulmonary disease. Clin Respir J 2017 Nov;11(6):875-880.
43. BorelB, PepinV, MahlerDA, NadreauÉ, MaltaisF. Prospective validation of the endurance shuttle walking test in the context of bronchodilation in COPD. Eur Respir J 2014 Nov;44(5):1166-1176.
44. Prinsen CA, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HC, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res 2018 May;27(5):1147-1157.
45. Fotheringham I, Meakin G, Punekar YS, Riley JH, Cockle SM, Singh SJ. Comparison of laboratory- and field-based exercise tests for COPD: a systematic review. Int J Chron

Obstruct Pulmon Dis 2015Mar;10:625-643.

46. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. Am J Respir Crit Care Med 2002;166(1):111-117.
47. 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 May;9(2):99-106.
48. O'Farrell R, Gargoum F, O'Connor B, Khan F, Magro E, O'Mahony A, et al. Comparison of incremental and endurance shuttle walk testing after pulmonary rehabilitation. American Thoracic Society 2011;183:A2024.
49. Borel B, Provencher S, Saey D, Maltais F. Responsiveness of various exercise-testing protocols to therapeutic interventions in COPD. Pulmonary Medicine 2013;2013.
50. Revill SM, Singh SJ, Morgan MD. Randomized controlled trial of ambulatory oxygen and an ambulatory ventilator on endurance exercise in COPD. Respiratory Medicine 2000;94(8):778-783.
51. Eves ND, Petersen SR, Haykowsky MJ, Wong EY, Jones RL. Helium-hyperoxia, exercise, and respiratory mechanics in chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2006 Oct;174(7):763-771.
52. Dyer F, Callaghan J, Cheema K, Bott J. Ambulatory oxygen improves the effectiveness of pulmonary rehabilitation in selected patients with chronic obstructive pulmonary disease. Chron Respir Dis 2012 May;9(2):83-91.
53. Cook CE. Clinimetricscorner: the minimal clinically important change score (MCID): a necessary pretense. J Man ManipTher 2008;16(4):E82-E83.
54. Rai SK, Yazdany J, Fortin PR, Aviña-Zubieta JA. Approaches for estimating minimal clinically important differences in systemic lupus erythematosus. Arthritis Res Ther 2015 Jun;17(1):143.