Chronic Respiratory Disease
Volume 16: 1–7

DOI: 10.1177/1479973119853828
journals.sagepub.com/home/crd
SAGE

The minimal important difference for the endurance shuttle walk test in individuals with chronic obstructive pulmonary disease following a course of pulmonary rehabilitation

Jakub Zatloukal[1], Sarah Ward[1], Linzy Houchen-Wolloff[2], Theresa Harvey-Dunstan[2] and Sally Singh[1,2]

Abstract

The endurance shuttle walk test (ESWT) is frequently used as an outcome measure for pulmonary rehabilitation (PR). The minimal important difference (MID) for the ESWT after a course of rehabilitation has not been conclusively confirmed in the literature. The aim was to establish the MID for the ESWT following the 6-week PR programme in patients with chronic obstructive pulmonary disease (COPD). Following the completion of the 6-week PR programme, data from 531 participants were included in the analysis to estimate the MID for the ESWT using both anchor-based and distribution-based methods. Mean age (standard deviation (SD)) was 69.4 (9.1) years, 303 male, FEV_1/FVC 0.51 (0.16). The baseline incremental shuttle walk test (ISWT) was 217.7 (SD 139.8) metres and ESWT 195.8 (SD 118.8) seconds, which increased to 279.6 (SD 149.5) metres and 537.4 (SD 378.3) seconds, respectively, following PR. The mean change was 61.8 (95% confidence interval (CI) 56.0–67.5) metres for the ISWT and 342.0 (95% CI 312.4–371.6) seconds for the ESWT. The distribution method (0.5 × SD) yielded an MID of 173.7 seconds, the global rating of change scale method yielded a value of 279.2 (95% CI 244.9–313.5) seconds for those rating themselves as 'slightly improved' and the ROC method 207 seconds. There was no agreement between the approaches employed. However, we propose that the MID for the ESWT in COPD following a 6-week PR programme is between 174 and 279 seconds.

Keywords

Rehabilitation, exercise test, outcome assessment, exercise, walking

Date received: 4 September 2018; accepted: 3 May 2019

Introduction

Exercise tolerance is a commonly used outcome measure for evaluating treatment interventions in individuals with chronic obstructive pulmonary disease (COPD) in both clinical practice and research. It can be assessed through standard cardiopulmonary exercise testing (CPET, in laboratory setting)[1] or using field walking tests such as the incremental shuttle walk test (ISWT), endurance shuttle walk test (ESWT) or six-minute walk test (6MWT).[2] These

[1] Pulmonary Rehabilitation Department, Glenfield Hospital, University Hospitals of Leicester NHS Trust, Leicester, UK
[2] Centre for Exercise and Rehabilitation Science (CERS), Respiratory Biomedical Research Unit (BRU), Glenfield Hospital, Leicester, UK

Corresponding author:
Jakub Zatloukal, Pulmonary Rehabilitation Department, Glenfield Hospital, University Hospitals of Leicester NHS Trust, Groby Road, Leicester, LE3 9QP, UK.
Email: jakub.zatloukal@uhl-tr.nhs.uk

tests elicit different physiological responses[3,4] and, consequently, when compared, may be more or less responsive following an intervention.

The extent of change post-intervention, while being statistically significant, may not be meaningful for the patient, and therefore a concept for determining the responsiveness and minimal clinically important difference (MCID) for patient-reported outcomes has been derived. MCID was first defined by Jaeschke et al.[5] as 'the smallest difference in score in the domain of interest which patients perceive as beneficial and which would mandate a change in the patient's management'. Commonly, this approach requires patients to report their perceived change using a global rating of change scale. Further, a concept of minimal important difference (MID) has been established which is defined as 'the smallest difference in score in the outcome of interest that informed patients or informed proxies perceive as important and which would lead the patient or clinician to consider a change in the management'.[6] The MID is usually derived from anchor-based methods, which use an external criterion as an anchor which already has an established MID. Another alternative for deriving the MID are distribution-based methodologies that use variability calculations like standard error of measurement, standard deviation (SD) or effect size[7,8]

The MIDs and MCIDs for exercise tests such as 6MWT, ISWT and standard cycle CPET have been previously studied in individuals with COPD.[3,9–14] However, in pulmonary rehabilitation (PR) programmes, treatment is based around aerobic training at a prescribed intensity, for a prolonged period of time; and using a constant work rate test (CWR), might be more responsive to any change elicited by the treatment intervention.[10] The most frequently reported CWR walking test is the ESWT; however, there are limited data describing the MID for the ESWT derived from rehabilitation programmes. To our knowledge, only two studies managed to estimate a MID for the ESWT in PR. The first of these studies reported a value using a distribution-based estimate (186 seconds, $n = 132$ participants) following a 7-week PR programme.[15] A second study estimated the MID using both anchor- and distribution-based methods of analysis (2 values 186–199 seconds, $n = 55$ participants); however, this study was conducted on a limited and specific population of individuals (COPD with chronic respiratory failure), who attended a 12-week PR programme.[16]

The aim of our study was to establish an MID for the ESWT following the completion of a shorter (6-week) PR programme in individuals with COPD with a broad range of disease severity using both anchor-based and distribution-based methodologies.

Methods

Design, setting and participants

This was an evaluation of the PR programme in individuals with a confirmed diagnosis of COPD who completed the 6-week course at the University Hospitals of Leicester NHS Trust. Participants were referred to outpatient PR either during the stable phase of the disease or within 4 weeks of an exacerbation. Analysis included individuals who completed the 6-week PR programme over a time period of 5 years (2013–2017). All participants routinely confirmed to data being collected within the service.

Outcome measures

The ISWT and ESWT were used to assess exercise tolerance and for prescribing the walking training. During the initial assessment, two ISWTs[17] and one ESWT[18] were performed to recommended guidelines.[2] The ESWT was prescribed at 85% of the predicted peak oxygen consumption (VO_{2Peak}) estimated from the ISWT.[19] Health-related quality of life was recorded with the use of self-reported Chronic Respiratory Questionnaire (CRQ-SR).[20]

Calculating the MID

Global rating of change scale. After completing the ESWT at the end of the programme, participants were asked to rate how they found their exercise tolerance using the following question: 'Compared to your endurance walk test before your rehabilitation programme, how would you rate your exercise tolerance now?' Answers were categorized on a 7-point Likert-type scale as (−3) 'large deterioration', (−2) 'moderate deterioration', (−1) 'slight deterioration', (0) 'no change', (1) 'slight improvement', (2) 'moderate improvement' and (3) 'large improvement'.

Anchor-based methods. We used multiple anchors with known MIDs that were acceptably correlated with the ESWT time change ($r \geq 0.3$, $p < 0.05$). The used anchors were ISWT MCID of 48 metres[14] and CRQ-SR with the MID of 0.5 points for each domain.[5] For the ISWT anchor, participants were considered as responders if they had walked a distance greater or equal to 50 metres and non-responders if this was shorter than 50 metres. For the CRQ-SR anchor, participants were considered as responders if the reported change in breathlessness was greater or equal to 0.5 points and as non-responders if the change was smaller than 0.5.

Distribution-based methods. As one of the distribution-based approaches, we used calculation of half of the standard deviation ($0.5 \times SD$) of the change, as it had been previously used in other studies.[15,16,21] As the other approach, we performed Receiver Operating Characteristic analyses (ROC analyses) for ESWT change coupled with anchor variable of a reasonable correlation ($r \geq 0.3$, $p < 0.05$) to determine the MID for ESWT change only using ROC describing meaningful relationship [area under curve (AUC) > 0.7].[8]

Intervention

Participants attended an outpatient PR programme for 6 weeks as per international guidelines[22,23] and adhered to British Thoracic Society guidelines throughout the duration of the study. Sessions were supervised twice weekly and consisted of a warm up, individually prescribed walking training at 85% of predicted VO_2peak derived from the best of two ISWTs, cycling exercise and two upper and two lower limb strength exercises with dumbbells (3 sets of 10 repetitions). Participants were also asked to perform unsupervised walking exercises on a daily basis at home using their prescribed walking speed and one additional session of strength exercises; and to record all exercise in a training diary. Participants were encouraged to gradually progress all exercises whenever possible based on the self-reported Borg dyspnoea scale (0–10)[24] and perceived exertion scale (6–20)[25]; however, this was guided by the physiotherapists in the supervised sessions.

Statistical analysis

Data were analysed using SPSS (version 20, IBM UK Ltd, Hampshire, UK). Baseline variables were normally distributed. Parametric and non-parametric statistics were used and data reported as mean, standard deviation and 95% confidence intervals (CIs), where appropriate. Relationships between various anchors and outcome measures were analysed using the Spearman's correlation coefficient (r). ROC analyses was only performed on variables with meaningful relationship ($r \geq 0.3$, $p < 0.05$) presenting with acceptable probability (AUC > 0.7). The power calculation was based on the change in ESWT of 184.5 seconds (SD 247.4) from a previous study[26] in a balanced one-way analysis of variance. For a medium effect size of 0.5[27] and assuming equal group sizes in the first 4 global rating of change categories (no change, slight, moderate and large improvement), 57 participants would be needed per category (90% power, 5% significance). Cohen's κ was used for analysing the agreement between methods ($\kappa > 0.2$, $p < 0.05$). Statistical significance was set as $p < 0.05$.

Table 1. Baseline characteristics of participants.

Characteristic	Mean (SD) unless stated otherwise
Age (years)	69.4 (9.1)
Gender	
Male (%)	303 (57)
Female (%)	228 (43)
BMI (kg m^{-2})	27.9 (7.1)
FEV_1 (l)	1.29 (0.58)
FVC (l)	2.54 (0.86)
FEV_1/FVC (%)	51.2 (16.4)
MRC	3.29 (1.0)
MRC 1 (%)	6 (1.1)
MRC 2 (%)	110 (20.7)
MRC 3 (%)	195 (36.7)
MRC 4 (%)	163 (30.7)
MRC 5 (%)	57 (10.7)
ISWT (m)	217.7 (139.8)
ESWT (s)	195.8 (118.8)

BMI: body mass index; SD: standard deviation; ESWT: endurance shuttle walk test; ISWT: incremental shuttle walk test; MRC: Medical Research Council dyspnoea scale.

Results

Study population

Data from 531 participants (303 male) who completed the 6-week PR programme were analysed. The mean (SD) age of participants at baseline was 69.4 years (9.1) and BMI 27.9 (7.1). The participant characteristics are shown in Table 1.

Participants' response to pulmonary rehabilitation

The mean (SD) baseline ISWT for all participants was 217.7 (139.8) metres, which increased to 279.6 (149.5) metres ($p < 0.001$) post-rehabilitation, exceeding the MCID for ISWT[14] with a 61.9 (67.1) metres improvement. Participants improved in the ESWT from 195.8 (118.8) seconds at baseline to

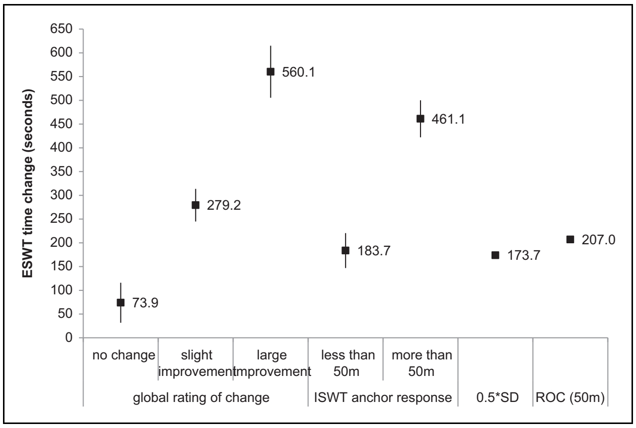

Figure 1. The means for the ESWT time change following 6-week pulmonary rehabilitation in COPD from the present study. ESWT: endurance shuttle walk test; COPD: chronic obstructive pulmonary disease.

537.4 (378.3) seconds ($p < 0.001$) at graduation, which represents a change of 341.6 (347.3) seconds (5 minutes and 41 seconds) for the group.

Anchor-based and distribution-based MID

The relationship between anchors was established using the Spearman's correlations to confirm which anchors could be used for further analysis. Anchors for the ISWT and the CRQ-dyspnoea change were correlated with the change in ESWT (time). Correlations were 0.468 ($p < 0.01$) for the ISWT change response and 0.143 ($p < 0.01$) for CRQ-dyspnoea. The CRQ-dyspnoea was therefore subsequently excluded from any further analysis.

The distribution of data using the MCID anchor for change in ISWT was 'less than 50 metres' in 42.9% ($n = 228$) and 'more than 50 metres' in 57.1% ($n = 303$). The mean change (95% CI) for the ESWT were 183.7 (146.9–220.4) and 461.1 (422.1–500.1) seconds, respectively, according to the ISWT change response anchor. Figure 1 shows the absolute mean changes for each category following the programme completion, which were statistically significantly different ($p < 0.001$).

The distribution of responses in the global rating of change method was 'large improvement' in 30.9% ($n = 164$), 'moderate improvement' in 40.1% ($n = 213$), 'slight improvement' in 17.3% ($n = 92$), 'no change' in 7.5% ($n = 40$) and 'slight deterioration' in 4.1% ($n = 22$). There were no significant differences between 'moderate improvement' and 'slight improvement' ($p = 0.324$), or between 'no change' and 'slight deterioration' ($p = 0.833$). Therefore, these were condensed to form new categories: 'no change' ($n = 62$), 'slight improvement' ($n = 305$) and 'large improvement' ($n = 164$). The mean change (95% CI) for the ESWT based on the global rating of change was 73.9 (31.7–116.0), 279.2 (244.9–313.5) and 560.1 (505.3–614.9) seconds for response categories 'no change', 'slight improvement' and 'large improvement', respectively. The differences between all categories were statistically significant ($p < 0.001$).

The ROC curve analyses was performed using the ISWT anchor change ('50 or more meters') and global rating of change (slight improvement category) that showed a meaningful correlation with the ESWT change ($r = 0.468$ and $r = 0.464$ respectively; $p < 0.01$) with AUC 0.773 and 0.396, respectively. Therefore, the global rating of change was further excluded from ROC analyses (AUC < 0.7). The ISWT change anchor '50 or more meters' was associated with a change of 207 seconds (0.702 sensitivity and 0.699 specificity).

The 0.5 $\times$ SD method yielded an MID value of 173.7 seconds.

Agreement between all methods

To enable the analysis of agreement between the different methods used, participants were divided into non-responder and responder categories within each change variable – ISWT change response anchor ('50 or more metres'), global rating of change (slight and large improvements were merged together as responders) and 0.5 $\times$ SD.

There was poor agreement ($\kappa < 0.2$) between the suggested MID derived from different approaches (global rating of change, distribution-based and anchor-based methods).

Discussion

To our knowledge, this is the largest study to report an MID of the ESWT, and the volume of data pre- and post-rehabilitation exceeds all of the studies included in the ERS statement about field walking tests.[3]

In our study, we have explored the MID using several different approaches. Each approach yielded a slightly different result. This disparity of outcome has previously been documented[3] in relation to the 6MWT. It is therefore perhaps unsurprising that there is a lack of agreement for the proposed MID. The global rating of change method tends to yield a higher value, and of course is the approach that aligns most closely with the definition of a minimum clinically important difference being '*the smallest difference in score in the domain of interest which patients perceive as beneficial*'.[5] The anchor-based methods assume that the change in outcomes are correlated; however, it is consistently documented in the literature that these alternative outcomes reflect different constructs and it is recognized that these are not consistently related.[28] It is plausible that there would be a lack of a convincing relationship between changes in health-related quality of life and exercise performance. Similarly, the two exercise tests, that is, incremental and endurance have fundamentally different properties, and again, you could anticipate that their response to rehabilitation would not necessarily be associated.

Previous literature suggests that the MID for the ESWT, following PR for individuals with COPD and chronic respiratory failure is 186–199 seconds when using the anchor-based method and 144 seconds when using the distribution-based approach.[16] Pepin et al.[15] also suggested an MID of 180 seconds following PR using only a distribution-based estimation ($0.5 \times SD$), as weak correlations between anchors precluded an anchor-based analysis. These values are similar to our current study analysis establishing an MID of around 3 minutes derived from the $0.5 \times SD$ and ROC analyses method (173.7 and 207.0 seconds). Interestingly, participants in our study who did not improve their ISWT by 50 or more metres did still improve in ESWT by 183 seconds ('three minutes') currently recognized as the MID for this test. However, the correlation between ISWT change and ESWT time change in our study was 0.464 ($p < 0.01$), which is acceptable for the anchor analysis ($r > 0.3$, $p < 0.05$), but we do not feel very confident to draw any conclusions from it as the correlation is not greater than 0.5.[13] Given the different constructs of these tests, it would be conceivable that individuals have the capacity to improve their endurance capacity despite no meaningful change in the ISWT.

The value determined from the global rating of change analysis was considerably greater than from the other methods. This is consistent with the literature where these techniques have been compared previously.[11] The improvement necessary is 279 seconds for a 'slight improvement' using the global rating of change technique. We therefore are presented with diverging results ranging from 3 minutes to approximately 4.5 minutes. This is a consequence of different approaches. In the European Respiratory Society/American Thoracic Society statement, the majority of papers included in the analysis were based upon distribution techniques; therefore, in line with this approach, we can confirm that the MID is at least 174 seconds (95% CI 139–198 seconds). The applicability of this value to other interventions, for example, bronchodilation and surgery needs to be confirmed. Currently, the value is only relevant to exercise-based interventions over a 6-week course of PR.

There was no relationship between ESWT walking speed and ESWT change ($r = 0.089$, $p = 0.049$), which suggests that participants walking at a higher speed can still improve in ESWT to a similar degree as less fit participants, and therefore our suggested MID can be used across the whole spectrum of individuals regardless of their baseline fitness level. This is a reflection of the fact that the selected walking speed is relative rather than absolute.

A potential limitation could be our decision of merging categories within the global rating of change. Participants scoring themselves with 'slight' or 'moderate improvement' (92 and 213 participants respectively) had very similar outcomes ($p = 0.324$) with a significantly overlapping CIs where the difference in means of ESWT change between these two categories was 37 seconds (253 and 290 seconds respectively). Therefore, we merged these two categories even though one can argue it does not truly reflect the one's perception of a slight improvement. Another reason for this was that it increased the power of the study, which would otherwise be underpowered due to low number of participants classed as 'no change'. Participants reporting themselves as 'no change' and 'slight deterioration' ($n = 40$ and $n = 22$, respectively; $p = 0.833$) were merged to form new 'no change' category.

Furthermore, we are presenting results of the MID change for ESWT only as a time reference (in seconds), because we believe that the ESWT change in distance has different meanings when related to the different walking speeds. If a distance MID is required, it should be looked into separately for each speed level of the ESWT. Also, it is debatable whether to look into the ESWT change expressed as percentage of change. It is possible that individuals (usually those with low ISWT) may terminate ESWT due to symptom limitation throughout the actual warm-up part of the test, which determines their baseline ESWT value as 0 seconds or 0 metres. In such a case, it is subsequently difficult to determine a percentage change following the intervention.

Conclusion

This is the largest study to propose MID values for the ESWT in individuals with COPD following 6-weeks of PR. We suggest an MID of between 174 and 279 seconds, which is derived from combining a distribution-based and anchor-based method (global rating of change).

Authors' note

This work was supported by the NIHR Leicester Biomedical Research Centre (BRC) – Respiratory. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. Partial data have been presented on the European Respiratory Society congress in Amsterdam, 2015. Abstract can be found in the ERJ supplement as follows: Zatloukal J, Houchen-Wolloff L, Ward S, Singh S. Clinical significance for ESWT in COPD after a course of pulmonary rehabilitation. *Eur Respir J* 2015; 46(Suppl. 59): 2065.

Acknowledgement

The authors would like to thank Dr Matthew Richardson (BRC statistician) for his advice regarding sample size estimation.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Jakub Zatloukal https://orcid.org/0000-0001-9000-4707
Linzy Houchen-Wolloff https://orcid.org/0000-0003-4940-8835
Theresa Harvey-Dunstan https://orcid.org/0000-0001-9654-7075

References

1. ATS/ACCP. ATS/ACCP statement on cardiopulmonary exercise testing. *Am J Respir Crit Care Med* 2003; 167: 211–277.
2. 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. *Eur Respir J* 2014; 44(6): 1428–1446.
3. 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.
4. Casas A, Vilaro J, Rabinovich R, et al. Encouraged 6-min walking test indicates maximum sustainable exercise in COPD patients. *Chest* 2005; 128(1): 55–61.
5. Jaeschke R, Singer J and Guyatt GH. Measurement of health status: ascertaining the minimal clinically important difference. *Control Clin Trials* 1989; 10: 407–415.
6. Schunemann HJ and Guyatt GH. Commentary-goodbye M(C)ID! Hello MID, where do you come from? *Health Serv Res* 2005; 40(2): 593–597.
7. Revicki D, Hays RD, Cella D, et al. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008; 61(2): 102–109.
8. Copay AG, Subach BR, Glassman SD, et al. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J* 2007; 7(5): 541–546.
9. Sutherland ER and Make BJ. Maximum exercise as an outcome in COPD: minimal clinically important difference. *COPD* 2005; 2(1): 137–141.
10. 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.
11. Holland AE, Hill CJ, Rasekaba T, et al. Updating the minimal important difference for six-minute walk distance in patients with chronic obstructive pulmonary disease. *Arch Phys Med Rehabil* 2010; 91(2): 221–225.
12. Redelmeier D, Bayoumi A, Goldstein R, 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(4): 1278–1282.
13. Puhan MA, Chandra D, Mosenifar Z, et al. The minimal important difference of exercise tests in severe COPD. *Eur Respir J* 2011; 37(4): 784–790.
14. Singh SJ, Jones PW, Evans R, et al. Minimum clinically important improvement for the incremental shuttle walking test. *Thorax* 2008; 63(9): 775–777.
15. Pepin V, Laviolette L, Brouillard C, et al. Significance of changes in endurance shuttle walking performance. *Thorax* 2011; 66(2): 115–120.
16. Altenburg WA, Duiverman ML, Ten Hacken NH, 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: 27.
17. 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(12): 1019–1024.
18. Revill SM, Williams J, Sewell L, et al. Within-day repeatability of the endurance shuttle walk test. *Physiotherapy* 2009; 95(2): 140–143.
19. Revill SM, Morgan MD, Singh SJ, et al. The endurance shuttle walk test: a new field test for the assessment of endurance capacity in chronic obstructive pulmonary disease. *Thorax* 1999; 54(3): 213–222.
20. Williams JE, Singh SJ, Sewell L, et al. Development of a self-reported Chronic Respiratory Questionnaire (CRQ-SR). *Thorax* 2001; 56(12): 954–959.
21. Houchen-Wolloff L, Boyce S and Singh S. The minimum clinically important improvement in the incremental shuttle walk test following cardiac rehabilitation. *Eur J Prev Cardiol* 2015; 22(8): 972–978.
22. Bolton CE, Bevan-Smith EF, Blakey JD, et al. British thoracic society guideline on pulmonary rehabilitation in adults. *Thorax* 2013; 68(Suppl 2): ii1–ii30.
23. Spruit MA, Singh SJ, Garvey C, et al. An official American Thoracic Society/European Respiratory Society statement: key concept and advances in pulmonary rehabilitation. *Am J Respir Crit Care Med* 2013; 188(8): e13–e64.
24. Mahler DA, Rosiello RA, Harver A, et al. Comparison of clinical dyspnea ratings and psychophysical measurements of respiratory sensation in obstructive airway disease. *Am Rev Respir Dis* 1987; 135(6): 1229–1233.
25. Borg G. *Borg's perceived exertion and pain scales*. Champaign: Human Kinetics, 1998.
26. Horton EJ, Mitchell KE, Johnson-Warrington V, et al. Comparison of a structured home-based rehabilitation programme with conventional supervised pulmonary rehabilitation: a randomised non-inferiority trial. *Thorax* 2018; 73: 29–36.
27. Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale: Lawrence Erlbaum Associates, 1988.
28. Curtis JR and Patrick DL. The assessment of health status among patients with COPD. *Eur Respir J* 2003; 21(41): 36s–45s.