



Prospective validation of the endurance shuttle walking test in the context of bronchodilation in COPD

Benoit Borel^{1,2}, Véronique Pepin³, Donald A. Mahler⁴, Éric Nadreau¹
and François Maltais¹

Affiliations: ¹Centre de recherche, Institut Universitaire de Cardiologie et de Pneumologie de Québec – Université Laval, Québec, Canada. ²Laboratoire HAVAE (EA6310), Université de Limoges, Limoges, France. ³Centre de recherche, Hôpital du Sacré-Coeur de Montréal, Montreal, Canada. ⁴Geisel School of Medicine at Dartmouth, Hanover, NH, USA.

Correspondence: François Maltais, Centre de Pneumologie, Institut Universitaire de Cardiologie et de Pneumologie de Québec – Université Laval, 2725 Chemin Sainte-Foy, Québec, G1V 4G5, Canada.
E-mail: Francois.Maltais@fmed.ulaval.ca

ABSTRACT This study focused on repeatability data and minimal important difference (MID) estimates of the endurance shuttle walking test (ESWT).

255 chronic obstructive pulmonary disease patients (forced expiratory volume in 1 s $54.7 \pm 13.2\%$ predicted) completed four ESWTs at different times during the 8-week study: two under baseline conditions 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. 97 patients performed all the tests with a portable metabolic system. Reproducibility of test performance and cardiorespiratory response was investigated with the data obtained on the first two ESWTs.

The mean differences between the first two ESWT performances (-6.7 ± 72.2 s and -7.3 ± 113.1 m for endurance time and walking distance, respectively) were not statistically significant. The between-test end-exercise and isotime values for each cardiorespiratory parameter were not significantly different from each other. With the exception of arterial oxygen saturation by pulse oximetry, the repeatability of cardiorespiratory adaptations to ESWT was also confirmed with strong Pearson and intraclass correlation coefficients. Finally, changes of 56–61 s and 70–82 m in endurance time and distance walked, respectively, were perceived by patients.

This study provides methodological information supporting the reliability of the ESWT and suggests MID estimates for this test.



@ERSpublications

This study provides a prospective validation of ESWT reliability and minimal important difference in COPD <http://ow.ly/zuLew>

Received: Feb 05 2014 | Accepted after revision: June 26 2014

Clinical trial: This study is registered at clinicaltrials.gov (identification number NCT01124422).

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

Copyright ©ERS 2014

Introduction

Exercise intolerance is a hallmark feature of chronic obstructive pulmonary disease (COPD), and leads to a sedentary lifestyle, poor quality of life and even reduced survival [1, 2]. As such, measuring the endurance to constant load exercise has been an important objective of several clinical trials evaluating the impact of pharmacological or rehabilitative interventions in COPD [3–5]. In this context, the endurance shuttle walking test (ESWT) is gaining acceptance due to its minimal requirement for equipment, low cost and ease of administration [6], and to its responsiveness and sensitivity to interventions, at least in single centre studies [7–12]. With the advancement of portable cardiopulmonary monitoring, it is also possible to obtain a thorough physiological assessment during the ESWT. However, some methodological aspects of the ESWT deserve further evaluation. The reproducibility of the ESWT time has been reported in only one study [12], involving a limited sample size and without measurement of cardiopulmonary variables during the test. Another issue is the interpretation of changes in ESWT performance after an intervention. In a prior study from our group, estimates of the minimal important difference (MID) from the patients' perspective were provided for changes in ESWT performance [13]. In this prior study, it was determined that a change in ESWT performance of 45–85 s or 60–115 m after bronchodilation could be perceived by patients ($n=69$). Prospective validation of these estimates in a larger number of patients seems warranted given the widespread use of MID values in clinical and administrative decision-making processes.

Therefore, the aims of this prospective study were two-fold: 1) to evaluate the reproducibility of the ESWT performance and of the cardiorespiratory variables during the test; and 2) to prospectively validate MID estimates of ESWT performance in patients with COPD.

Subjects and methods

Ethics statement

The study was conducted in 24 sites, in the USA and Canada, between July 2010 and May 2011. The study was approved by all the local ethics review committees and was conducted in accordance with the Declaration of Helsinki and Good Clinical Practice guidelines.

Subjects

Patients with stable COPD agreed to participate and signed a written informed consent form. Inclusion criteria were as follow: age ≥ 40 years; a clinical history of COPD; post-bronchodilator forced expiratory volume in 1 s (FEV₁) 30–80% predicted; FEV₁/forced vital capacity ratio ≤ 0.70 ; and a current or past history of ≥ 10 pack-years of cigarette smoking. Exclusion criteria included: diagnosis of asthma; an exacerbation requiring antibiotics and/or systemic corticosteroids within the 6 weeks preceding inclusion into the study; a body mass index ≥ 40 kg·m⁻²; and any clinically significant abnormal ECG.

Study design

This study represents a pre-specified secondary analysis of an 8-week multicentre, placebo-controlled, double blind, randomised, parallel-group study (clinicaltrials.gov identification number NCT01124422) investigating whether the addition of fluticasone/salmeterol fixed-dose combination (FSC) to open label tiotropium further improved walking capacity in comparison to tiotropium alone [5]. Figure 1 provides a schematic representation of study design.

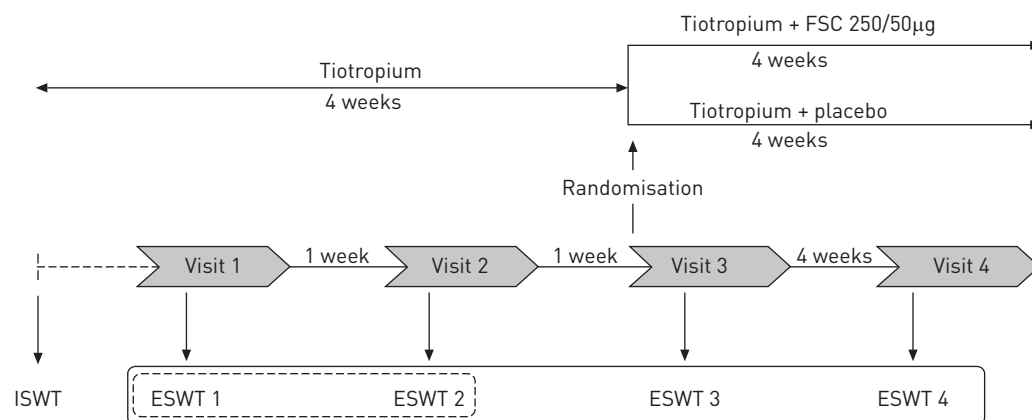


FIGURE 1 Study timeline. FSC: fluticasone/salmeterol fixed-dose combination; ISWT: incremental shuttle walking test; ESWT: endurance shuttle walking test.

Study participants completed a 4-week run-in period of open-label tiotropium (Spiriva; Boehringer Ingelheim GmbH, Ingelheim, Germany). The incremental shuttle walking test (ISWT) was first completed without prior practice. After ISWT completion, patients were familiarised with the ESWT procedures and were asked to walk on the same course for a few minutes at constant speed. After at least 2 weeks of tiotropium therapy, two ESWTs (ESWT 1 and ESWT 2) were performed under baseline condition. A third and fourth ESWT (ESWT 3 and ESWT 4) were performed after the first dose and after 4 weeks of study medications, respectively (FSC 250/50 µg administered *via* Advair/Seretide DISKUS twice daily (GlaxoSmithKline, Research Triangle Park, NC, USA) or placebo DISKUS twice daily). These additional treatments were administered under a background of open-label tiotropium. As per protocol design, patients with exercise duration >20 min and those in whom ESWT 1 and ESWT 2 exercise duration was >2 min apart were excluded from further study procedures in the original clinical trial [5].

Endurance shuttle walking test

All subjects initially performed an ISWT, as originally described by SINGH *et al.* [14], in order to determine their own maximal walking speed and distance. From the data obtained with the ISWT, peak oxygen uptake ($\dot{V}O_{2\text{peak}}$) was estimated and the walking speed corresponding to 85% of estimated $\dot{V}O_{2\text{peak}}$ was determined in order to define the walking speed for the ESWT, according to the procedure described by REVILL *et al.* [12]. The ESWTs were performed on the same course as the ISWT (10 m course in a flat corridor), in accordance with published guidelines [12]. After a 90-s warm-up period, walking speed increased at the speed corresponding to 85% of estimated $\dot{V}O_{2\text{peak}}$ and subjects were instructed to walk as long as they could at the pace dictated by the auditory signal. The endurance time was defined as the duration of the ESWT and the distance covered during the test was recorded, excluding the warm-up period. With the objective of standardising exercise testing procedures across participating centres, investigators and research personnel had the opportunity to familiarise themselves with the ESWT by performing a few practice tests under the supervision of an exercise physiologist during the investigator meeting. Research teams were also provided with an instruction manual and the required materials covering the methodological aspects of the ESWT.

Subjective rating of change

After each ESWT and before any feedback was given to the subject about the test result, subjects rated their performance of the day in comparison to the previous one using a 7-point Likert scale. The scale ranged from -3 to +3 and included the following ratings: -3 (large deterioration); -2 (moderate deterioration); -1 (slight deterioration); 0 (no change); +1 (slight improvement); +2 (moderate improvement); and +3 (large improvement). Subjects were asked to answer the following question: "In comparison with your previous test, how would you rate your performance on today's test using the present scale?"

Physiological measurements

In a subset of participating centres (eight out of 24) and subjects ($n=97$, 35% of the total study population), physiological measurements were collected during the ESWT, using a portable gas exchange analyser (Oxycon Mobil; Viasys Healthcare, Care Fusion, Germany). Subjects wore a facemask, which allowed measurements of $\dot{V}O_2$, carbon dioxide output ($\dot{V}CO_2$), minute ventilation ($\dot{V}E$) and respiratory exchange ratio (RER) on a breath-by-breath basis. The device was calibrated before each test according to manufacturer's recommended procedures. Arterial oxygen saturation measured by pulse oximetry (SpO_2) was monitored during all tests using a finger oximeter. A 12-lead electrocardiogram (Cardiosoft; Corina, Milwaukee, WI, USA) was used to monitor cardiac activity and heart rate throughout each test.

Statistical analysis

The analyses for this study were pre-specified at the beginning of the study. All variables are expressed as mean \pm SD, unless otherwise stated. For all analyses, the level of significance was set at $p=0.05$. The reproducibility of the ESWT performance (time and distance) was evaluated by comparing values obtained at ESWT 1 and 2 using a paired t-test. The presence of a systematic bias and the agreement between these two ESWTs was assessed using a Bland–Altman plot [15], and the relationship between exercise performance at ESWTs was evaluated by the Pearson product moment correlation. The analysis of ESWT performance reproducibility was completed with the determination of intraclass correlation coefficient (ICC), a valuable way to determine the degree of association and agreement between variables [16]. End-exercise and submaximal $\dot{V}O_2$, $\dot{V}CO_2$, $\dot{V}E$, heart rate, SpO_2 and RER values were compared using paired t-tests. Time courses of each parameter during ESWTs were compared using two-way repeated measure ANOVA, with "visit" and "isotime" as factors. Isotime was defined as exercise time-points corresponding to 25%, 50%, 75% and 100% of the latest exercise time reached on both ESWTs. When the ANOVA F-ratio was significant, *post hoc* analysis was then conducted with a Tukey test to perform pairwise multiple

comparison procedures. Test–retest reproducibility of cardiorespiratory adaptations to ESWT was also assessed using the ICC and within-subject coefficient of variation.

To provide MID estimates for ESWT time and distance, data from the four ESWTs were used. MID values were estimated using both distribution- and anchor-based approaches, similarly to our previous investigation [13]. For the distribution-based approach, half the SD of the change in ESWT performance provided the MID estimates [17]. In the anchor-based approach, differences in performance between two ESWTs were calculated for consecutive tests and for each patient. Relationships between objective and subjective changes in ESWT performance were established using Pearson product moment correlations. The coefficient correlations between the anchor and target needed to be ≥ 0.3 for the regression to be valid [18]. If this condition was fulfilled, the association between the subjective rating of change and the corresponding difference in performance was plotted. The resulting slopes, which represented the change needed in walking time or distance for participants' ratings to move one unit on the Likert perception scale, were considered as estimates of the MID.

All the data were analysed using SPSS (version 21.0; SPSS Inc., Chicago, IL, USA) and GraphPad Prism 6 (GraphPad Software Inc., La Jolla, CA, USA).

Results

Anthropometric and pulmonary data

One patient had an ESWT duration >20 min and was excluded from further study procedures. Because of this, it was not possible to obtain reproducibility and MID data in this individual. Nine patients were excluded from the treatment phase of the original clinical trial because their two baseline ESWTs did not meet the *a priori* defined reproducibility criteria for the trial (endurance time <2 min apart). 12 other patients performed both baseline ESWTs but were not included in the treatment phase due to other reasons. These 21 patients contributed to the ESWT reproducibility analysis but not to the MID analysis. Lastly, 255 patients were randomised into the treatment phase of the study and could contribute to both aspects of the present report.

As shown in table 1, the characteristics of the entire population ($n=276$) and the subgroup of subjects in whom physiological measurements were obtained during the ESWTs ($n=97$) were similar. On average, subjects exhibited moderate-to-severe airflow obstruction.

Reproducibility of ESWT performance

Mean values for ESWT time and distance obtained at ESWT 1 and ESWT 2 are reported in table 2. These ESWT performance parameters were not statistically different between the two tests.

The reproducibility of ESWT performance was also confirmed with Pearson correlations and the Bland–Altman plot (fig. 2). Strong linear relationships were found for endurance times and distances obtained at ESWT 1 and ESWT 2, with correlation coefficients of 0.92 and 0.90, respectively. The graphical interpretation of the Bland–Altman plot further confirms the reproducibility of ESWT performance, with a mean bias (95% CI) of -6.7 s (-148.3 – 134.8 s) and -7.3 m (-228.9 – 214.3 m) for endurance time and distance, respectively (fig. 2). These small biases represent less than one shuttle at the fastest walking cadence. The reproducibility of ESWT performance was also assessed using ICC, with an ICC (95% CI) of 0.96 (0.95–0.97) and 0.95 (0.94–0.96) for endurance time and endurance distance, respectively.

Reproducibility of metabolic and cardiorespiratory parameters measured during ESWT

Mean end-exercise values for metabolic and cardiorespiratory parameters were not significantly different between ESWT 1 and ESWT 2 (table 3).

Apart from SpO_2 , end-exercise values reached during both ESWTs showed strong Pearson correlations coefficients (table 3). The time course of each physiological parameter during ESWT 1 and ESWT 2 was practically identical (fig. 3).

Similarly, ANOVA results show no significant interaction of “visit” and “isotime” factors for the physiological parameters, which means there was no significant isotime difference for these parameters between both ESWTs. ICC and coefficients of variation are provided in tables 4 and 5. With the exception of SpO_2 , the reproducibility of all the other cardiorespiratory parameters was excellent at different percentages of isotime or at the maximal level.

MID values determination

Using the distribution-based approach, our data suggest MID values of 61 s or 82 m. The correlation coefficients between the anchor (the subjective rating of change) and the changes in ESWT performance were 0.48 and 0.45 for the ESWT time and distance, respectively (fig. 4). The slopes of these relationships

TABLE 1 Baseline characteristics of the entire study group and the subgroup for which physiological measurements were obtained

	All subjects	Physiological subgroup
Subjects	276	97
Patients characteristics		
Males/females	152/124	58/39
Age years	62.9 ± 9.4	61.6 ± 7.9
BMI kg·m ⁻²	27.5 ± 5.6	27.5 ± 5.8
Smoking history pack-years	56 ± 28	58 ± 26
COPD duration years	7.0 ± 6.4	6.6 ± 6.5
Pulmonary function post-bronchodilation		
FEV ₁ L	1.6 ± 0.5	1.6 ± 0.5
FEV ₁ % predicted	54.9 ± 13.1	53.7 ± 12.7
FVC L	3.1 ± 0.9	3.2 ± 1.0
FVC % predicted	89.0 ± 18.1	89.7 ± 20.4
FEV ₁ /FVC %	53 ± 11	52 ± 11
Pulmonary function pre-bronchodilation		
FRC % predicted	129.7 ± 30.2	131.4 ± 34.0
TLC % predicted	105.8 ± 17.0	106.2 ± 19.3
RV % predicted	149.4 ± 40.8	152.4 ± 43.4
IC L	2.30 ± 0.64	2.27 ± 0.71
DLC₀ % predicted	74.1 ± 51.6	57.3 ± 16.2
Peak exercise capacity during ISWT		
Maximal distance m	412.8 ± 179.8	369.7 ± 145.0
Maximal velocity km·h ⁻¹	5.6 ± 1.1	5.4 ± 1.0

Data are presented as n or mean ± SD. BMI: body mass index; COPD: chronic obstructive pulmonary disease; FEV₁: forced expiratory volume in 1 s; FVC: forced vital capacity; FRC: functional residual capacity; TLC: total lung capacity; RV: residual volume; IC: inspiratory capacity; DLC₀: diffusing capacity of the lung for carbon monoxide; ISWT: incremental shuttle walking test.

suggested that a change in walking time of 56.3 s (95% CI 48.9–63.6 s) and a change in walking distance of 69.6 m (95% CI 59.5–79.6 m) should be perceived by subjects. These changes represent 15.7% (95% CI 13.2–18.1%) and 14.8% (95% CI 12.6–17.1%) of the baseline value, respectively.

Discussion

The present results indicate high reproducibility of ESWT time and distance in the context of a multicentre clinical trial. In addition, cardiorespiratory parameters obtained during the ESWT were highly reproducible, with the exception of SpO₂. Finally, we found that changes in ESWT time and distance of 56–61 s and 70–82 m, respectively, were likely to be perceived by subjects and could therefore be considered clinically meaningful.

Reproducibility of ESWT performance

The use of the ESWT for the evaluation of patients with COPD has become increasingly popular in research and clinical fields [8, 9, 13]. However, the methodological validation of the ESWT is limited. Before the present investigation, the reproducibility of the ESWT was documented only once, in a sample of 11 subjects in the original development study by REVILL *et al.* [12]. Although this paper concluded on the good reproducibility of ESWT performance, the precision of the limits of agreement remained questionable in light of the small sample size, as highlighted by the authors themselves [12]. Accordingly, our study provides

TABLE 2 Mean endurance shuttle walking test (ESWT) performances and mean performance differences

Variable	ESWT 1	ESWT 2	Mean difference	ICC value (95% CI)
Endurance time s	381 ± 229	374 ± 220	-6.7 ± 72.2	0.96 (0.95–0.97)
Endurance distance m	486 ± 342	479 ± 332	-7.3 ± 113.1	0.95 (0.94–0.96)

Data are presented as mean ± SD, unless otherwise stated. ICC: interclass correlation coefficient.

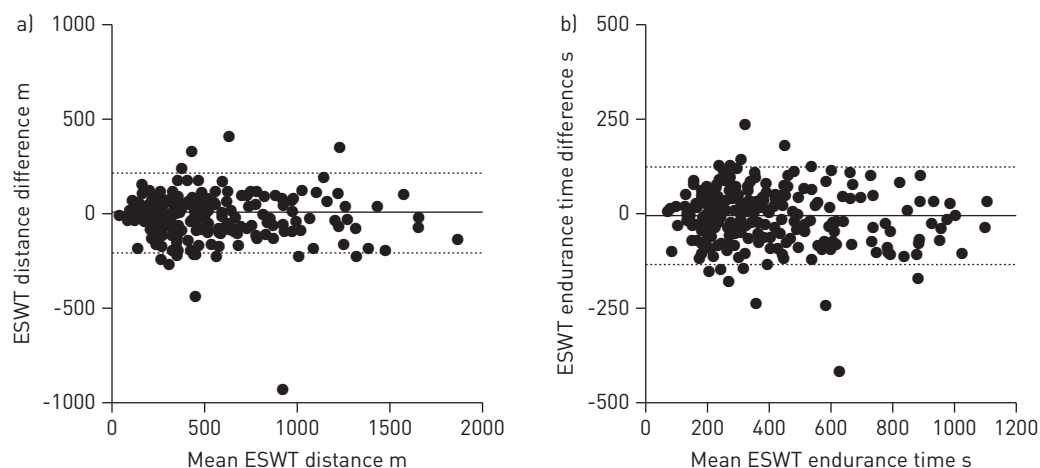


FIGURE 2 Bland–Altman plot for a) endurance shuttle walking test (ESWT) distance and b) ESWT endurance time. —: the mean difference between ESWT 1 and ESWT 2; ·····: 95% confidence intervals of the difference in ESWT performance.

an important contribution by highlighting the absence of a significant difference in mean endurance time and distance between two ESWT performed 1 week apart, under stable conditions. The graphical analyses of the Bland–Altman plot confirm the absence of systematic differences across the test performance spectrum (*i.e.* from low to high performance) and ICC also confirm the strong association between the performances achieved on both ESWTs. These results are in accordance with the conclusions of REVILL *et al.* [12] and thus add weight to these findings. In addition, the reproducibility and limits of agreement of the ESWT over time can now be interpreted with more confidence given the multicentre nature and large sample size of our study and the fact that both tests were conducted 1 week apart.

Nevertheless, an intra-individual variability was observed in ESWT performance, as indicated by the SD of the mean ESWT performance differences between ESWT 1 and ESWT 2 (table 2) and by the limits of agreement of the differences in ESWT performance (fig. 2). The randomness of this variability suggests that it was not related to a learning effect, as confirmed by the Bland–Altman plot. The issue about intra-individual variability in exercise performance was previously discussed for the ESWT [12, 19–21], but also for constant work-rate cycling exercise [22]. REVILL *et al.* [21] reported greater variability in ESWT performance with longer endurance time. One suggested explanation for this observation was the possibility that longer walking times were more sensitive to environmental factors and other influences (motivation, mood, *etc.*) than shorter tests [21]. In our study, examinations of the Bland–Altman plot (fig. 2) did not reveal such influence of walking duration on the between-test variability in ESWT performance. Day-to-day variations in airway function, differences in medication use, subject encouragement, prior pre-test activity or differences in time-of-day of the assessments were some factors proposed to explain performance variability in constant work-rate cycle exercise tests [22]. Finally, there was a small subgroup of patients (21 patients corresponding to 8% of the total sample size) that had not been included in the treatment phase. Among these 21 patients, nine were unable to perform reproducible ESWTs defined as a <2-min difference between the two baseline tests. Although the inclusion of these individuals did not alter our conclusion

TABLE 3 End-exercise values for cardiorespiratory measures at endurance shuttle walking test (ESWT) 1 and 2

Variables	ESWT 1	ESWT 2	Difference	p-value	r-value
$\dot{V}O_2$ mL·min ⁻¹	1143±384	1122±400	-13.8±186	0.49	0.89
$\dot{V}CO_2$ mL·min ⁻¹	1059±394	1037±402	-13.7±194	0.51	0.88
RER	0.92±0.09	0.91±0.10	-0.01±0.07	0.40	0.73
$\dot{V}E$ L·min ⁻¹	38.8±13.9	38.0±13.2	-1.1±11.3	0.36	0.66
Heart rate beats·min ⁻¹	117.4±18.4	116.5±16.8	-0.91±11.4	0.44	0.79
SpO ₂ %	88.1±12.5	89.1±11.2	1.3±13.9	0.36	0.32

Data are presented as mean ± SD, unless otherwise stated. All p-values were nonsignificant. $\dot{V}O_2$: oxygen uptake; $\dot{V}CO_2$: carbon dioxide output; RER: respiratory exchange ratio; $\dot{V}E$: minute ventilation; SpO₂: arterial oxygen saturation measured by pulse oximetry.

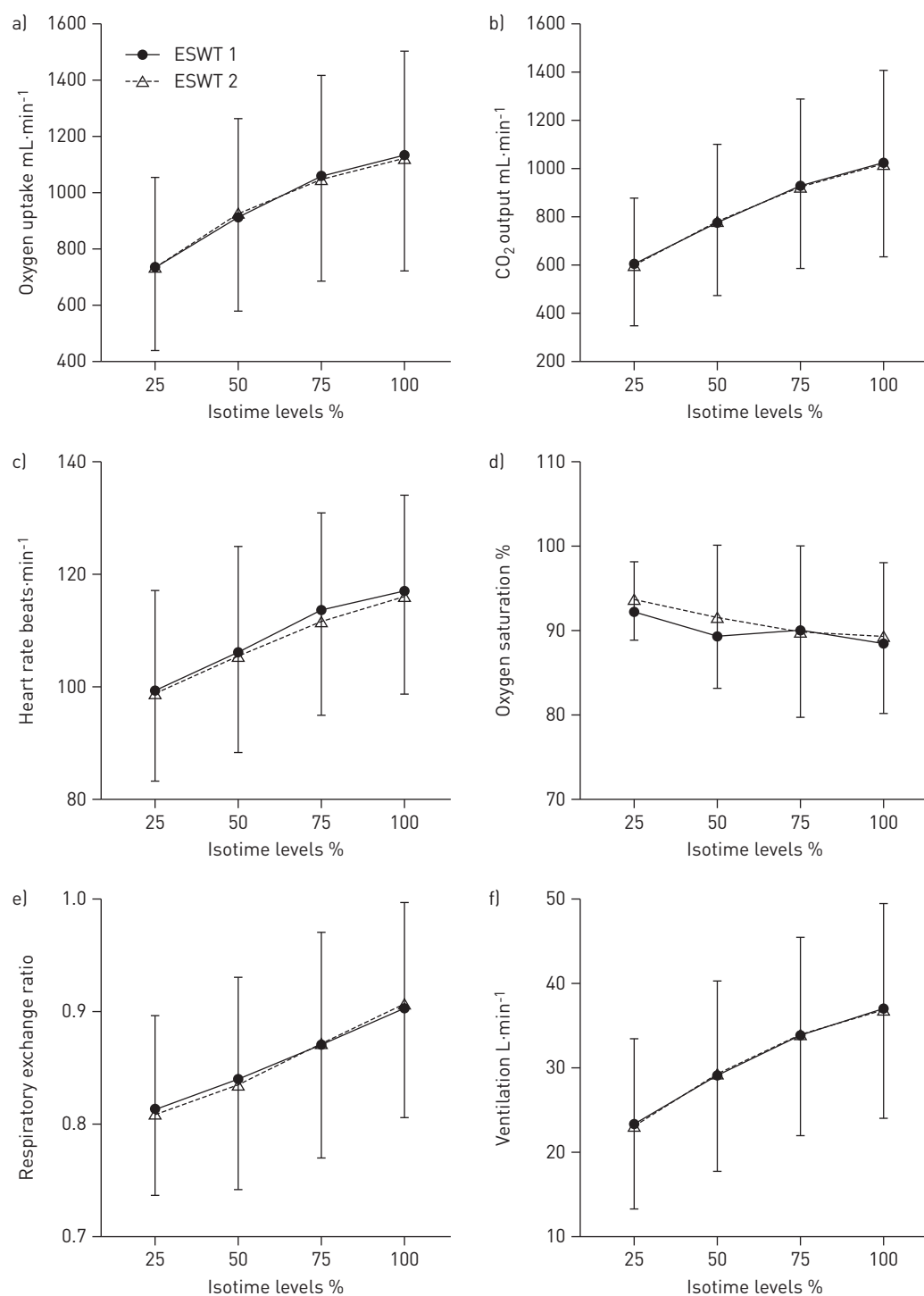


FIGURE 3 Time course of cardiorespiratory parameters at 25%, 50%, 75% and 100% isotime levels. ESWT: endurance shuttle walking test.

regarding the reproducibility of the ESWT, we would recommend excluding them from randomisation in the treatment period of clinical trials aiming at evaluating the impact of therapies on exercise endurance.

Repeatability of cardiorespiratory adaptations to ESWT

Peak values and time course of cardiorespiratory and metabolic parameters were compared between the first two ESWTs performed in 97 out of 255 subjects. No significant difference between mean values was found for any parameter supporting good repeatability of cardiopulmonary physiological variables during ESWT.

TABLE 4 Intraclass correlation coefficients for cardiorespiratory parameters at different isotime levels

	Isotime levels				
	25%	50%	75%	100%	Peak
$\dot{V}O_2$	0.80 [0.71–0.86]	0.88 [0.83–0.92]	0.95 [0.92–0.96]	0.93 [0.89–0.95]	0.89 [0.84–0.93]
$\dot{V}CO_2$	0.79 [0.69–0.86]	0.87 [0.80–0.91]	0.93 [0.90–0.96]	0.94 [0.90–0.96]	0.88 [0.83–0.92]
Heart rate beats·min ⁻¹	0.84 [0.76–0.89]	0.83 [0.76–0.89]	0.88 [0.83–0.92]	0.88 [0.82–0.92]	0.79 [0.71–0.86]
SpO ₂	0.30 [0.10–0.48]	0.23 [0.02–0.41]	0.44 [0.25–0.60]	0.32 [0.10–0.50]	0.32 [0.12–0.49]
RER	0.66 [0.52–0.76]	0.82 [0.74–0.88]	0.88 [0.83–0.92]	0.87 [0.81–0.91]	0.73 [0.62–0.81]
$\dot{V}E$	0.79 [0.70–0.86]	0.84 [0.77–0.89]	0.89 [0.84–0.93]	0.92 [0.88–0.95]	0.66 [0.52–0.76]

Data are presented as intraclass correlation [95% CI]. $\dot{V}O_2$: oxygen uptake; $\dot{V}CO_2$: carbon dioxide output; SpO₂: arterial oxygen saturation measured by pulse oximetry; RER: respiratory exchange ratio; $\dot{V}E$: minute ventilation. 25%, 50%, 75%, 100% correspond to exercise time-points corresponding to 25%, 50%, 75% and 100% of the latest exercise time reached in both endurance shuttle walking tests. Peak corresponds to the highest value reached during the two first endurance shuttle walking tests.

This contention was further supported by Pearson correlation coefficients, ICC and coefficients of variation values. ICC values reflect the within-subject variance, referenced to as the degree of between-subject variance [23, 24]. It is generally assumed that an ICC near 0 reflects an absence of reliability between measurements and an ICC of 1 indicates a perfect reliability, with a cut-off value of ≥ 0.75 for the lower limit of the 95% confidence interval for defining a very high level of repeatability [16, 24]. With the exception of oxygen saturation, the majority of ICC values found in our study, presented a lower limit of 95% confidence interval >0.75 . To the best of our knowledge, no study has documented the repeatability of cardiorespiratory adaptations in two consecutive ESWTs. O'DONNELL *et al.* [22] reported on the reproducibility of cardiorespiratory parameters during constant work rate cycling exercise. Even if the methodology of ESWT and constant work rate tests are different, both tests are defined as endurance tests; interestingly, the ICC and coefficients of variation observed in our study were similar to the values reported in the study by O'DONNELL *et al.* [22].

Determination of MID values for ESWT

The determination of MID values for ESWT performances was based on two different methods: the distribution-based and the anchor-based approaches. As previously suggested [13, 25, 26], different methods should be combined for the determination of MID values, considering that there is no consensus in favour of a single best methodology for the determination of MID estimate and that this estimate should be viewed as a range of probable values rather than as a discrete value.

In the present study, the distribution-based and anchor-based approaches led to similar MID values, which is reassuring regarding the validity of our estimates. In addition, these values are in accordance with previously suggested MID values obtained in patients with COPD [13]. The present estimates were

TABLE 5 Coefficients of variation for cardiorespiratory parameters at different isotime levels

	Isotime levels				
	25%	50%	75%	100%	Peak
$\dot{V}O_2$	11.8	9.6	7.2	7.0	8.6
$\dot{V}CO_2$	12.1	10.4	7.9	7.3	9.7
Heart rate beats·min ⁻¹	4.8	5.2	4.1	4.1	5.0
SpO ₂	3.1	5.5	4.7	5.5	7.0
RER	4.4	3.6	3.0	3.0	5.2
$\dot{V}E$	14.4	11.7	8.2	7.3	11.3

Data are presented as mean within-subject coefficient of variation %. $\dot{V}O_2$: oxygen uptake; $\dot{V}CO_2$: carbon dioxide output; SpO₂: arterial oxygen saturation measured by pulse oximetry; RER: respiratory exchange ratio; $\dot{V}E$: minute ventilation. 25%, 50%, 75%, 100% correspond to exercise time-points corresponding to 25%, 50%, 75% and 100% of the latest exercise time reached in both endurance shuttle walking tests. Peak corresponds to the highest value reached during the two first endurance shuttle walking tests.

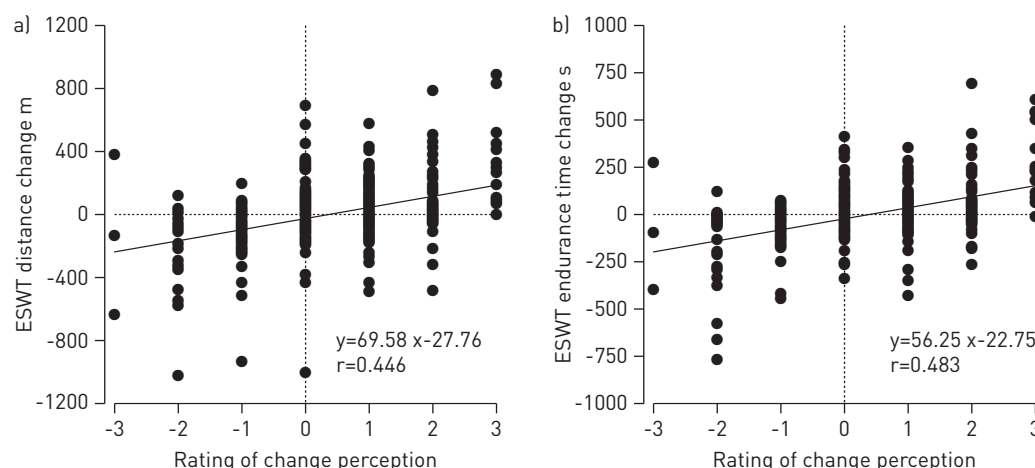


FIGURE 4 Relationships between a) endurance shuttle walking test (ESWT) distance change and b) ESWT time changes and the subjective ratings of change. Each data point represents one subject.

prospectively determined in a larger sample size, thereby adding weight to previous findings and allowing for better precision of these estimates with reduced 95% CI [13].

Nonetheless, these results need to be interpreted by taking into account the modest relationships between changes in ESWT performance and perception of ESWT change. In fact, one important requirement for anchor-based approach interpretability is that the association between the target (change of ESWT performance) and the anchor (perception of performance change) should be appreciable [27]. In this context, some authors have suggested that correlation coefficients of the relationship should be ≥ 0.5 for interpreting anchor-based approaches results satisfactorily. In the case of correlation coefficients between 0.3 and 0.5, the analyses can be performed but the interpretation of the results and estimates requires some caution [26, 28]. With correlation coefficients amounting to 0.48 and 0.45 for endurance time and endurance distance, respectively, and considering the present estimates are very similar to what was previously reported, we are confident about the validity of the proposed estimates.

Our MID results were obtained in the context of a pharmacological intervention and as such may not be applicable to other types of interventions [13], although it has been argued that the same MID values should be applied irrespective of the intervention [29]. All patients involved in this clinical trial had a background therapy of tiotropium and whether this may have influenced the MID results is uncertain. The presence of a baseline potent inhaled therapy probably reduced the increase in endurance time seen with the addition of a second inhaled therapy, a situation that could have decreased the strength of the association between changes in endurance time and changes in patients' perception. Despite this, the wide range of changes in endurance time seen in the trial (fig. 4) suggests that the experimental conditions were adequate to provide valid MID estimates. As in our initial publication on this topic [13], with our anchor we have defined a threshold of change in ESWT that could be perceived by patients. In addition, we refer to this threshold as the MID [30]. The use of additional anchors (e.g. dyspnoea or quality of life questionnaires) would have been useful to ascertain the clinical importance of the observed changes.

One interesting methodological consideration is that patients involved in our trial performed only one ISWT. This study procedure is at variance with the suggestion that two ISWTs should be performed to account for a small but significant learning effect seen with this test [31–37]. This proposed methodology is likely to be important to accurately determine ISWT performance. However, our data would support the idea that performing two ISWTs may not be essential when the purpose of the test is simply to determine the walking speed to be used during the ESWT, as indicated by the good reproducibility of the ESWT coupled with the fact that only one patient reached the maximum 20-min ESWT duration.

In conclusion, this study provides prospective methodological validation of the ESWT by confirming good reproducibility of ESWT performance and by demonstrating the high repeatability of cardiorespiratory measurement under stable condition. With MID values ranging from 56 to 61 s and 70 to 82 m for endurance time and endurance distance, respectively, our study also provides additional information for the clinical interpretation of changes in ESWT performance after a therapeutic intervention. Thus, the results of this study highlight the fact that the ESWT can be considered as a valid alternative for the evaluation of exercise tolerance in COPD, whether in a clinical setting or in the context of scientific research.

Acknowledgements

We would like to thank Glenn D. Carter, Andrea M. Morris, Amanda H. Emmett and Tom J. Ferro (GlaxoSmithKline, Research Triangle Park, NC, USA), and all the research staff for their commitment and involvement in this study. We also want to thank the study subjects for their participation in this project.

References

- Martinez FJ, Foster G, Curtis JL, *et al.* Predictors of mortality in patients with emphysema and severe airflow obstruction. *Am J Respir Crit Care Med* 2006; 173: 1326–1334.
- Oga T, Nishimura K, Tsukino M, *et al.* Analysis of the factors related to mortality in chronic obstructive pulmonary disease: role of exercise capacity and health status. *Am J Respir Crit Care Med* 2003; 167: 544–549.
- O'Donnell DE, Flüge T, Gerken F, *et al.* Effects of tiotropium on lung hyperinflation, dyspnoea and exercise tolerance in COPD. *Eur Respir J* 2004; 23: 832–840.
- Laviolette L, Bourbeau J, Bernard S, *et al.* Assessing the impact of pulmonary rehabilitation on functional status in COPD. *Thorax* 2008; 63: 115–121.
- Maltais F, Mahler DA, Pepin V, *et al.* Effect of fluticasone propionate/salmeterol plus tiotropium *versus* tiotropium on walking endurance in COPD. *Eur Respir J* 2013; 42: 539–541.
- Solway S, Brooks D, Lacasse Y, *et al.* A qualitative systematic overview of the measurement properties of functional walk tests used in the cardiorespiratory domain. *Chest* 2001; 119: 256–270.
- Eaton T, Young P, Nicol K, *et al.* The endurance shuttle walking test: a responsive measure in pulmonary rehabilitation for COPD patients. *Chron Respir Dis* 2006; 3: 3–9.
- Bédard ME, Brouillard C, Pepin V, *et al.* Tiotropium improves walking endurance in COPD. *Eur Respir J* 2012; 39: 265–271.
- Brouillard C, Pepin V, Milot J, *et al.* Endurance shuttle walking test: responsiveness to salmeterol in COPD. *Eur Respir J* 2008; 31: 579–584.
- Pepin V, Saey D, Laviolette L, *et al.* Exercise capacity in chronic obstructive pulmonary disease: mechanisms of limitation. *COPD* 2007; 4: 195–204.
- Pepin V, Saey D, Whittom F, *et al.* Walking *versus* cycling: sensitivity to bronchodilation in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2005; 172: 1517–1522.
- Revill SM, Morgan MD, Singh SJ, *et al.* The endurance shuttle walk: a new field test for the assessment of endurance capacity in chronic obstructive pulmonary disease. *Thorax* 1999; 54: 213–222.
- Pepin V, Laviolette L, Brouillard C, *et al.* Significance of changes in endurance shuttle walking performance. *Thorax* 2011; 66: 115–120.
- Singh SJ, Morgan MD, Scott S, *et al.* Development of a shuttle walking test of disability in patients with chronic airways obstruction. *Thorax* 1992; 47: 1019–1024.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 1: 307–310.
- Lee J, Koh D, Ong CN. Statistical evaluation of agreement between two methods for measuring a quantitative variable. *Comput Biol Med* 1989; 19: 61–70.
- Norman GR, Sloan JA, Wywich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care* 2003; 41: 582–592.
- Guyatt GH, Osoba D, Wu AW, *et al.* Methods to explain the clinical significance of health status measures. *Mayo Clin Proc* 2002; 77: 371–383.
- Revill SM, Noor MZ, Butcher G, *et al.* The endurance shuttle walk test: an alternative to the six-minute walk test for the assessment of ambulatory oxygen. *Chron Respir Dis* 2010; 7: 239–245.
- 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: 291–298.
- Revill SM, Williams J, Sewell L, *et al.* Within-day repeatability of the endurance shuttle walk test. *Physiotherapy* 2009; 95: 140–403.
- O'Donnell DE, Travers J, Webb KA, *et al.* Reliability of ventilatory parameters during cycle ergometry in multicentre trials in COPD. *Eur Respir J* 2009; 34: 866–874.
- Muller R, Buttner P. A critical discussion of intraclass correlation coefficients. *Stat Med* 1994; 13: 2465–2476.
- Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res* 2005; 19: 231–240.
- Borel B, Provencher S, Saey D, *et al.* Responsiveness of various exercise-testing protocols to therapeutic interventions in COPD. *Pulm Med* 2013; 2013: 410748.
- Puhan MA, Mador MJ, Held U, *et al.* Interpretation of treatment changes in 6-minute walk distance in patients with COPD. *Eur Respir J* 2008; 32: 637–643.
- Schunemann HJ, Griffith L, Jaeschke R, *et al.* Evaluation of the minimal important difference for the feeling thermometer and the St. George's Respiratory Questionnaire in patients with chronic airflow obstruction. *J Clin Epidemiol* 2003; 56: 1170–1176.
- Puhan MA, Chandra D, Mosenifar Z, *et al.* The minimal important difference of exercise tests in severe COPD. *Eur Respir J* 2011; 37: 784–790.
- Jones PW, Beeh KM, Chapman KR, *et al.* Minimal clinically important differences in pharmacological trials. *Am J Respir Crit Care Med* 2014; 189: 250–255.
- Redelmeier DA, Bayoumi AM, Goldstein RS, *et al.* Interpreting small differences in functional status: the six minute walk test in chronic lung disease patients. *Am J Respir Crit Care Med* 1997; 155: 1278–1282.
- Dyer CA, Singh SJ, Stockley RA, *et al.* The incremental shuttle walking test in elderly people with chronic airflow limitation. *Thorax* 2002; 57: 34–38.
- Dyer F, Marriner P, Cheema K, *et al.* Is a practice incremental shuttle walk test really necessary? *Chron Respir Dis* 2011; 8: 201–205.
- Eiser N, Willsher D, Dore CJ. Reliability, repeatability and sensitivity to change of externally and self-paced walking tests in COPD patients. *Respir Med* 2003; 97: 407–414.

- 34 Arnardottir RH, Emtner M, Hedenstrom H, *et al.* Peak exercise capacity estimated from incremental shuttle walking test in patients with COPD: a methodological study. *Respir Res* 2006; 7: 127.
- 35 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 Physl Med Rehabil* 2011; 90: 35–39.
- 36 Hill K, Dolmage TE, Woon L, *et al.* Comparing peak and submaximal cardiorespiratory responses during field walking tests with incremental cycle ergometry in COPD. *Respirology* 2012; 17: 278–284.
- 37 Zainuldin R, Mackey MG, Alison JA. Prescription of walking exercise intensity from the incremental shuttle walk test in people with chronic obstructive pulmonary disease. *Am J Physl Med Rehabil* 2012; 91: 592–600.