ORIGINAL RESEARCH

Reduced Variability of Endurance Time in New Protocols for Exercise Tests in COPD

Ellen Tufvesson [ID][1]
Finn Radner [ID][1]
Georgia Papapostolou[1]
Linnea Jarenbäck [ID][1]
Saga Jönsson[1]
Ulf Nihlén[1]
Jaro Ankerst[1]
Alf Tunsäter [ID][1]
Stefan Peterson[2]
Leif Bjermer [ID][1]
Göran Eriksson [ID][1]

[1]Department of Clinical Sciences Lund, Respiratory Medicine and Allergology, Lund University, Skane University Hospital, Lund, Sweden; [2]Regional Cancer Centre South, Region Skane, Lund, Sweden

Purpose: For exercise testing of COPD patients, a standard endurance test (ET) with constant workload is recommended. The test suffers from large inter-individual variability and need for large sample sizes in order to evaluate treatment effects.

Methods: A new protocol for ET in COPD was designed. In contrast to the standard ET, the new ET involved an increasing workload in order to reduce the standard deviation of endurance time. Two new ETs were compared with the standard ET. In Study A, the new ET started at 75% of the patient's maximum workload (W_{MAX}) and increased stepwise with 3%/2 min until exhaustion. Study B started at 70% of W_{MAX} and increased linearly with 1%/min.

Results: In Study A, that included 15 patients, the standard deviation and range for endurance time and work capacity were narrower for the new versus the standard ET. However, the higher mean workload at end and the low mean work capacity relative to the standard ET indicated that the stepwise increase was too aggressive. In Study B, that included 18 patients, with a modified protocol, the averages for endurance time, workload at end and work capacity were similar for new and standard ET, while the standard deviations and ranges for endurance time and work capacity were kept more narrow in the new ET. The variances for endurance time were not equal between the standard ET and the two new ETs ($p < 0.05$ for both according to Levene's test).

Conclusion: The new ET reduced the number of patients with extreme endurance times (short and long) compared to the standard test. The new test showed a significant lower variance for endurance time, which potentially can lead to fewer patients needed in comparative studies. The overall best results were observed with a low linear increase during endurance.

Keywords: COPD, exercise testing, constant endurance test, endurance time

Introduction

Response to therapy in COPD patients has been mainly assessed by symptoms, number of exacerbations and resting pulmonary function. However, prognosis and quality of life is more accurately associated with exercise (in)tolerance.[1–6] There are several ways of classifying exercise tests, with only minor changes in contemporary publications in comparison to a printed version from WHO in 1971.[7] Various exercise tests can be mainly categorized based on the intensity of effort, either as a maximum or a submaximum test, or based on applying the workload as either on a single constant level throughout the test, or as continuous (linear or stepwise) increased load.[7]

A particular approach of cardiopulmonary exercise testing has gained popularity over the years in evaluating the effect of interventions in COPD patients, namely

Correspondence: Ellen Tufvesson
Department of Clinical Sciences Lund, Respiratory Medicine and Allergology, Lund University, Skane University Hospital, Lund 22185, Sweden
Tel +46 73 640 19 16
Email ellen.tufvesson@med.lu.se

the constant work-rate exercise test.[8] This is a sub-maximum constant cycle endurance test (ET) at 75% of maximum achieved workload, which in the following text will be referred to as the standard ET.[8–12] Exercise time until exhaustion measured by a standard ET is a more sensitive method in detecting changes after intervention in comparison to maximum incremental tests and 6-minute walk tests.[8] Despite this proven superiority, there is a limited usage in COPD of the standard ET, both in clinical trials and clinical practice, due to large intra- and interindividual variability. Exceedingly long endurance times have repeatedly been reported. In the study by Worth et al,[12] one patient bicycled for 30 min, and in the study by Maltais et al,[9] 2 patients had endurance times of around 50 min. Too short endurance times are also problematic for the evaluation of the study results. Ten out of 92 patients had an endurance time of less than 3 min during a standard ET in a study by Vivodtzev et al,[11] and 43 of 463 had an endurance time of less than 3.4 min in a study by O'Donnell et al.[5] Because of the great variability, large sample sizes are needed to evaluate treatment effects in clinical trials.[9–11]

Elimination of long endurance times during the standard ET will reduce the phenomenon of very long post-intervention tests, requiring premature termination by the investigators and therefore invalidating interpretation of the magnitude of intervention-related endurance time change using parametric statistics.[13–15] In an official statement by the ERS,[8] the target endurance time of a pre-intervention standard ET should be between 3 and 8 min, eg, for evaluation of rehabilitation. There are some important physiological, statistical, and practical reasons necessitating this relatively narrow range.[1,4,14,15]

The initial step of cardiopulmonary exercise testing involves acquiring a maximum workload (W_{MAX}) by a maximum incremental exercise test (standard W_{MAX} test). The existing protocols for a standard W_{MAX} test may lead to either an underestimation or overestimation of maximum workload, leading to too long and short endurance time, respectively. In a previous study,[16] a prediction algorithm for calculation of W_{MAX} in COPD patients was developed based on bicycle exercise ET results from large industry-sponsored clinical trials. The best algorithm included anthropometrics and six lung function parameters, of which forced expiratory volume in 1 second (FEV_1) and diffusion capacity for carbon monoxide (DLCO) were the best predictors. The study also included a new W_{MAX} test where the participants

started biking with a 3-min warm-up at 40% of their predicted W_{MAX}, followed by an instant increase to 70% and then a linear increase in workload to achieve the predicted W_{MAX} within 8 min. The two tests correlated well, but the new W_{MAX} test had benefits: continuous (not categorical) increase, total work performed was higher with less deviation and better correlation to the prediction algorithm even though the algorithm was developed from the standard W_{MAX} test.

Although it is difficult to specify all the contributors resulting in large variability in endurance time, it is suggested that an ET performed at a fixed proportion of maximum incremental workload capacity does not provide the same physiological stress among individuals and this approach is unlikely to result in the optimal endurance time.[11]

Our overall aim is to improve the ET to a level where smaller studies comparing treatments can be performed and still conclusive outcomes can be obtained. We therefore aim to reduce the variability of the ET, ie, so that too short and too long endurance times can be avoided according to recommendations of ERS.[8,11,15] In this methodological study, we specifically aimed to investigate a new sub-maximum, dis-continuous ET protocol on COPD patients.

Patients and Methods
Patients
Both studies enrolled outpatients with COPD, with a post-bronchodilator FEV_1 of 40–80% of predicted normal, a ratio of FEV_1 to forced vital capacity (FVC) of less than 0.7 and no cardiovascular co-morbidity preventing exercise testing or any other disease that could put the patients at increased risk or interfere with the study results. Subjects with a history of COPD exacerbation within 6 months prior to the study were excluded.

Ethics
The Regional Ethical Review Board in Lund, Sweden, approved the studies, which complied with the Declaration of Helsinki. Written informed consent was obtained from all patients prior to any study-related procedure.

Study Design
Studies A and B were randomized single-centre cross-over trials. Patients not fulfilling the inclusion/exclusion criteria at visit 1 were excluded. Baseline demographic and lung function data (Table 1) were used to calculate the predicted maximum workload (W_{MAX}) in W for each patient as

Table 1 Patient Baseline Characteristics

	Study A	Study B
Number of patients	15	18
Age, years	71 ± 5	71 ± 5
Male, n (%)	11 (73)	10 (56)
Height, cm	172 ± 9	171 ± 10
Weight, kg	77 ± 12	77 ± 11
BMI, kg/m^2	26 ± 3	26 ± 3
Previous/current smokers, n (%)	14 (93)/1 (7)	17 (94)/1 (6)
Pack years, median (range)	30 (10–68)	33 (10–65)
Time since diagnosis, 2–5 years/>5 years, n	3/12	4/14
FEV_1, L[a]	1.80 ± 0.44	1.76 ± 0.37
FEV_1, % of predicted[a]	60.3 ± 8.6	64.5 ± 9.7
FVC, L[a]	3.62 ± 1.07	3.39 ± 0.84
FEV_1/FVC[a]	0.51 ± 0.09	0.53 ± 0.10
Total CCQ score, median (range)	1.0 (0.2–2.2)	0.9 (0–3.0)
CCQ symptom score, median (range)	1.3 (0.3–2.3)	1.3 (0–3.8)
CCQ mental score, median (range)	0.0 (0–2.0)	0.5 (0–2.5)
CCQ function score, median (range)	1.0 (0–2.8)	0.5 (0–3.0)
Medications used at study entry, n (%):		
Any ICS	9 (60)	5 (28)
Any LABA	8 (53)	9 (50)
Any LAMA	14 (93)	16 (89)

Notes: [a]Post-bronchodilator values. Values are mean ± standard deviation unless otherwise stated.
Abbreviations: BMI, body-mass index; CCQ, Clinical COPD Questionnaire; FEV_1, forced expiratory volume in 1 second; FVC, forced vital capacity; ICS, inhaled corticosteroid; LABA, long-acting β2-receptor agonist; LAMA, long-acting muscarinic receptor antagonist.

previously described.[16] At all visits, patients received a fast-acting bronchodilator, ie, all measurements and exercise tests are post-bronchodilator administration. Before all exercise tests, patients filled out the Clinical COPD Questionnaire (CCQ) and underwent a physical examination including vital signs, ECG and dynamic spirometry. A physician assessed the patient´s clinical status and approved or disapproved further participation. All exercise tests involved continuous measurement of VO_2, VCO_2, respiratory rate and heart rate. Borg dyspnoea score and Borg leg discomfort score[17] were reported every second minute, and measurements of ECG and blood pressure were performed before and after exercise. Both studies included measurements of W_{MAX} from the standard maximum test (Study A) and/or the new maximum test (Study A[16] and Study B). The new maximum test included an initial phase of pedalling for 3 min at 40% (Study A) or 30% (Study B) of the predicted W_{MAX}, followed by an instant increase to 75% (Study A) or 70% (Study B), followed by a stepwise (Study A) or linear (Study B) increase in workload until exhaustion.

Study A

Study A, consisted of a screening visit (visit 1), not including any exercise tests, followed by 4 visits with exercise testing (Figure 1; Study A). At visit 2, enrolled patients were randomized to either the standard W_{MAX} test or the new W_{MAX} test as previously reported,[16] in order to obtain the measured W_{MAX} to be used as the basis for the starting values for each ET.

Patients randomized to the standard W_{MAX} test (performed at visit 2), performed a standard ET (at visit 3) at 75% of measured W_{MAX} until exhaustion was reached in accordance with the standard methodology. Patients randomized to the new W_{MAX} test (performed at visit 2) instead performed a new ET (ET_A; at visit 3). After a 3-minute period of low workload cycling (40% of the measured W_{MAX}), the load was increased to 75% and then stepwise increased with 3% every second minute until the patient stopped due to exhaustion (Figure 2; Study A). At visits 4 and 5 the patients were crossed over to complete the opposite tests (Figure 1; Study A).

After completion of Study A, the new ET_A was modified to aim for a decrease in the higher mean workload and lower work capacity compared to the standard ET. A number of modifications (decreased start workload and less escalation intensity) were implemented to achieve a less intense escalation and for subsequent use in Study B.

Study B

Study B consisted of a screening visit including the new W_{MAX} test (visit 1), followed by two visits with ETs (Figure 1: Study B). At visit 1, eligible patients performed the new W_{MAX} test. At visit 2, enrolled patients were randomized to either the standard ET or the new modified ET (ET_B). The standard ET was the same as in Study A, albeit with a constant workload of 75% of measured W_{MAX} obtained from the new W_{MAX} test instead of the standard W_{MAX} test as in Study A.

The new modified ET_B started with an initial workload for 3 minutes at 30% of measured W_{MAX} obtained from the new W_{MAX} test. Then, the workload was set to 70%, followed by a linear increase of 1.0%-units per min until the patient reached his/her maximum workload and stopped due to exhaustion (Figure 2: Study B). At visit 3 the patients were crossed-over to the other test (Figure 1: Study B).
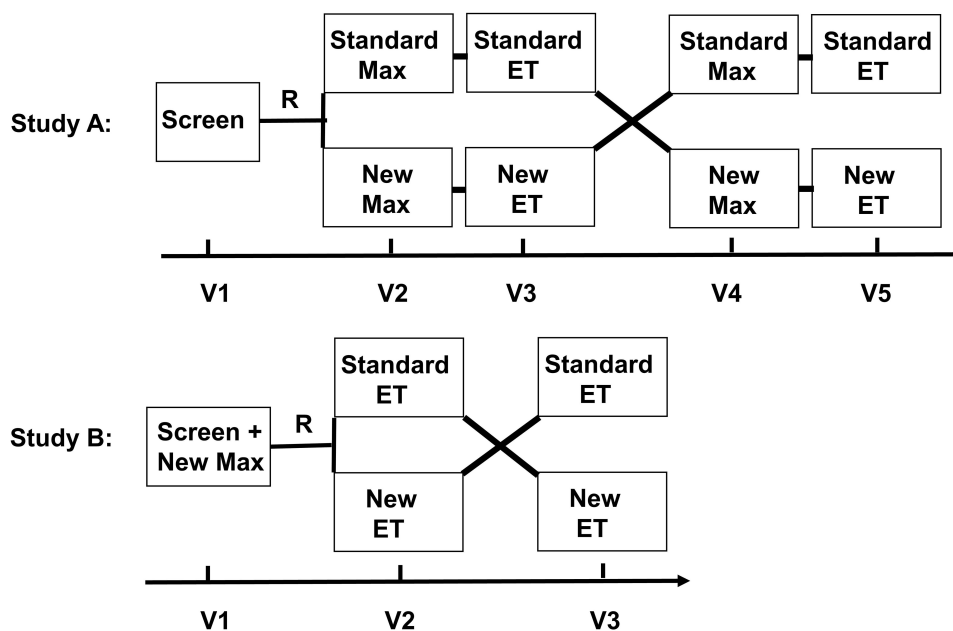
Figure I Schematic picture of study design for standard and new tests in Study A and Study B.
Abbreviations: ET, endurance test; Max, W_{MAX} test; R, randomisation; Screen, screening visit; V, visit.

Study Equipment

Patients performed flow-volume spirometry (MasterScreen, Erich Jaeger GmbH, Würzburg, Germany) and ergospirometry (Oxycon Pro™, Erich Jaeger GmbH, Würzburg, Germany).

Outcome Measures

In both studies, the primary outcome measure was the non-equivalence of the variance of the endurance times for the standard ET and the new ET_A/ET_B, obtained from Levene's test.

The secondary outcome variable was a descriptive comparison between the standard ET and the new ET_A/ET_B for the range of the endurance times. The work capacity (kWs) was compared between the standard ET and the new ET_A/ET_B in the same way as endurance time.

Other outcome variables such as workload at end (W), Borg scale results, oxygen kinetics measurements and reasons for stopping exercise were descriptively compared between the standard and new ET_A/ET_B.

Statistical Analyses

We designed the studies to include 15–20 patients in order to obtain enough data to compare the new ET_A/ET_B with the standard ET, but no formal power calculations were performed. Levene's test for equal variance was performed in IBM SPSS (v 25 for Windows). Further comparisons between

the two tests were performed using descriptive statistics and boxplots using Microsoft Excel for Office 365. Demographics and patient data were expressed as mean ± standard deviation unless otherwise stated. Work capacity in kWs was calculated as area-under-curve from workload (W) x time (s) (Figure 2).

Results
Patient Populations
Study A
In total, 19 patients were enrolled between July and October 2015. Among them, three patients were excluded due to $FEV_1 \geq 80\%$ predicted, and one patient due to significant cardiovascular co-morbidity. The remaining 15 patients were included in the study and performed all study procedures. They had a mean age of 71 years, a mean FEV_1 of 60% of predicted normal, and an FEV_1/FVC ratio of 0.51. Further patient baseline characteristics are presented in Table 1.

Study B
In total, 30 patients were enrolled between July 2016 and February 2018. Six patients were excluded due to $FEV_1 \geq 80\%$ predicted, two patients due to significant cardiovascular co-morbidity and one patient due to other relevant co-morbidity (inguinal hernia). Twenty-one patients were included in the study and performed the new W_{MAX} test at visit 1. Among these 21 patients, one left the study between visit 1 and 2, and two patients completed visit
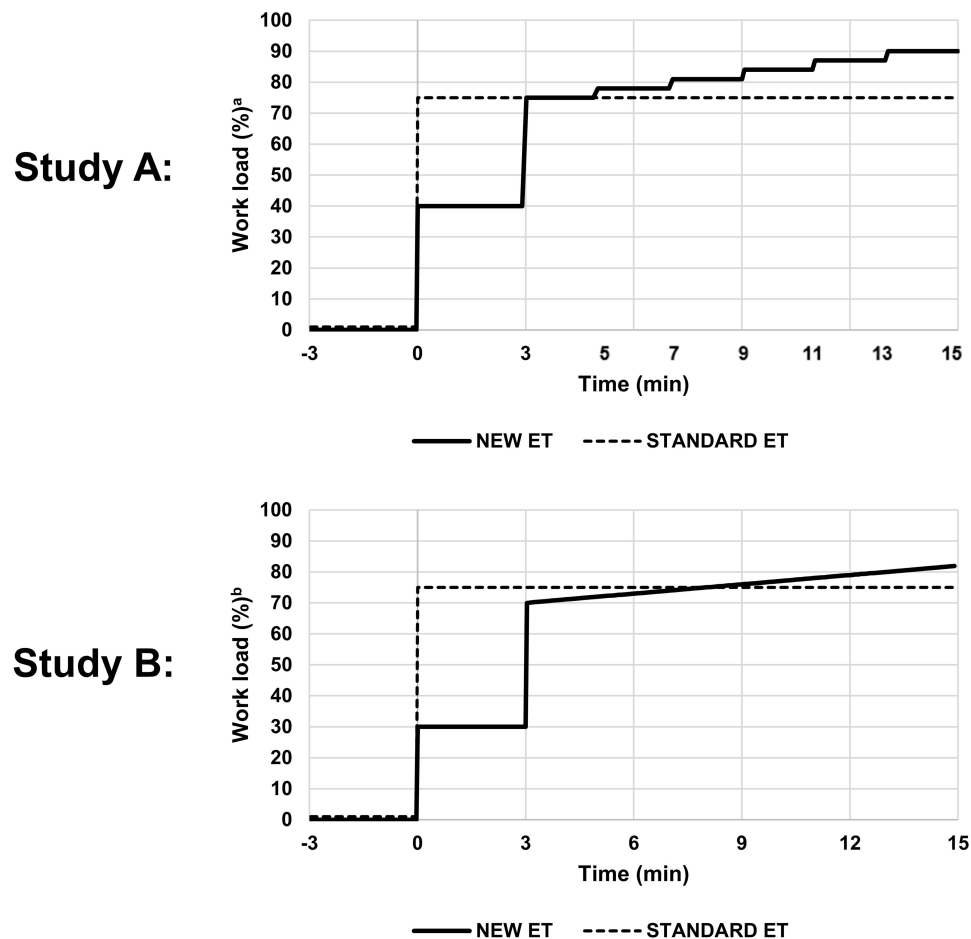
Figure 2 Schematic design of the endurance tests in Study A and Study B.
Notes: [a]Workload in % of maximum value obtained from the new maximum test (in the new endurance test) or from the standard maximum test (in the standard endurance test). [b]Workload in % of maximum value obtained from the new maximum test. Baseline (−3 to 0 min) and recovery phase are similar in the new endurance test and the standard endurance test.
Abbreviations: ET, endurance test; min, minutes.

2, but then left the study, leaving 18 patients who completed all visits. These patients had a mean age of 70 years, a mean FEV_1 of 64% of predicted normal and an FEV_1/FVC ratio of 0.55. Further demographics and baseline characteristics are provided in Table 1.

Results from the W_{MAX} tests in Studies A and B are shown in Table S1.

Comparison of the New and the Standard Endurance Test in Study A

The patients tested with the new ET_A compared to the standard ET reached a numerically higher measured workload at the end of the test (89 versus 80 W). The endurance time was similar between new and standard ET, albeit the new ET_A included 3 minutes on a low workload (40%).

The new ET_A had a narrower range of the endurance time (6.2–14.7 versus 3.3–20.9), showing that fewer patients had a very short or very long endurance time on the new ET_A compared to the standard ET. Two patients had an endurance time ≥20 min and two patients had an endurance time below 5 min during the standard ET, compared to 0 and 0, respectively, with the new ET_A (Tables S2 and S3).

The standard deviation in % of mean endurance time for the new ET_A was less than half that of the standard ET (24% versus 51%). Levene's test showed that the variances for the endurance time of the two tests were not equal; F (1.28) = 12.96; p=0.001.

The new ET_A used less work capacity (46 versus 56 kWs; Table 2, Figure 3) and a narrower range and a lower standard deviation than the standard ET was demonstrated. Levene's test did not show a statistically significant difference in variance (p=0.066).

Table 2 Results from the Endurance Tests in Study A and Study B

	Study A		Study B	
	Standard ET	**New ET$_A$**	**Standard ET**	**New ET$_B$**
Workload				
Workload at start, W	80 ± 17	42 ± 9	75 ± 24	30 ± 9
Workload at 3 min, W	80 ± 17	79 ± 17	75 ± 24	70 ± 23
Workload at end, W	80 ± 17	89 ± 23	75 ± 24	78 ± 27
SD in % of Work load at end	21%	26%	32%	34%
Range of Work load at end, W	53–113	68–153	33–137	33–146
Endurance time				
Endurance time, min	10.9 ± 5.6	10.3 ±2 0.5	9.9 ± 5.6	10.7 ± 3.0
SD in % of endurance time	51%	24%	57%	28%
Median endurance time, min	12.1	10.7	7.9	9.4
Range of endurance time, min	3.3–20.9	6.2–14.7	2.8–21.7	7.4–16.3
Levene's test: Standard vs New	F (1.28) = 12.96; p=0.001		F (1.34) = 4.51; p=0.041	
Work performed				
Work capacity, kWs	57 ± 38	46 ± 21	46 ± 35	42 ± 25
SD in % of work performed	69%	48%	76%	59%
Median Work capacity, kWs	54	39	40	34
Range of Work capacity, kWs	12–142	18–86	9–143	12–97
Levene's test: Standard vs New	F (1.28) = 3.66; p=0.066		F (1.34) = 0.168; p=0.46	

Notes: Values are mean ± standard deviation unless otherwise specified.
Abbreviation: vs, versus.

Five patients stopped the new ET$_A$ due to leg discomfort compared to one patient at the standard ET, while six patients at each test stopped due to dyspnea. Borg scale scores at end and peak VO$_2$ VCO$_2$, minute ventilation and respiratory rate-values were similar between the tests (Table 3).

The higher workload at the end, and the lower work capacity relative to the standard ET, indicated that the step up of workload was too aggressive. Also, more patients stopped the ET due to leg discomfort during the new ET$_A$. A less powerful step up was therefore chosen for Study B, as described in the 'Methods' section.

Comparison of the New and the Standard Endurance Test in Study B

The patients tested with the new ET$_B$ compared to the standard ET reached slightly higher workload at the end of exercise (78 versus 75 W). The mean endurance time was similar to that in study A comparing the new and standard ET, albeit new ET$_B$ included 3 minutes on a low workload (30%).

The new ET$_B$ had a narrower range of the endurance time (7.4–16.3 versus 2.8–21.7) compared to the standard ET, showing that fewer patients bicycled very short or very long (Figure 3). Two patients had an endurance time ≥20 min during the standard ET and 3 patients had
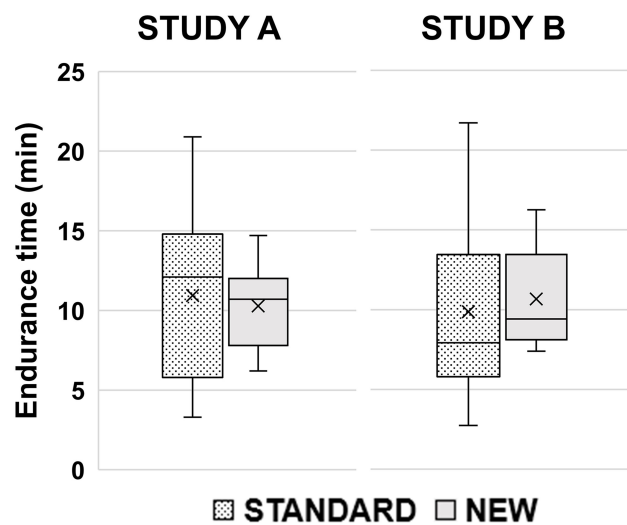

Figure 3 Box plots of endurance times in minutes (min) in the standard and new endurance tests in Study A and Study B.
Notes: Boxes show 25–75 percentiles, whiskers show minimum and maximum values, lines through the boxes show median and crosses show mean values.

Table 3 Borg Scale Scoring, Stopping Reasons and Ergospirometry Results from the Endurance Tests in Study A and Study B

Study	Endurance Test	Borg Dyspnea Score, at End	Borg Leg Discomfort Score, at End	Reason for Stopping Exercise, n (%)				Ergospirometry results			
				Dyspnea	Dyspnea and Leg Discomfort	Leg Discomfort	Other Reason[a]	Peak VO$_2$ (L/min)	Peak VCO$_2$ (L/min)	Peak VE (L/min)	Peak RR (min^{-1})
A	Standard	8.9 ± 1.1	17.7 ± 1.5	6 (40%)	7 (47%)	1 (7%)	1 (7%)	1.46 ± 0.33	1.59 ± 0.30	67 ± 15	46 ± 7
A	New	9.1 ± 1.1	18.1 ± 1.8	6 (40%)	4 (27%)	5 (33%)	0	1.50 ± 0.37	1.64 ± 0.39	69 ± 15	46 ± 8
B	Standard	7.7 ± 2.4	16.9 ± 2.9	7 (39%)	6 (33%)	4 (22%)	1 (6%)	1.41 ± 0.31	1.48 ± 0.36	61 ± 16	40 ± 6
B	New	7.7 ± 2.4	17.2 ± 2.2	5 (28%)	7 (39%)	5 (28%)	1 (6%)	1.39 ± 0.33	1.48 ± 0.41	63 ± 17	40 ± 7

Notes: Values are mean ± standard deviation unless otherwise specified. [a] For specifications see Table S2, S4 and S5.
Abbreviations: VE, minute ventilation; RR, respiratory rate.

an endurance time below 5 min compared to 0 and 0, respectively, with the new ET$_B$ (Tables S4 and S5). The standard deviation in % of mean endurance time for the new ET$_B$ was less than half that of the standard ET (28 versus 57%). Levene's test showed that the variances for the endurance time of the two tests were not equal (F (1.34) = 4.51; p=0.041).

The new ET$_B$ used slightly less work capacity (42 versus 46 kWs; Table 2, Figure 3) and demonstrated a narrower range and a lower standard deviation in comparison to the standard ET, but Levene's test did not demonstrate a significant difference in variance (p=0.46).

Table 3 presents Borg scale results and reasons for stopping the ET. Five patients stopped the new ET$_B$ due to leg discomfort compared to four for the standard ET, while five and seven patients, respectively, stopped due to dyspnea. Borg scale scores at end and peak VO$_2$, VCO$_2$, minute ventilation and respiratory rate-values were similar between the tests (Table 3).

Comparison of the Two New ET Tests (ET$_A$ and ET$_B$) Using the Standard ET as Reference

The mean endurance time was similar across all three ET tests. However, the standard deviation for endurance time (tested as difference in variance using Levene's test) was statistically significantly lower for both the two new ET tests as compared to the standard ET. In addition, the narrower range of endurance time in both studies resulted in similarly low numbers of patients cycling very short and very long compared to the standard ET.

However, compared to the standard ET the mean workload at end differed numerically less in ET$_B$ (78 versus 74 W=4%) than in ET$_A$ (89 versus 80 W=11%). In addition, the mean work capacity also differed less in ET$_B$ (42 versus 46=−9%) than in ET$_A$ (46 versus 56=−18%) compared to the standard ET. The finding of a lowered standard deviation for work capacity was similar in the two new ET tests, although the test of variance did not reach significance (ET$_A$: 48 versus 69%, p=0.066 and ET$_B$: 59 versus 76%, p=0.46). In addition, the narrower range of work capacity was seen for both new ETs compared to the standard ET.

Borg scale scores at end of dyspnea and leg discomfort, as well as oxygen kinetic data at peak, were similar across all three ETs. The number of patients that stopped due to dyspnea was also similar in all studies. However, the

number of patients who stopped due to leg discomfort was larger in ET_A compared to the standard ET (5 (33%) versus 1 patient (7%)), but in study B, the number of patients who stopped due to leg discomfort was similar to the standard ET (5 (28%) versus 4 (22%)).

Taking all the above results into consideration, ET_B was preferred over ET_A in relation to the standard ET.

Discussion

Our objective was to develop a new sub-maximum ET for COPD patients, with only a low number of patients required for making a clinically relevant evaluation of treatment effects. It was designed to achieve both a reasonably narrow range of endurance time (avoiding too short and too long endurance times), and a variance well below that of the standard ET. The results of this study suggest that a new sub-maximum ET with increasing workload is capable to narrow the range of endurance time and reduce its variance compared to the standard constant workload ET.

In Study A, the sub-maximum ET started at 40% of W_{MAX} from the preceding W_{MAX} test, increased to 75% after three minutes and thereafter the workload increased stepwise with 3%/2min until exhaustion. This protocol significantly reduced standard deviation with 50% (ie, the variance differed significantly) and reduced the range for endurance time compared to the standard ET. Also, the work capacity had lower standard deviation (30%, NS). The mean endurance times were similar between the two tests, but workload at the end increased (+11%) and work capacity decreased (−18%) with the new protocol compared to the standard ET. Our conclusion was that the new test fulfilled our objectives, but the stepwise escalation resulted in a workload at end close to 90% of W_{MAX}. In Study B we therefore reduced the workload to 30% during the initial three minutes, followed by an instant increase to 70%, ie, lower than in the first study and also lower than the 75% in the standard ET. Thereafter, the workload was increased with 1% per minute in a linear fashion until exhaustion was reached. This protocol also resulted in a significantly reduced standard deviation with 50% (ie, the variance differed significantly) and reduced the range for endurance time. Mean workload at the end and mean work capacity came closer to the standard ET.

Both protocols fulfilled the objective of lowering the variance and making the range of the endurance time narrower. However, the sub-maximum linear new test protocol in Study B was superior. Firstly, because many

mean values of outcomes were similar to the standard ET, including similar of patients that stopped due to leg fatigue. Secondly, because the lower starting loads of 30% and 70% would benefit patients with a preceding low W_{MAX} test and where W_{MAX} was over-estimated, ie, 70% will be more forgiving than a 75% start. Thirdly, because the linear 1%/min escalation meant a slower continuous increase compared to a higher stepwise (discontinuous) increase, while keeping endurance phase well below 20 min.

A constant work rate at 75% is most commonly used in exercise testing in COPD but also a constant work rate at 85% has been used. The constant work rate test has been shown to be reliable and reproducible but there are some limitations. As the relation between endurance time and work capacity is not linear, a fixed constant percentage of the ET will be dependent on whether the patients are below or above the time versus workload curve and the accuracy of the incremental maximum test. Both can lead to extremely short and long endurance times. In Study B, the range of endurance time was 2.8 to 21.7 minutes for the standard ET. In two studies[5,11] (one including pooled data from two separate studies[5]), the ranges for the standard ET were 1–25 minutes and it has been reported that COPD patients have kept on bicycling for 30 min[12] and 50 min,[9] respectively. The shorter endurance times (<2min) can be unreliable and difficult for recordings, while longer times (>20min) will lead to stopping independent of work capacity (investigator decision, boredom, seat pain, exhaustion of the muscles, etc.).[1,4,8]

Puente-Maestu et al[8] recommended a target of 3–8 min for the pre-intervention standard ET at 75% to achieve a narrower and common baseline in rehabilitation studies, which can reduce the number of patients needed to show effect.[8,14,15] Fifty-seven percent of COPD patients (25% <3min and 18%>8min) was within this range in the two studies described above.[5,11] A solution to increase the number of patients with reasonable endurance times could be to redo the ET at 5% lower or higher dependent on the endurance time being low or high. The latter could lead to an additional 30% of the patients within the set range.[5,11,18] This solution will increase laboratory time and cost, but it is also less feasible in a cross-over study to investigate different pharmaceuticals.

Our research question was: Why should a sub-maximum ET be conducted at constant workload, given all data showing the skewness of the constant load and the knowledge that the endurance time and the work capacity

is not linear. We found one publication where the work-load was increased during the ET.[10] In their protocol, it was stated that patients reaching 15 minutes at 75% of W_{MAX} had the workload increased to 90%. In Study A, we introduced a step-wise escalation which made the range of the endurance narrower with some patients reaching 90% of W_{MAX}. In Study B, as explained above, the strategy was somewhat changed, partly to address patients with short endurance time, ie, to start the escalation phase at 70%. With a linear increase of workload, it will take 5, 10 and 15 min on the escalation phase to reach 75%, 80% and 85% of W_{MAX}, respectively. Out of 18 patients, the shortest and longest time on this escalation phase were 4.4 and 13.7 minutes, compared to 2.8–21.7 for the standard test. This protocol is adapted to reduce the number of outliers, who probably do not represent specific COPD patients but rather under- or over-estimation of the preceding incremental maximum test.[8]

The inclusion and exclusion criteria in the two studies were standard, eg, FEV_1 40–80%, FEV_1/FVC ratio<70%, no exacerbations the preceding 6 weeks and no cardiovascular comorbidities preventing the test. There were no further selection criteria to recruit patients more prone to response in treatment intervention studies, such as FRC>120%[9,10,12,19] and reversibility of IC[19] used in other studies.

The initial 3-minute phase at a low workload (30–40% of W_{MAX}) was added to investigate oxygen kinetics (VO_2 and VCO_2) in order to add information about endurance time or work capacity. This initial phase can be deleted, especially if no additional information is achieved. However, patients have reported that this stepwise start is preferred in contrast to a direct start at 75%.

Some limitations of our study could be further discussed. Both studies had a small number of participants and the studies are mainly descriptive. Neither study had a formal power calculation. However, the results for the new ET in the two studies were consistent versus the standard test and met the objectives. Our methodology could also be criticised for not being a true constant-workload test since the workload is increased during the endurance phase. The question then arises of why this test should be constant and whether the patients notice the slow linear increase (1%/min). Another limitation was that the standard ET was terminated by the investigator after 20 min for two patients, due to protocol programming up to 20 min only. This pre-termination resulted in a lower standard deviation than would have been collected without the pre-termination for the standard ET. The presented data including the limitations render a validation study evaluating the benefit of reduced standard variation in a double-blind comparative study.

Conclusion

In summary, a new protocol for exercise testing in COPD has been developed. This new ET was shown to have less variance and a narrower range for endurance time and work capacity, as compared to the standard ET, and hence a reduced number of patients bicycling with extreme endurance times (both short and long). This can potentially lead to fewer patients needed in comparative studies. A sub-maximum ET with flat linear escalation, starting at 70% of W_{MAX}, seemed to provide the best overall results and will be selected for our future studies of endurance time in response to bronchodilators.

Acknowledgments

Funding

Disclosure

The authors report no conflicts of interest in this work.

References

1. American Thoracic S. American College of Chest P. ATS/ACCP statement on cardiopulmonary exercise testing. *Am J Respir Crit Care Med*. 2003;167(2):211–277.
2. 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(12):1326–1334. doi:10.1164/rccm.200510-1677OC
3. Oga T, Nishimura K, Tsukino M, Sato S, Hajiro T. 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(4):544–549. doi:10.1164/rccm.200206-583OC
4. Palange P, Ward SA, Carlsen KH, et al. Recommendations on the use of exercise testing in clinical practice. *Eur Respir J*. 2007;29(1):185–209. doi:10.1183/09031936.00046906
5. 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(4):866–874. doi:10.1183/09031936.00168708
6. Radtke T, Crook S, Kaltsakas G, et al. ERS statement on standardisation of cardiopulmonary exercise testing in chronic lung diseases. *Eur Respir Rev*. 2019;28:154.

7. World Health O.; Andersen KL, Shephard RJ, Denolin H, Varnauskas E, Masironi R. *Fundamentals of Exercise Testing*. Geneva: World Health Organization; 1971.

8. Puente-Maestu L, Palange P, Casaburi R, et al. Use of exercise testing in the evaluation of interventional efficacy: an official ERS statement. *Eur Respir J*. 2016;47(2):429–460. doi:10.1183/13993003.00745-2015

9. Maltais F, Hamilton A, Marciniuk D, et al. Improvements in symptom-limited exercise performance over 8 h with once-daily tiotropium in patients with COPD. *Chest*. 2005;128(3):1168–1178. doi:10.1378/chest.128.3.1168

10. Magnussen H, Paggiaro P, Schmidt H, Kesten S, Metzdorf N, Maltais F. Effect of combination treatment on lung volumes and exercise endurance time in COPD. *Respir Med*. 2012;106(10):1413–1420.

11. Vivodtzev I, Gagnon P, Pepin V, et al. Physiological correlates of endurance time variability during constant-workrate cycling exercise in patients with COPD. *PLoS One*. 2011;6(2):e17007. doi:10.1371/journal.pone.0017007

12. Worth H, Forster K, Eriksson G, Nihlen U, Peterson S, Magnussen H. Budesonide added to formoterol contributes to improved exercise tolerance in patients with COPD. *Respir Med*. 2010;104(10):1450–1459. doi:10.1016/j.rmed.2010.07.006

13. Puente-Maestu L, SantaCruz A, Vargas T, Martinez-Abad Y, Whipp BJ. Effects of training on the tolerance to high-intensity exercise in patients with severe COPD. *Respiration*. 2003;70(4):367–370. doi:10.1159/000072899

14. Whipp BJ, Ward SA. Quantifying intervention-related improvements in exercise tolerance. *Eur Respir J*. 2009;33(6):1254–1260. doi:10.1183/09031936.00110108

15. van der Vaart H, Murgatroyd SR, Rossiter HB, Chen C, Casaburi R, Porszasz J. Selecting constant work rates for endurance testing in COPD: the role of the power-duration relationship. *COPD*. 2014;11(3):267–276.

16. Eriksson G, Radner F, Peterson S, et al. A new maximal bicycle test using a prediction algorithm developed from four large COPD studies. *Eur Clin Respir J*. 2020;7(1):1692645. doi:10.1080/20018525.2019.1692645

17. Borg GA. Psychophysical bases of perceived exertion. *Med Sci Sports Exerc*. 1982;14(5):377–381. doi:10.1249/00005768-198205000-00012

18. Andrianopoulos V, Wagers SS, Groenen MT, et al. Characteristics and determinants of endurance cycle ergometry and six-minute walk distance in patients with COPD. *BMC Pulm Med*. 2014;14:97. doi:10.1186/1471-2466-14-97

19. Berton DC, Marques RD, Palmer B, O'Donnell DE, Neder JA. Effects of lung deflation induced by tiotropium/olodaterol on the cardiocirculatory responses to exertion in COPD. *Respir Med*. 2019;157:59–68. doi:10.1016/j.rmed.2019.09.006