Systematic Review and Meta-analysis

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-110025, India

Received: 18 June 2021
Accepted: 27 February 2022

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

DOI 10.5001/omj.2023.21

ABSTRACT

The endurance shuttle walk test (ESWT) is a simple, acceptable, field-based test which was first established in 1999 to measure endurance exercise capacity in patients with chronic obstructive pulmonary disease (COPD). The aim of this systematic review is to examine the reliability and responsiveness of ESWT in COPD. Among 791 articles identified through electronic databases, 17 articles were included in the present review. Qualitative and quantitative analyses were conducted according to PRISMA and as per COSMIN the quality of the studies were graded as low for reliability and moderate for responsiveness. Qualitative analysis indicated inadequate evidence for reliability of the ESWT in patients with COPD. The meta-analysis found strong evidence that ESWT was responsive to change following pulmonary rehabilitation (PR) with an estimated mean difference (ESWT time, seconds) 303.19 s (95% C.I 175.63 to 430.75 s; $p < 0.00001$), ambulatory oxygen (AO) with mean difference (ESWT time, seconds)129.04s (95% C.I

47.98 to 210.09s; $p = 0.002$) and (ESWT mean distance, meters) 80.71m (95% CI 38.66 to 122.76 m; $p = 0.0002$).The ESWT was also responsive to Bronchodilation with mean difference of 168.62m (95% C.I 117.03to 220.21m; $p < 0.00001$).The present finding suggests strong potential of ESWT as responsive test in COPD, but to draw definitive conclusion regarding reliability of the ESWT, further research is needed in this population.

Keywords: Endurance shuttle walk test, ESWT, Chronic obstructive pulmonary disease, COPD, Responsiveness, Reliability.

INTRODUCTION

The Endurance shuttle walk test (ESWT) is a test of endurance capacity, described first in 1999 in chronic obstructive pulmonary disease (COPD) patients.[1] The measurement of endurance capacity, by the ESWT is 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 evaluation of exercise tolerance in COPD.[3] The ESWT was designed to complement the incremental shuttle walk test (ISWT) and uses the same 10m shuttle course.[4] The pace of the ESWT is traditionally calculated at predefined percentage of peak performance on the ISWT around 70–85% estimated peak oxygen consumption (VO_2peak).[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 (HR> 85%) predicted maximum,[4] 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 causes of worldwide mortality and morbidity is attributed to COPD ranked eighth causing disability and disease burden in 2015 by disability-adjusted life years (DALYs).[7–9] The symptoms of breathlessness and reduced exercise capacity is characteristic 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 in COPD impairs exercise tolerance, peripheral muscle endurance and quality of life.[13–15]

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

Measurement property of any test is vital in 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 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 insignificant difference in test-retest reliability and repeatability in 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 standardized response mean (SRM) ranging from 0.52 to 1.27.[1,2,23–27] However, till date, there is no systematic review and meta-analysis which can qualitatively and quantitatively summarize the findings obtained from these studies. Therefore, the objective of this systematic review is to examine the reliability and responsiveness of ESWT in patients with COPD.

METHODS

Protocol and Registration

The protocol for this systematic review was registered in the International Prospective Register of Systematic Reviews (CRD42020217847) on November 28, 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 in this systematic review.[29]

Eligibility criteria

The studies describing 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 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, Pub Med, and Cochrane (via CENTRAL) till 18th March, 2021.The keywords used were, "endurance shuttle walk test", "chronic obstructive pulmonary disease", "ESWT", "COPD", "responsiveness", "reliability". The systematic search of articles were 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 minimal change in score when 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

Literature search was conducted and the studies were imported on Mendeley Desktop, reference manager. The articles from all the databases (web of science, PubMed, 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, the review authors (S.A, A.M) retrieved full-text articles and individually evaluated them. In case of disagreement at any stage it was resolved through discussion with the third reviewer (J.M).

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 property (reliability, responsiveness, MCID) and the main findings observed. Review was carried out in accordance with PRISMA statement.[28] The mean difference (MD) and standard error (SE), was extracted for 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 was used for assessing the risk of bias (RoB) of the selected studies,[29] and the data was extracted for the reliability and responsiveness properties. COSMIN consists of 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 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 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 when n=(10-19) and inadequate if n<=10.[34] The COSMIN has a set of questions for each reliability and responsiveness. In 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 was 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) is reflected by the I^2 in meta-analysis. I^2 value> 75% depicts high heterogeneity, 50-75% depicts moderate heterogeneity and <25% reflects low heterogeneity.[35] The overall quality of evidence of pooled result 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, including (n=435, web of science), (n=239, PubMed), and (n= 117, Cochrane data base), 17 met the inclusion

criteria, with 3 reporting on more than one measurement property. After duplicates removal, the remaining articles (n=117) were screened for eligibility through the title and the abstract and 22 were retrieved full text, among which 5 were removed as it did not meet the eligibility criteria. Eight articles included data on reliability of ESWT in patients with COPD, eight on responsiveness, and four on interpretability. Among 17 studies, 9 articles[1,24–27,37–40] were analyzed quantitatively. The remaining studies were analyzed qualitatively as they didn't report intra-class coefficient (ICC),[1,6,21,22,37,41,42] or mean change in distance or time.[2,23,43] Only one study,[43] reported the ICC which cannot be pooled for conducting meta-analysis.

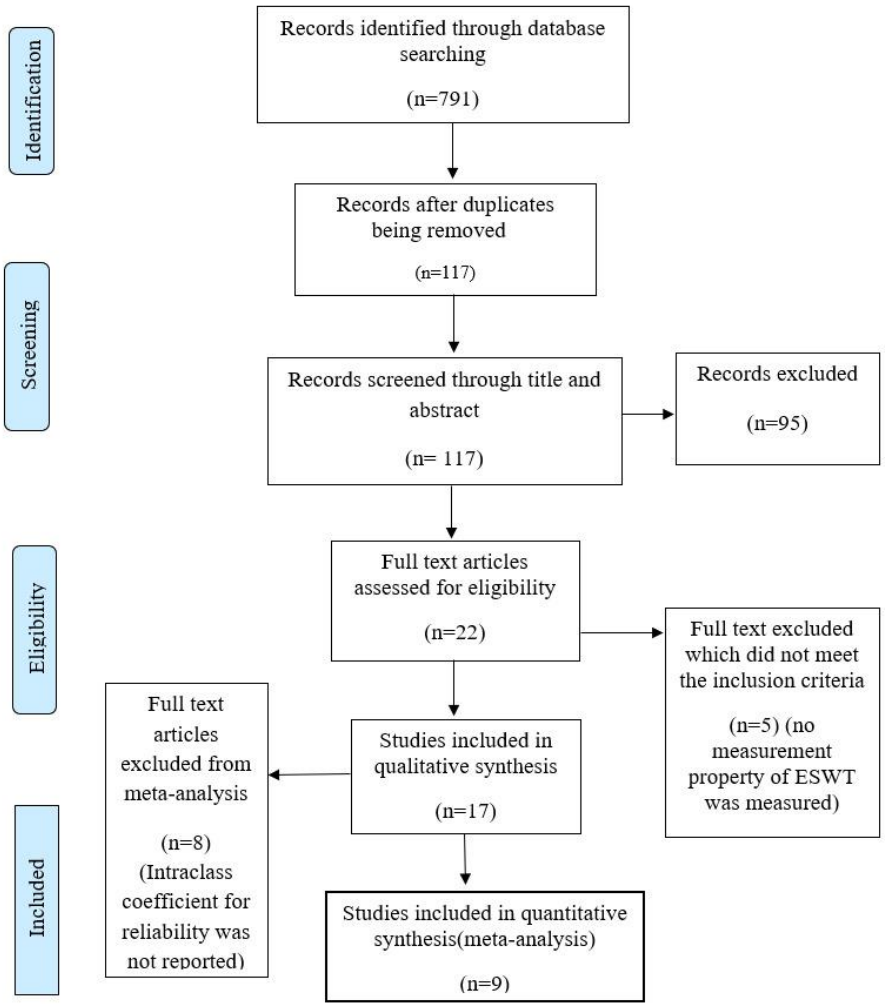

Figure 1: PRISMA flow chart-study selection process.

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 the seventeen included articles met the criteria of generalizability, however two studies did not mention about the distribution of gender,[23,25] three studies did not mention the disease severity.[2,23,26]

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

Measurement property	**First author (year of publication)**	**Population**			**Disease severity**	**Instrument administration**		
Reliability		**n**	**Mean Age**	**Gender % (female)**		**Setting**	**Country**	**Language**
	McKeough, Z.(2018)	66	70	46%	Moderate to severe	General hospital	Australia	English
	Ngai, S.P.C. (2017)	22	71	50%	moderate	Clinical setting	Australia	English
	Borel, B. (2014)	97	63	40%	stable	City hospital	USA, Canada	English
	Hill, K. (2012)	24	67	36%	Mild to severe	Clinical setting	Australia	English
	McKeough, Z.J.(2011)	53	72	35%	moderate	General hospital	Australia	English
	Butcher, G. (2010)	23	67	43.4%	COPD with exertional desaturation	Community hospital	Nottinghamshire	English
	Revill, S.M.(2009)	44	68	25%	severe	Outpatient dept	UK	English
	Revill,S.M. (2009)	11	66	47.6%	Moderate to severe	hospital	UK	English
Pooled Result		**340**						
Responsiveness	Zatloukal, J. (2019)	531	69.4	43%	stable	Hospital	UK	English
	Altenburg, W.A. (2015)	55	62	42%	Stage II	hospital	Netherlands	English
	Borel, B. (2014)	255	63	45%	stable	City hospital	USA, Canada	English
	Pepin, V. (2011)	210	68	36%	Stable	Hospital	Canada, UK	English

	Leung, R.W.M (2010)	32	71	30.5%	Stage II-IV	General hospital	Australia	English
	Butcher, G. (2010)	23	67	43.4%	COPD with exertional desaturation	Community hospital	Nottinghamshire	English
	Brouillard, B.(2008)	20	65	30%	-	hospital	Canada	English
	Sandland, S.J. (2008)	41	71	29%	severe	hospital	UK	English
	Pepin, V. (2007)	14	64	-	Stage II-III	General hospital	Canada	English
	Eaton, T. (2006)	20	71	45%	-	City hospital	New Zealand	English
	Pepin, V. (2005)	17	65	-	-	General hospital	Canada	English
	Revill, S.M. (1999)	21	66	47.6%	Moderate to severe	hospital	UK	English
Pooled Result		**1239**						

Reliability

The eight included studies[1,6,21,22, 37,41-43,] depict the test-retest reliability and repeatability of ESWT in COPD (Tables 2). ICCs was 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 agreement between test and re-test. The correlation coefficient was 0.92 and 0.90 for endurance time and distance respectively.[43]In a study,[41] with moderate to severe COPD with exercise-induced oxygen desaturation, the coefficient of repeatability of ESWT was 19 seconds whereas in another study,[6] the coefficient of repeatability was narrow in comparison to the ISWT. There was insignificant difference in test-retest reliability and repeatability among two studies.[22,21] There was good repeatability of ESWT to AO in COPD patients with exertional desaturation with wider limits of agreement (LOA), and MD of 0.91m.[37] Ngai et al,[42] reported that ESWT is repeatable in moderate COPD without learning effect. In one study[1] significant difference was observed in first two tests while no statistical significant difference was observed in test 2 and 3.

Table 2: Studies that assessed Reliability of the ESWT.

First author (year of publication)	Population	Measurement Property	Main Result
McKeough, Z.(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,S.P.C. (2017)	Patients with COPD (moderate) (n=22)	Reliability (Test 1 and Test 2 was performed on the same day, 30 min apart; Test 3 within a week from 2; Test 4 one week after Test 3.)	ESWT is repeatable in people with moderate COPD. (Learning effect was not evident) The mean durations of Test1 to Test4 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, B. (2014)	Patients with Stable COPD (n=97)	Reliability test 1(ESWT) and test 2 (ESWT) in 1 week interval (repeatability and reproducibility)	The ICC value at 95% CI is 0.96 (0.95–0.97) for endurance time and 0.95 (0.94–0.96) for endurance distance. The mean difference -6.7 ±72.2 s and -7.3 ±113.1m
Hill, K. (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, Z.J.(2011)	Patients with COPD (moderate) (n=53)	Reliability (Test 1 versus test 2 (pre-exercise training), test 3 versus 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))
Butcher, G. (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 limits of agreement (95%) were wide with a small number of individuals showing a larger variation in response. For the eleven patients that repeated the oxygen walk, the mean increase was still significant compared to the performance on air.

Revill, S.M. (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 shows limits of agreement −88 to +112 seconds.
Revill, S.M. (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 limits of agreement (2SD) between tests 2 and 3 was +15 (42) s (p>0.05).

Abbreviation: AUC: Area under Curve; COPD: chronic obstructive pulmonary disease; ISWT: Incremental Shuttle Walk Test; ESWT: Endurance Shuttle Walk Test; GOLD: Global Initiative for Chronic Obstructive Lung Disease; 6MWT: 6-min walk test; ICC: intra class co-efficient

Responsiveness

Eight studies described the responsiveness of ESWT in patients with COPD summarized in Table 3. Two studies,[1,2] reported the responsiveness to PR with ES moderate and large respectively. Two studies[23], [24] showed large SRM (1.46) sensitivity to change,[24] and 0.93,[23] whereas the remaining studies showed a moderate sensitivity index.[25–27]

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

First author (year of publication)	Population	Measurement Property	Main Result

Zatloukal, J. (2019)	Patients with Stable COPD (n=531)	6 weeks of PR and mean change of ESWT was 342.0 (95% CI 312.4–371.6) seconds.	MCID of ESWT in COPD after 6-week PR was between 174 and 279 seconds. By the distribution method (0.5 SD) MCID of 173.7 seconds, the global rating of change scale 279.2 (95% CI 244.9–313.5) seconds and the ROC method 207 seconds. The mean change was 341.6 (347.3) s with 2.87(ES) and 0.98(SRM).
Altenburg, W.A. (2015)	Patients with COPD (GOLD stage IV) (n=55)	6 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. The distribution-based method the MCID was 144 s, 61% and 137 m. The mean change was 121(290) s and 153(274) m. The effect size was 0.18 and SRM 0.41.
Borel, B. (2014)	Patients with Stable COPD (n=255)	8 weeks of study to bronchodilation with two ESWT (baseline) with tiotropium (1 week apart), one after a single dose and one after 4 weeks of either fluticasone propionate/salmeterol combination or placebo in addition to tiotropium.	MCID values ranging 56 to 61 s and 70 to 82 m in endurance time and endurance distance, respectively.
Pepin, V. (2011)	Patients with COPD(n=201)	PR (n=132 for 7weeks 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) is 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(374.5) s with ES 2.2. The mean change in the bronchodilation group was 90.62(120) s with SRM 0.75.
Leung, R.W.M (2010)	Patients with COPD (GOLD stage I–IV) (n=32)	Responsiveness to walking (PR) (eight weeks)	The endurance walking time of walking training group was (mean change 439 ±346 s) (95% CI 70 to 483) more than the cycle training group (mean change, 160 ± 204s). The ES and SRM for the walking group was 2.23 and 1.27(SRM).
Butcher, G. (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 (47, 49) m and between endurance time was 3.6 (63, 56) s. 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 2 and 3, 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, B.(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, S.J. (2008)	Patients with COPD (Severe hypoxemic) (n=41)	Responsiveness to ambulatory oxygen 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, V. (2007)	Patients with COPD (GOLD stage II–III) (n=14)	Responsiveness to bronchodilation	There was 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 0.66
Eaton, T. (2006)	Patients with COPD (n=20)	Responsiveness to PR (seven weeks)	There was significant improvement in ESWT of 92% (302m, 95%C1 104, 501) in ESWT distance. The ES was 0.54. (SRM=0.78)
Pepin, V. (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. 0.93(SRM)
Revill, S.M. (1999)	Patients with COPD (moderate to severe) (n=44) (21 participated in 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; AUC: Area under Curve; COPD: chronic obstructive pulmonary disease; ISWT: Incremental Shuttle Walk Test; ESWT: Endurance Shuttle Walk Test; GOLD: Global Initiative for Chronic Obstructive Lung Disease; 6MWT: 6-min walk test; MCID: minimal clinical important difference; ICC: intraclass correlation co-efficient

The four studies,[38–40,43] assessed the interpretability of the ESWT in 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 56 to 61s and 70 to 82 m in response to BD. While remaining three showed response to PR and BD ranging MCID values (144 s, 61%and 137 m),[39] 45-85 s (60-115 m, 95%CI),[38] and (174 -279 s, 95%CI).[40]

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 responsiveness.[29] The articles were individually and independently judged, further resolving the inconsistencies through discussion. The COSMIN checklist consists of set of questions for all the measurement property. 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 COSIMN recommendations in Table 4. In reliability studies ICC $\geq$ 0.70 is sufficient, ICC < 0.70 (insufficient) and indeterminate if ICC not reported. In studies of responsiveness, the result being in accordance with the hypothesis or the area under the curve (AUC) $\geq$ 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 is shown in Table 5.The data was pooled on the basis of the MD obtained.

Table 4: Methodological quality of the included studies.

First author (year of publication)	Reliability (quality)	Rating	Responsiveness (quality)	Rating
Zatloukal, J. (2019)			very good	?
McKeough, Z.(2018)	inadequate	?		
Ngai, S.P.C. (2017)	inadequate	?		
Altenburg, W.A. (2015)			adequate	?
Borel, B. (2014)	very good	+	very good	?
Hill, K. (2012)	inadequate	?		
McKeough, Z.J.(2011)	inadequate	?		
Pepin, V. (2011)			adequate	+
Leung, R.W.M (2010)			adequate	+
Butcher, G. (2010)	inadequate	?	doubtful	+
Revill, S.M. (2009)	inadequate	?		
Brouillard, B.(2008)			adequate	+
Sandland, S.J. (2008)			doubtful	?
Pepin, V. (2007)			inadequate	+
Eaton, T. (2006)			adequate	+
Pepin, V. (2005)			doubtful	+
Revill, S.M. (1999)	doubtful	?	adequate	+

Abbreviations: (+) 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 to 430.75s ; $p < 0.00001$) Sample size- 756 **BD** – Mean difference(distance)= 168.62 m (95% CI 117.03 to 220.21 m; $p < 0.00001$) Sample size- 103 **AO** – Mean difference(time)= 129.04s (95% CI 47.98 to 210.09 s; $p = 0.002$) Sample size- 64 **AO** – Mean difference(distance)= 80.71 m (95% CI 38.66 to 122.76 m; $p = 0.0002$)Sample size- 64	+	Moderate

CI: class interval; PR: Pulmonary Rehabilitation,; BD: Bronchodilation; AO: Ambulatory Oxygen; (+) sufficient.

Quantitative Assessment

The whole measurement property was 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, A.M) assessed the article and resolved doubts through discussion.

Meta-analysis of responsiveness

The MD (time and distance) was significant for PR ($p<0.00001$), BD ($p<0.00001$) and (AO), ESWT time ($p=0.002$); ESWT distance ($p=0.0002$). The results indicated high heterogeneity ($I^2=96\%$ [PR], $I^2 =75\%$ [AO]). The heterogeneity in the studies could be due to the clinical factors such as gender, age and disease severity. It can also be due to less number of studies and quality of the selected study.

Responsiveness to PR- The five studies reporting the mean change in the ESWT (time) to PR was included in the meta-analysis. The effect size (ES) ranging from 0.18 to 2.9. Figure 2 shows the MD in ESWT time following PR (mean 303.19 s, 756 participants, 95% CI 175.63 to 430.75 s; $p< 0.00001$) in COPD patients.

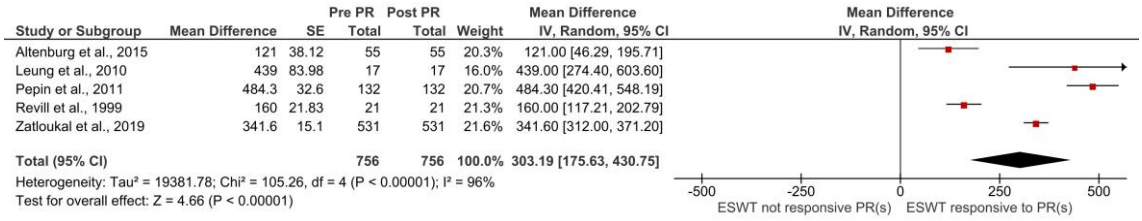

Figure 2: Responsiveness of ESWT following pulmonary rehabilitation.

Responsiveness to BD-Three studies including 103 participants reporting the mean change in the ESWT (distance) to BD was included in the meta-analysis with SRM ranging from 0.56 to 0.93. The MD in ESWT distance following BD was 168.62 m (95% CI 117.03 to 220.21m; p<0.00001) in COPD patients (Figure 3).

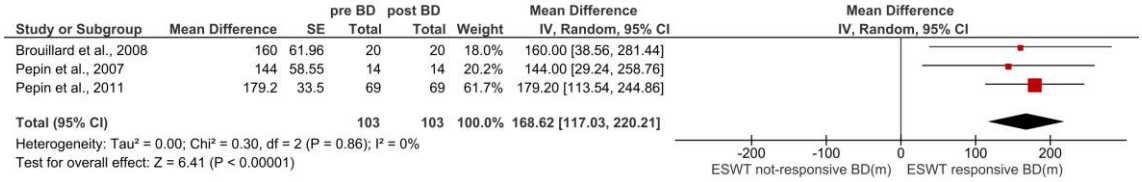

Figure 3: Responsiveness of ESWT following bronchodilator.

Responsiveness to AO-Two studies reporting the mean change in the ESWT (time) to AO was included in the meta-analysis with ES of 0.69 and 1.08. The MD in ESWT time (mean 129.04s, 64 participants, 95% CI 47.98 to 210.09 s; p=0.002); (Figure 4a) and ESWT distance (mean 80.71 m, 64 participants, 95% CI 38.66 to 122.76 m; p=0.0002) (Figure 4b) following AO.

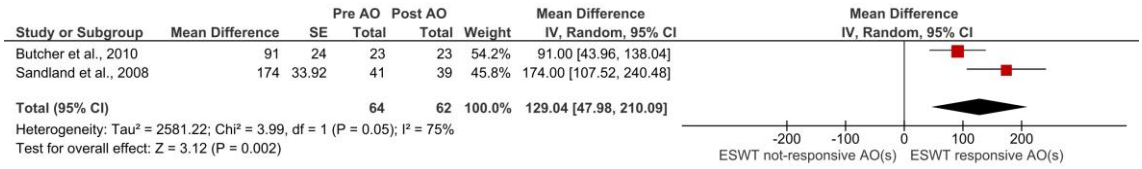

Figure 4a: Responsiveness of ESWT following ambulatory oxygen (seconds)

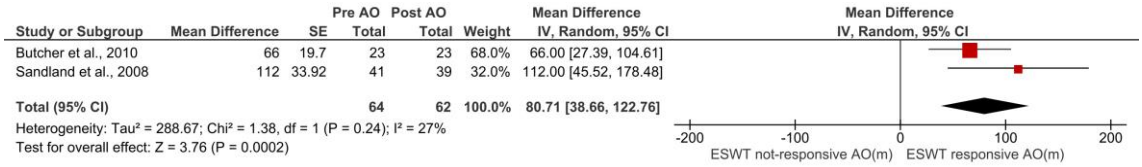

Figure 4b: Responsiveness of ESWT following ambulatory oxygen (minutes)

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 high heterogeneity obtained in the included studies.

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. The Pearson's correlations also confirmed the reproducibility of ESWT performance following 8 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. According to Fotheringham et al.[45] ESWT has a good repeatability and reproducibility with prior familiarization. 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 (EID) showed small difference between the first and second tests.[41]There was a statistically significant

difference with the end-test dyspnea score for the ISWT and ESWT with 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 to 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] Revillet al,[1] and McKeough et al,[22] reported insignificant changes in test-retest reliability to ESWT. The measurements of oxygen saturation (SpO2), HR, and modified Borg dyspnoea scale repeated well during the test.[6,22] One investigation[22] 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 reported 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 the 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 as compare to 6MWT and endurance cycle ergometer tests (ECET).[23,25] The improvements observed in the performances of two studies assessing ESWT was exceeding MCID, following BD therapy.[23,25]

The supplemental oxygen might have an impact on 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 as compared with air ranged from 70s-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 m). 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 the oxygen shall not be inferred in all the patients with COPD. Three studies[27,37,50] reported the potential use of SpO_2 in recording exertional desaturation during ESWT in either AO assessment or breathing air. As compared with 6MWT a study[37] reported greater desaturation with the ESWT in COPD. MCID scores reflects changes in clinical intervention which are meaningful for the patient.[53] The MCID was estimated either with anchor (45-85s, 186-199s, 174-279s, 56-61s),[38–40,43] or distribution-based method (203m, 137m, 173.7s, 81m).[38–40,43] The distribution-based considers measurement error but lacks clinical explanation in different samples. The anchor-based approach offers clinical significance of MCID, through the external changes to an anchor, but it do not consider measurement error.[54] There were no studies which 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, 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 COSMIN checklist allows sufficient flexibility in quality interpretation making the results of study's quality in absolute.[44] The meta-analysis of ESWT responsiveness was significant following PR, BD and AO in patients with COPD.

STRENGTH AND LIMITATIONS

This is the first study to systematically review and meta-analyse 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's, 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's 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 the 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. Majority 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 setting and in research due to the high heterogeneity obtained in the included studies. ESWT lacks enough studies reporting ICC to reach definitive conclusion as a reliable tool in patients with COPD. Further research examining the reliability of ESWT is required reporting ICC values.

REFERENCES

1. Revill SM, Morgan MDL, 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. doi:10.1136/thx.54.3.213

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;3:3-9.

3. Puente-Maestu L, Villar F, De Miguel J, et al. Clinical relevance of constant power exercise duration changes in COPD. Eur Respir J. 2008;34(2):340-345. doi:10.1183/09031936.00078308

4. Singh SJ, Morgan MDL, 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. doi:10.1136/thx.47.12.1019

5. Singh SJ, Puhan MA, Andrianopoulos V, 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;44(6):1447-1478. doi:10.1183/09031936.00150414

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;17(2):278-284. doi:10.1111/j.1440-1843.2011.02089.x

7. Incidence I, Collaborators P. 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. Published online 2017:1990-2015. doi:10.1016/S0140-6736(16)31678-6

8. Collaborators D. Global , regional , and national life expectancy , all-cause mortality , and cause-specifi c mortality for 249 causes of death , 1980 – 2015 : a systematic analysis for the Global Burden of Disease Study 2015. Published online 2016:1980-2015. doi:10.1016/S0140-6736(16)31012-1

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;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. Rev Port Pneumol. 2001;7(4-5):398-400. doi:10.1016/s0873-2159(15)30846-1

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;35(3):1-8. doi:10.5001/omj.2020.54

12. Corhay JL, Dang DN, Van Cauwenberge H, Louis R. Pulmonary rehabilitation and copd: Providing patients a good environment for optimizing therapy. Int J COPD. 2013;9:27-39. doi:10.2147/COPD.S52012

13. Rabinovich R, Gonzalez-desuso JM, Rodríguez D. Clinical Assessment of Peripheral. 2009;(January):39-46. doi:10.1097/PHM.0b013e31818dff86

14. Arne M, Janson C, Janson S, 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. doi:10.1080/02813430902808643

15. Viegi G, Pistelli F, Sherrill DL, Maio S, Baldacci S, Carrozzi L. Definition, epidemiology and natural history of COPD. Eur Respir J. 2007;30(5):993-1013. doi:10.1183/09031936.00082507

16. Moloney E, Clayton N, Mukherjee DK, Gallagher CG, Egan JJ. The shuttle walk exercise test in idiopathic pulmonary fibrosis. Respir Med. 2003;97(6):682-687. doi:10.1053/rmed.2003.1501

17. Ushiki A, Nozawa S, Yasuo 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;55(1):33-38. doi:10.1016/j.resinv.2016.08.004

18. Ishrat, R., Mujaddadi, A., Ali, M. S., Talwar, D., & Hussain, M. E. Reliability and responsiveness of incremental shuttle walk test to estimate exercise capacity in patients with pulmonary arterial hypertension. Comp Exer Physio.2020;16(3):179-185.

19. Holland AE, Spruit MA, Troosters T, et al. An official European Respiratory Society / American Thoracic Society technical standard : field walking tests in chronic respiratory disease. Published online 2014:1428-1446. doi:10.1183/09031936.00150314

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;145(6):1357-1369. doi:10.1378/chest.13-2071

21. Revill SM, Williams J, Sewell L, Collier R, Singh SJ. Within-day repeatability of the endurance shuttle walk test. 2009;95:140-143. doi:10.1016/j.physio.2009.02.001

22. McKeough ZJ, Leung RWM, 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. doi:10.1097/PHM.0b013e3181fc7f42

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;172(12):1517-1522. doi:10.1164/rccm.200507-1037OC

24. Leung RWM, 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. doi:10.1016/S1836-9553(10)70040-0

25. 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. doi:10.1136/thx.2006.065540

26. Brouillard C, Pepin V, Milot J, Lacasse Y, Maltais F. responsiveness to salmeterol in COPD. 2008;31(3):579-584. doi:10.1183/09031936.00119007

27. Sandland CJ, Morgan MDL, Singh SJ. Detecting oxygen desaturation in patients with COPD : Incremental versus endurance shuttle walking. Published online 2008:1148-1152. doi:10.1016/j.rmed.2008.03.007

28. Kamioka H. Preferred reporting items for systematic review and meta-analysis protocols (prisma-p) 2015 statement. Japanese Pharmacol Ther. 2019;47(8):1177-1185.

29. Mokkink LB, Prinsen CA, Patrick D, et al. COSMIN Study Design checklist for Patient-reported outcome measurement instruments. Dep Epidemiol Biostat Amsterdam Public Heal Res Inst Amsterdam Univ Med Centers, Locat VUmc. 2019;(July):1-32. www.cosmin.nl

30. Terwee CB, Mokkink LB, Knol DL, Ostelo RWJG, Bouter LM, De Vet HCW. Rating the methodological quality in systematic reviews of studies on measurement properties: A scoring system for the COSMIN checklist. Qual Life Res. 2012;21(4):651-657. doi:10.1007/s11136-011-9960-1

31. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. Stat Med. 1998;17(1):101-110. doi:10.1002/(SICI)1097-0258(19980115)17:1<101::AID-SIM727>3.0.CO;2-E

32. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: A critical review and recommendations. J Clin Epidemiol. 2000;53(5):459-468. doi:10.1016/S0895-4356(99)00206-1

33. Mokkink LB, Prinsen CAC, Patrick DL, et al. COSMIN methodology for systematic reviews of Patient - Reported Outcome Measures (PROMs). User Manual. 2018;(February):1-78.

34. Editors S, Thomson CC, Wilson KC, et al. ATS CLINICAL PRACTICE GUIDELINE : SUMMARY FOR CLINICIANS. 2015;12(3):446-447. doi:10.1513/AnnalsATS.201501-057OT

35. Zlowodzki M, Poolman RW, Kerkhoffs GM, Tornetta P, Bhandari M. How to interpret a meta-analysis and judge its value as a guide for clinical practice. Acta Orthop. 2007;78(5):598-609. doi:10.1080/17453670710014284

36. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: An emerging consensus on rating

quality of evidence and strength of recommendations. Chinese J Evidence-Based Med. 2009;9(1):8-11.

37. Butcher G. 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 doi:10.1177/1479972310378311

38. Pepin V, Laviolette L, Brouillard C, et al. Significance of changes in endurance shuttle walking performance. Thorax. 2011;66(2):115-120. doi:10.1136/thx.2010.146159

39. Altenburg WA, Duiverman ML, ten Hacken NHT, 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;16(1):1-7. doi:10.1186/s12931-015-0182-x

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. doi:10.1177/1479973119853828

41. McKeough Z, Leung R, Neo JH, 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;15(2):131-137. doi:10.1177/1479972317729051

42. Ngai SPC, Spencer LM, Jones AYM, Alison JA. Repeatability of the endurance shuttle walk test in people with chronic obstructive pulmonary disease. Clin Respir J. 2017;11(6):875-880. doi:10.1111/crj.12430

43. Borel B, Pepin V, Mahler DA, Nadreau É, Maltais F. Prospective validation of the endurance shuttle walking test in the context of bronchodilation in COPD. Eur Respir J. 2014;44(5):1166-1176. doi:10.1183/09031936.00024314

44. Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res. 2018;27(5):1147-1157. doi:10.1007/s11136-018-1798-3

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 COPD. 2015;10:625-643. doi:10.2147/COPD.S70518

46. Issues S, Test MW, Equipment R, Preparation P. American Thoracic Society ATS Statement : Guidelines for the Six-Minute Walk Test. 2002;166:111-117. doi:10.1164/rccm.166/1/111

47. Greening NJ, Evans RA, Williams JEA, 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. doi:10.1177/1479972312439317

48. O'Farrell R, Gargoum F, O'Connor B, et al. Comparison Of Incremental And Endurance Shuttle Walk Testing After Pulmonary Rehabilitation. Published online 2011:A2024-A2024. doi:10.1164/ajrccm-conference.2011.183.1_meetingabstracts.a2024

49. Borel B, Provencher S, Saey D, Maltais F. Responsiveness of Various Exercise-Testing Protocols to Therapeutic Interventions in COPD. 2013;2013(Mid).

50. Organ MDLM. Randomized controlled trial of ambulatory oxygen and an ambulatory ventilator on endurance exercise in COPD. Published online 2000:778-783. doi:10.1053/rmed.2000.0824

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;174(7):763-771. doi:10.1164/rccm.200509-1533OC

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;9(2):83-91. doi:10.1177/1479972312438702

53. Cook CE. Clinimetrics Corner: The Minimal Clinically Important Change Score (MCID): A Necessary Pretense. J Man Manip Ther. 2008;16(4):82E-83E. doi:10.1179/jmt.2008.16.4.82e

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;17(1):1-8. doi:10.1186/s13075-015-0658-6