Systematic review of the association between exercise tests and patient-reported outcomes in patients with chronic obstructive pulmonary disease

Yogesh Suresh Punekar1
John H Riley2
Emily Lloyd3
Maurice Driessen2
Sally J Singh4
1Health Outcomes, GlaxoSmithKline, Uxbridge, 2MDC Global Clinical Development UK, Respiratory Research and Development, GlaxoSmithKline, Uxbridge, 3Value Evidence and Outcomes, GlaxoSmithKline, Brentford, 4Centre for Exercise and Rehabilitation Science, University Hospitals of Leicester NHS Trust, Leicester, UK

Introduction: Chronic obstructive pulmonary disease (COPD) is an increasingly common cause of death worldwide. Its cardinal symptoms include breathlessness and severely reduced exercise capacity. Several patient-reported outcome (PRO) measures are used to assess health-related quality of life (HRQoL), functional performance, and breathlessness in patients with COPD. Exercise testing is employed to measure functional performance objectively, which is generally believed to impact on overall HRQoL. However, the extent to which commonly used laboratory- and field-based exercise test results correlate with PROs has not been systematically assessed.

Materials and methods: A search of Embase, MedLine, and the Cochrane Library identified primary publications in English that reported data on the correlations (Pearson’s r or Spearman’s ρ) between the outcomes of exercise tests and HRQoL and breathlessness PROs. Studies reporting on the following tests were included: 6-minute walk test (6MWT), 12MWT, incremental and endurance shuttle walk tests, incremental and endurance cycle ergometer tests, and treadmill tests.

Results: Of 3,205 articles screened, 28 were deemed eligible for inclusion. The most commonly reported HRQoL PRO measure was the St George’s Respiratory Questionnaire (13 studies), and the most commonly reported breathlessness PRO measure was the Baseline Dyspnea Index (six studies). The St George’s Respiratory Questionnaire appears to correlate very weakly to moderately with the 6MWT, and breathlessness PROs appear to be moderately to strongly associated with 6MWT outcomes. Across all studies, the 6MWT was the most commonly reported exercise test. Very few publications reporting associations between other exercise tests and PRO measures were found.

Conclusion: This review found evidence to support the association of 6MWT outcomes with HRQoL and breathlessness PROs. There were limited data showing correlations with the outcomes of other exercise tests. Further work is required to examine the associations between these PROs and exercise test outcomes.

Keywords: exercise tests, COPD, patient-reported outcomes

Introduction
Chronic obstructive pulmonary disease (COPD) is a leading cause of death worldwide. The prevalence of the disease is projected to increase as the population ages and as exposure to risk factors, such as smoking, continues.1–3 COPD is characterized by breathlessness, episodes of exacerbations, and reduced exercise capacity.4 Decreased
exercise capacity can result in reduced ability to perform the activities of daily living, and the resultant inactivity and sedentary lifestyle can further exacerbate exercise-capacity impairment.\textsuperscript{5,6}

In clinical practice, spirometry is recommended by the Global Initiative for Chronic Obstructive Lung Disease (GOLD) for the diagnosis of COPD.\textsuperscript{4} However, spirometry is a poor predictor of disability and health-related quality of life (HRQoL) in patients with COPD,\textsuperscript{7} and correlates weakly with dyspnea, exercise capacity, and health status.\textsuperscript{8–10} Functional evaluations, such as exercise tests, are thus recommended in addition to spirometry. However, there is currently no single standard test for the assessment of exercise capacity. Field-based tests routinely used to measure exercise capacity include the 6-minute walk test (6MWT),\textsuperscript{11} 12MWT,\textsuperscript{12} incremental shuttle walk test (ISWT),\textsuperscript{13} and endurance SWT (ESWT).\textsuperscript{14} Laboratory-based assessments allow the researcher to monitor several concomitant physiological variables, such as heart rate, workload performed, and oxygen consumption. Such tests include the incremental cycle ergometer test (ICET), endurance CET (ECET), and treadmill test (TT). It has been demonstrated that COPD and its consequent limitation of daily activity affect the HRQoL of patients.\textsuperscript{15} Exercise tests are designed to reflect a patient’s exercise capacity, but how accurately they reflect patients’ HRQoL as assessed by patient-reported outcomes (PROs) of HRQoL and breathlessness in individuals with COPD is unclear. While there have been studies in which exercise tests have been included in the validation of new PRO instruments,\textsuperscript{16,17} it has elsewhere been reported that improvements in exercise capacity elicited by rehabilitation do not necessarily correlate with HRQoL PROs in patients with COPD.\textsuperscript{18}

Several questions remain regarding the associations between exercise tests and PROs. Which PROs are independent of exercise and which are not? Which, if any, PROs accurately and consistently reflect the ability of a patient with COPD to perform exercise? The purpose of this systematic review was to assess the strength of the available evidence supporting correlations between the outcomes of different exercise tests and the PROs most commonly used to assess HRQoL and breathlessness.

Materials and methods

Search strategy

Literature searches were conducted using Ovid, incorporating MedLine (1948 to January 22, 2013, then updated to include articles from July 23, 2012 to September 13, 2016), Embase (1974 to January 22, 2013 and July 23, 2012 to September 13, 2016), and the Cochrane Library (to January 22, 2013 and July 23, 2012 to September 12, 2016). Search strings were constructed to identify study publications (including those specific to emphysema and bronchitis) reporting primary data on the outcomes of the following exercise tests in patients with COPD: 6MWT, 12MWT, ISWT, ESWT, ICET, ECET, and TT. The full search strings used have been published previously.\textsuperscript{19}

Study selection

Study selection followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines for performing a systematic literature review.\textsuperscript{20} Publications were initially screened based on titles and abstracts, and full articles were reviewed when their relevance was unclear from the abstract. Publications were excluded if they were review articles, not in English, studied patients with confounding comorbidities (eg, cancers or diabetes), unclear on the precise variables used for regression analysis, or examined an inappropriate intervention (eg, nonbronchodilatory pharmacotherapy or homeopathy). Articles were subsequently included for assessment only if they reported data on the correlations (Pearson’s $r$ and or Spearman’s $\rho$) between any of the prespecified exercise tests and PRO measures, such as the St George’s Respiratory Questionnaire (SGRQ)\textsuperscript{17} total and individual domain scores (while the former is the usual means of reporting SGRQ data, significant correlations may exist within individual domains alone), the 36-item Short-Form Health Survey (SF-36),\textsuperscript{21} the five-domain European QoL questionnaire (EQ-5D),\textsuperscript{2,23} the Chronic Respiratory Disease Questionnaire (CRQ),\textsuperscript{24} the baseline dyspnea index (BDI),\textsuperscript{25} the oxygen-cost diagram (OCD),\textsuperscript{26} the Medical Research Council dyspnea scale (MRC),\textsuperscript{27} and the modified MRC dyspnea scale (mMRC).\textsuperscript{28} MRC and mMRC correlations were reported in combination, given that they report the same scale on different intervals.

Exercise tests and PRO measures

Exercise tests and respiratory-related PRO measures

The SGRQ is a disease-specific, self-administered questionnaire assessing symptoms, activity, and impacts on health status in COPD and asthma. Lower scores are associated with improved health status.\textsuperscript{29} The CRQ is a comprehensive HRQoL questionnaire specific for individuals with COPD that assesses dyspnea, fatigue, emotional function, and mastery (the feeling of control over the disease and its effects). Each component is scored using a 7-point Likert
scale, which can be combined to produce a total score of 20–140, with higher scores indicating improvement. The BDI questionnaire assesses the severity of dyspnea based on the three components of functional impairment, magnitude of task, and magnitude of effort, which are rated from 0 (very severe) to 4 (no impairment). These scores can be combined to produce a focal score of 0–12, with higher scores indicating improved health status. The OCD assesses dyspnea and measures the oxygen requirement of different activity levels. It is scored from 0 to 100 mm, with higher scores indicating greater improvement. The MRC is a simple questionnaire that evaluates the effect of breathlessness on daily activities by grading patients’ perceptions of breathlessness from 1 to 5, with lower grades indicating less breathlessness and improved health status. The mMRC is a 5-point version of this questionnaire.

Exercise tests and non-disease-specific PRO measures
The SF-36 is a generic questionnaire of 36 items, which assesses physical functioning, social functioning, role limitations (physical), role limitations (emotional), emotional well-being, mental health, energy and vitality, pain, general health perceptions, and current general health perceptions compared with the previous year to determine the general health status of a patient. Higher scores indicate better health status. The EQ-5D is another measure that assesses HRQoL across the five dimensions of mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension is scored from 1 (no problem) to 3 (severe problem). The test also includes a visual analog scale to measure general HRQoL from 0 to 100, with 100 representing the best health condition.

Inclusion/exclusion criteria and data abstraction
Owing to a lack of high-quality evidence for associations between these tests and the prespecified PRO measures, we included observational studies in our final analysis in addition to randomized controlled trials. In our literature search, we reviewed articles to identify those presenting Pearson’s and/or Spearman’s correlations between our stated PRO measures and the most commonly reported exercise test outcomes. Studies reporting lung-function variables only as a percentage of age-, sex-, and body-mass index-predicted values were excluded. Publications involving studies assessing multivariate regressions were also excluded, owing to the multifactorial nature of the statistical approach and the unsuitability of the output for aggregation.

Data were primarily abstracted by a single reviewer. A randomly generated selection of 30% of all articles was reviewed by a second reviewer in both phases for quality-control purposes. The following outcomes of exercise tests were recorded: distance or stages achieved for the 6MWT, 12MWT, and ISWT, duration of exercise for the ESWT and ECET, and the highest recorded volume of oxygen consumption (peak VO\textsubscript{2}) and maximum workload (W\textsubscript{max}) for the TT and ICET.

Statistical analysis
Pearson’s and Spearman’s correlations between PRO scores and the most commonly reported exercise test outcomes are presented. Pearson’s correlations are often used to describe the linear association between two variables when comparing continuous variable data. Spearman’s correlations are commonly used to describe the linear association between two sets of ranked (ordinal) data. Correlations are presented as the range of significant values reported in the study publications reviewed. The strength of correlations has been classified according to British Medical Journal guidelines, which regard significant correlation coefficients of 0–0.19 as very weak, 0.2–0.39 as weak, 0.4–0.59 as moderate, 0.6–0.79 as strong, and 0.8–1 as very strong.

Results
Overview of identified studies
The PRISMA-compliant search methodology used to identify relevant articles is summarized in Figure 1. Of 3,205 articles screened, 28 were ultimately deemed eligible for inclusion in this review. Table 1 provides a summary of the studies included.

Correlations between exercise test outcomes and patient-reported quality of life-measure outcomes
St George’s Respiratory Questionnaire total score
In total, 13 study publications reported correlations between the SGRQ total score (SGRQ\textsubscript{total}) and one or more exercise test outcomes (Table 2, Figure 2). Of these, correlations between outcomes from the SGRQ and distance covered in the 6MWT were most commonly reported. These correlations were typically significant, with six articles reporting weak negative Pearson’s correlations of −0.26, −0.26 (P<0.01), −0.37 (P<0.05), −0.37 (P=0.0228), −0.39 (P<0.01), and −0.39 (P=0.01), demonstrating that as distance covered in the 6MWT increases, SGRQ\textsubscript{total} scores decrease, indicating better
health status. There were also two studies reporting Spearman’s correlations of $-0.27^{42}$ (JP de Torres confirmed this was incorrectly reported as 0.27 in the article) and $-0.56^{50}$. Three papers noted nonsignificant associations between the 6MWT and SGRQ$_{activity}^{46,52,55}$. Limited data were available for other exercise tests, with moderate correlations given for the ISWT ($r=-0.55^{45}$, $\rho=-0.55^{43}$), weak–moderate correlations for the ICET ($r=-0.29$ to $-0.36^{50}$ depending on the outcome measure obtained [VO$_2$ or W$_{max}$]), and moderate Spearman’s correlations ($\rho=-0.54^{50}$ and $-0.58^{43}$) for the ECET and TT, respectively.

St George’s Respiratory Questionnaire activity domain
Correlations between exercise test outcomes and the SGRQ activity domain score (SGRQ$_{activity}$) were reported in nine study publications (Table 2).$^{37,43,45,50,55,59,60}$ Of these, four articles noted significant weak–strong negative correlations between 6MWT distance and SGRQ$_{activity}$ ($r=-0.35^{37}, -0.36^{53}$ [weak], and $-0.44^{45}$ [moderate]; $\rho=-0.68^{50}$ and an unspecified correlation of $-0.37^{40}$). Two further studies found strong correlations between SGRQ$_{activity}$ and ISWT outcome ($r=-0.67^{45}$ and $\rho=-0.62^{43}$), with two more finding no significant association for this relationship.$^{55,59}$ One study$^{50}$ reported moderate correlations between SGRQ$_{activity}$ and ICET peak VO$_2$ ($\rho=-0.51$) and strong correlations between SGRQ$_{activity}$ and ICET $W_{max}$ ($\rho=-0.62$) and ECET ($\rho=-0.62$), with a weak correlation for the ICET $W_{max}$ found in one other study ($r=-0.31$).$^{37}$ Finally, one publication reported moderate correlations between TT peak VO$_2$ and SGRQ$_{activity}$ ($\rho=-0.58$).$^{43}$ Again, these negative correlations suggest that as exercise performance increases, SGRQ$_{activity}$ scores decrease, indicating improved health status.

St George’s Respiratory Questionnaire impact domain
Correlations between exercise test outcomes and the SGRQ impact domain score (SGRQ$_{impact}$) were reported in nine articles (Table 2).$^{37,43,45,46,50,55,59,60}$ Five studies found weak–moderate correlations between the SGRQ$_{impact}$ and the 6MWT ($r=-0.22, -0.28, -0.37^{37,50,59}$ and $0.4^{60}$, $\rho=-0.51^{50}$, $\rho=-0.52^{43}$), two found moderate evidence for an association between the SGRQ$_{impact}$ and the ISWT ($r=-0.53^{45}$ and $\rho=-0.48^{45}$), and there was moderate evidence for an association between the SGRQ$_{impact}$ and the ECET ($\rho=-0.50^{50}$) or TT ($\rho=-0.54$).$^{43}$ A further study
Table 1 Summary of studies included

| Study            | Country, N | Study design/details                                                                 | Study duration/ follow-up period                                                                 | Inclusion and exclusion                                                                 | Age (years), sex (n), BMI (kg/m²) | Disease severity (staging method, score) | Pulmonary function |
|------------------|------------|--------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------|-----------------------------------|-----------------------------------------|-------------------|
| Agrawal et al⁵⁰ | India, N=129 | Aimed to identify the correlation between HRQoL using SGRQ and functional exercise capacity using 6MWT in patients with COPD | Patients screened July 1, 2009–June 30, 2010, and recalled their symptoms over 1 month–1 year | Inclusion: >40 years old, postbronchodilator FEV₁ / FVC <70% predicted Exclusion: coexisting lung lesions on chest radiograph, ischemic heart disease on electrocardiography, serum creatinine >1.4 mg/dL, those with confirmed COPD with any contraindications for 6MWT, as given by the ATS | Age: mean 60.96±11.5 Male: 99 (76.74%) BMI: 19.13±3.55 | GOLD Stage I (mild): 15 (11.6%) Stage II (moderate): 60 (46.5%) Stage III (severe): 37 (28.6%) Stage IV (very severe): 17 (13.3%) | FEV₁: L: 1.19±0.62 FEV₁, % predicted: 53.68±21.1 FEV₁/FVC: 58.67±9.83 |
| Bavarsad et al⁵⁰ | Iran, N=36  | Cross-sectional study to investigate relationships among 6MWT, dyspnea, QoL, and disease severity and to identify the predictors of 6MWT in patients with COPD | Patients assessed during an outpatient pulmonary clinic | Inclusion: mild–very severe COPD according to GOLD criteria, 40–70 years old, presenting stable clinical condition without exacerbations in preceding month Exclusion: comorbid conditions likely to reduce exercise capacity eg, unstable angina and MI during previous month and BMI >35 kg/m², weaning dose of oral corticosteroids, cognitive deficit and musculoskeletal disorders, increasing FEV₁ ≥12% following bronchodilator (salbutamol) therapy and consumption of bronchodilator during 6MWt | Age: mean 56.8±8.8 Male: 33 (91.7%) BMI: 24.9±4.7 | Mild–very severe COPD as per GOLD | FEV₁: L: 1.86±0.89 FEV₁, % predicted: 57.8±24.7 FEV₁/FVC: 3.19±1.06 FEV₁, % predicted: 802±22.4 FEV₁/FVC, %: 69.6±16.5 |
| Belza et al⁵⁴   | US, N=63    | Cross-sectional, descriptive Assessments conducted over 4 days                      |                                                                                                  | Inclusion: veterans with COPD about to enter PR program who had not been hospitalized within past 2 months for respiratory problems Exclusion: veterans with severe, unstable cardiovascular disease or rapidly declining clinical course | Age: 65.4±8 Male: 60 Female: 3 BMI: not stated | Moderate–severe COPD (based on FEV₁) | FEV₁, % predicted: 36±16 |

(Continued)
| Study | Country, N | Study design/details | Study duration/ follow-up period | Inclusion and exclusion | Age (years), sex (n), BMI (kg/m²) | Disease severity (staging method, score) | Pulmonary function |
|-------|------------|----------------------|----------------------------------|-------------------------|----------------------------------|----------------------------------------|-------------------|
| Benzo and Sciurba<sup>35</sup> | US, N=50 | Aimed to report range of VO<sub>2</sub> associated with each minute walked during shuttle walk test | Single shuttle walk test performed 30 minutes after bronchodilators during regular scheduled COPD clinic visit | Inclusion: stable COPD Exclusion: acute respiratory symptoms, using antibiotics or oral prednisone in preceding 2 weeks, any physical impediment to walking | Age: 60±12 Male: 24 Female: 26 BMI: 28±6 | Mild/severe, according to GOLD guidelines | FEV<sub>1</sub>, % predicted: 60.5±23.1 PEFV, % predicted: 77.7±17.9 |
| Boer et al<sup>36</sup> | the Netherlands, N=128 (64 with COPD) | Cross-sectional, to evaluate whether activity-based dyspnea scales can substitute for actual functional capacity testing in patients with COPD | Patients were enrolled between January 2001 and November 2001 | Inclusion: smokers with COPD or at risk of COPD Exclusion: smoking history ≤5 years, history of asthma, acute exacerbation or changes in respiratory treatment regimen in preceding 4 weeks, concomitant comorbidity that may influence functional capacity (ie, cardiovascular, neurologic, or endocrine diseases; and/or locomotive limitations) | Age: 51.7±7.7 Male: 57 Female: 71 BMI: 25±3.9 | GOLD At risk of COPD: 63 Stage I (mild): 31 Stage II (moderate): 22 Stage III (severe): 11 | FEV<sub>1</sub>, L: 2.74±0.92 FEV<sub>1</sub>, % predicted: 90.8±23.1 PEFV, % predicted: 4.03±1.04 PEFV, % predicted: 111.2±17.8 |
| Brown et al<sup>37</sup> | US, N=1,218 | Randomized, double-blind, two-period crossover to compare cardiopulmonary exercise testing and 6MWT in patients with emphysema to determine degree of impairment<sup>74,75</sup> | Assessments conducted within 6-week period during baseline evaluation, before pulmonary rehabilitation and randomization to National Emphysema Treatment Trial | Inclusion: with COPD, enrolled in National Emphysema Treatment Trial, with radiographic evidence of emphysema, FEV<sub>1</sub> ≤45% predicted, TLC ≥100% predicted, RV ≥150% predicted, and PaCO<sub>2</sub> ≤60 mmHg who had not smoked in prior 4 months and did not have severe comorbid conditions Exclusion: not stated | Age: 66.6±6.13 Male: 746 Female: 472 BMI: 24.7±3.88 | With COPD with severe or very severe emphysema who were participating in a trial of LVRS | FEV<sub>1</sub>, L: 0.77±0.24 FEV<sub>1</sub>, % predicted: 26.9±7.12 PEFV, L: 2.5±0.78 PEFV, % predicted: 66.8±15.2 |
| Callens et al<sup>38,74</sup> | France, N=50 | Two-part study to determine evidence for dynamic hyperinflation after walking with handheld spirometer and to determine functional consequences in patients with COPD | ND | Inclusion: consecutive patients with COPD receiving their regular treatment, current or past smokers (history of >15 pack-years), clinically stable for preceding 8 weeks | Group 1 (n=20) Median age: 61 (IQR 55–71) Male: 12 Female: 8 Median BMI: 23.6 (22.1–28) | GOLD: group 1, group 2 Stage I (mild): 4, 2 Stage II (moderate): 6, 10 Stage III (severe): 7, 11 | FEV<sub>1</sub>, % predicted: 52 (IQR 37–76), 42 (IQR: 30–57) FEV<sub>1</sub>/PFV, %: 55 (IQR 44–71), 51 (IQR 41–63) |
### Camargo and Pereira<sup>39</sup> Brazil, N=50

**To determine correlations among various dyspnea scales, spirometric data, and 6MWT in symptomatic patients with COPD**

**Patients enrolled between March 2008 and July 2009**

**Inclusion:** consecutive patients with symptomatic COPD (≥40 years old) treated between March 2008 and July 2009, with documented postbronchodilator FEV<sub>1</sub> ≤ 65% of predicted within last 12 months, with smoking history ≥ 10 pack-years

**Exclusion:** dyspnea from any other cause than COPD; those using supplemental oxygen, those unable to perform 6MWT, answer dyspnea questionnaires, or perform pulmonary function tests, presenting with exacerbation in the last 3 months, presenting with radiological abnormalities indicative of other conditions

**Group 2 (n=30)**
- Median age: 60 (IQR 54–68)
- Male: 27
- Female: 3
- Median BMI: 25.3 (21–26.8)

**Stage IV (very severe):** 3, 7

**Age:** 69±8
- Male: 35
- Female: 15
- BMI: 27±5

**Not stated**

### Chuang et al<sup>40</sup> US/Taiwan, N=27

**Measured self-assessed daily activities, scored using oxygen-cost diagram, pulmonary function testing, 6MWT, increasing ramp-pattern cardiopulmonary exercise test performed on cycle ergometer to maximum in patients with COPD**

**Assessments conducted at study entry and exercise tests completed in a random order**

**Inclusion:** outpatients with clinically stable COPD receiving regular schedule of administered bronchodilators with or without oral prednisolone (<10 mg/day), who had peak exercise heart rate ≥ 85% of maximally predicted, RER ≥ 1.09 at peak exercise, at least 4 mmol/L decrease from resting baseline in plasma bicarbonate at peak exercise.

**Age:** 65±6
- Male: 27
- Female: 0
- BMI: not stated

**Moderate–severe COPD, based on most (~90%) patients with FEV<sub>1</sub>/VC < 65%**

**FEV<sub>1</sub>, L:** 1.3±0.4
- FEV<sub>1</sub>, % predicted: 52±12
- FVC, L: 2.7±0.7
- FVC, % predicted: 85±14

(Continued)
Table 1 (Continued)

| Study          | Country, N | Study design/details                                                                 | Study duration/ follow-up period | Inclusion and exclusion                                                                 | Age (years), sex (n), BMI (kg/m²) | Disease severity (staging method, score) | Pulmonary function |
|----------------|------------|--------------------------------------------------------------------------------------|---------------------------------|----------------------------------------------------------------------------------------|-----------------------------------|------------------------------------------|-------------------|
| Cote et al11   | US, N=365  | Prospective study to compare capacity of the peak VO₂ and 6MWT in predicting mortality in COPD patients and to identify thresholds associated with this outcome | Assessments conducted at study entry and after mean follow-up of 67 months | Exclusion: significant arrhythmia or history of malignancy, cardiovascular, or peripheral vascular disease, or locomotion problems | Age: 67±8 Male/female: “mostly men” BMI: 26.8±5.4 | “Wide range of COPD severity” |                                |
| de Torres et al42 | Spain, N=146 | FEV₁-matched case series to explore factors contributing to sex differences in QoL of COPD patients | Patients recruited over 5 years, January 2000- December 2005 | Inclusion: 73 consecutive female patients with COPD attending a pulmonary clinic from January 2000 to December 2005 plus 73 randomly selected male patients with COPD with a similar degree of airflow obstruction; all patients with smoking history ≥20 pack-years, postbronchodilator FEV₁/V FVC <0.7, and clinically stable Exclusion: history of asthma, bronchiectasis, tuberculosis, or other confounding diseases | Age: 63±8 (male); 56±11 (female) Male: 73 Female: 73 BMI: 27±4 (male); 25±7 (female) | GOLD stages I–IV | Not stated |
| Dowson et al43 | UK, N=29   | Investigated whether density-mask analysis of high-resolution computed tomography or exercise capacity were better surrogates for health status in a well-defined, homogeneous group of patients with α₁-antitrypsin deficiency | Recruitment based on date of annual ATD assessment and program for treatment assessment; lung-function testing conducted on same day as clinical | Inclusion: α₁-ATD and macroscopic emphysema selected consecutively from α₁-ATD treatment center Exclusion: asthma, bronchiectasis, liver disease, or other medical problems likely to limit exercise or alter health status | Median age: 52 (IQR 46–60) Male: 19 Female: 10 BMI: not stated | Moderate–severe airflow obstruction | Median FEV₁, L: 1.03 (IQR 0.84–1.41) Median FEV₁, % predicted: 35 Median FEV₁/V C, L: 0.31 (IQR 0.25–0.43) Median FEV₁/V C, % predicted: 37 |
| Study                  | Country | N  | Design                        | Primary Outcome                                                                 | Participants                                                                 |
|-----------------------|---------|----|-------------------------------|---------------------------------------------------------------------------------|------------------------------------------------------------------------------|
| Eakin et al<sup>44</sup> | US      | 143 | To evaluate dimensions underpinning dyspnea ratings, lung function, and respiratory muscle pressures | After receiving salbutamol 5 mg or terbutaline 5 mg and ipratropium bromide 500 mg | Age: 67.5±9  
Male: 69  
Female: 74  
BMI: not stated  
COPD diagnosis by clinical history and pulmonary function tests  
FEV<sub>1</sub>, L: 1.22±0.57  
FEV<sub>1</sub>, % predicted:  
53±24  
PVC, L: 2.54±0.9  
PVC, % predicted:  
77±21 |
| Emtner et al<sup>45</sup> | Sweden  | 21  | Prospective study to evaluate independent contribution of exercise capacity (walking distance) to rehospitalization in COPD patients hospitalized with acute exacerbations of obstructive lung disease | After receiving salbutamol 5 mg or terbutaline 5 mg and ipratropium bromide 500 mg | Age: 65±10  
Male: 2  
Female: 10  
BMI: 23±4  
Hospitalization group (n=12)  
Age: 65±9  
Male: 5  
Female: 4  
BMI: 22±4  
No-hospitalization group (n=12)  
Age: 65±10  
Male: 2  
Female: 10  
BMI: 23±4  
Hospitalization group (n=9)  
Age: 65±9  
Male: 5  
Female: 4  
BMI: 22±4  
No-hospitalization, hospitalization disease severity not stated |
| Heijdra et al<sup>46</sup> | US      | 41  | Observational, prospective study to explore whether differences in QoL between smokers and ex-smokers could be explained by cough and phlegm, differences in pulmonary function tests, or exercise capacity | After receiving salbutamol 5 mg or terbutaline 5 mg and ipratropium bromide 500 mg | Age: 66±8  
Male: 19  
Female: 22  
BMI: 26±6  
COPD defined by ATS guidelines  
FEV<sub>1</sub>, L: 0.97±0.3  
FEV<sub>1</sub>, % predicted:  
38±11 |

(Continued)
| Study | Country, N | Study design/details | Study duration/ follow-up period | Inclusion and exclusion | Age (years), sex (n), BMI (kg/m²) | Disease severity (staging method, score) | Pulmonary function |
|-------|------------|----------------------|---------------------------------|-------------------------|----------------------------------|----------------------------------------|-----------------|
| Hillman et al<sup>27</sup> | Australia, N=26 | To evaluate influence of body composition and peripheral muscle strength on 6MWT using dual-energy X-ray absorptiometry scanning, spirometry, and dynamometry | ND | Exclusion: myocardial infarction within 6 months, ventilator dependency, malignancy, congestive heart failure, hepatic cirrhosis, end-stage renal disease, orthopedic conditions precluding 6MWT performance, or history of psychiatric or neurologic illness that interfered with the study. Inclusion: =85 years old with severe–very severe COPD and minimum smoking history of 20 pack-years recruited from three PR programs. | Age: 71±8 Male: 13 Female: 13 BMI: 25.4±4.6 | GOLD Stage III (severe): 14 Stage IV: 12 | FEV<sub>L</sub>: 0.9±0.4 FEV<sub>L</sub>, % predicted: 32±11 |
| Hodgev et al<sup>28</sup> | Bulgaria, N=20 | To compare cardiovascular and dyspnea responses to 6MWT and ISWT in patients with COPD | Assessment within 2 days | Inclusion: clinically stable COPD patients who had not received systemic steroids at least 2 months before study, but did receive therapy with bronchodilators during the study. Exclusion: history of asthma, allergic rhinitis, atopy, active lung tuberculosis, lung carcinoma, cardiovascular disorders, including myocardial infarction, angina pectoris, pericarditis, valvular diseases (except relative tricuspid insufficiency), arrhythmia, and arterial hypertension requiring drug treatment, disorders of locomotor apparatus, anemias, kidney, liver, or metabolic disorders. | Age: 55.9±8.7 Male: 20 Female: 0 BMI: 27.8±7.7 | As per guidelines recommended by National Consensus Conference | FEV<sub>L</sub>: 1.35±0.72 FEV<sub>L</sub>, % predicted: 42±19 |
| Study | Country | N | Study Design | Inclusion | Exclusion criteria | Age | Male | Female | BMI | COPD severity | FEV₁, L (%) | FEV₁/FVC (%)
|-------|---------|---|--------------|-----------|-------------------|-----|------|--------|-----|---------------|-------------|-------------|
| Kaplan et al | US, N=1,218 | Report to evaluate QoL measures before randomization in National Emphysema Treatment Trial | Assessment during 6–10 week rehabilitation program before randomization in National Emphysema Treatment Trial | Inclusion: from National Emphysema Treatment Trial with radiographic evidence of bilateral emphysema and severe airflow obstruction and hyperinflation, and had completed a PR program | Exclusion: high-risk characteristics for perioperative morbidity and/or mortality, with emphysema thought to be unsuitable for LVRS or with medical conditions or other circumstances making them unable to complete the trial | Mean age: 67 | Male: 746 | Female: 472 | BMI: not stated | Severe airflow obstruction | FEV₁: 0.68±0.22 | FVC: 2.14±0.72 |
| Mangueira et al | Brazil, N=30 | Cross-sectional study to investigate correlations between HRQoL and 6MWT in women with COPD | Assessed once, on day of first medical visit or at follow-up | Inclusion: women with stable COPD monitored via COPD control and treatment program August 4, 2005–October 30, 2006, no exacerbations in preceding 30 days | Exclusion: recent episode of unstable angina or acute myocardial infarction, presenting with cognitive defect, or having concomitant disease that might impede study performance | Age: 64.5±10.4 | Male: 0 | Female: 30 | BMI: 23.9±4.4 | GOLD | Stage I (mild): 14 | FEV₁, % predicted: 58.2±26.8 | FVC, % predicted: 69.2±26.4 |
| Oga et al | Japan, N=36 | Randomized, double-blind, placebo controlled, crossover study to compare 6MWT, progressive cycle ergometry, and cycle endurance test in COPD patients | Assessments conducted on separate days over 2-week period | Inclusion: consecutive male patients with stable COPD from placebo arm of previously reported clinical trial aged >45 years, smoking history >20 pack-years, chest radiographs showing hyperinflation, FEV₁ <80% predicted, postbronchodilator FEV₁/FVC <0.7 | Exclusion: Patients with exacerbations in preceding 3 months, history of asthma, other diseases likely to affect exercise, or hypoxemia at rest | Age: 69±7 | Male: 36 | Female: 0 | BMI: 20.3±3.2 | COPD as per ATS guidelines | FEV₁: 1.07±0.45 | FEV₁, % predicted: 40.3±16.7 |

(Continued)
| Study                | Country, N | Study design/details                                                                 | Study duration/ follow-up period | Inclusion and exclusion                                                                 | Age (years), sex (n), BMI (kg/m²) | Disease severity (staging method, score) | Pulmonary function |
|---------------------|------------|---------------------------------------------------------------------------------------|---------------------------------|-----------------------------------------------------------------------------------------|----------------------------------|----------------------------------------|--------------------|
| O' Reilly et al.   | UK, N=10   | Double-blind, placebo-controlled study to evaluate effects of prednisone 30 mg/day on respiratory function | Assessments conducted before and after three 2-week treatment periods | Inclusion: males with chronic airway obstruction Exclusion: not stated | Mean age: 61 (range 52–70) Male: 10 Female: 0 BMI: not stated | Chronic bronchitis (n=8) | FEV₁ L: 0.81±0.21 FVC L: 2.56±0.593 |
| Pelegrino et al.   | Brazil, N=68 | Cross-sectional study to analyze cardiopulmonary variables in COPD patients with or without depleted lean body mass, before and after 6MWT | ND | Inclusion: stable COPD and postbronchodilator FEV₁/FVC <70% Exclusion: exacerbations within preceding 3 months, signs of water retention, cardiovascular or osteoarticular diseases | Age: 64.3±9.2 Male: 49 Female: 19 BMI: 25.3±4.9 | COPD stage Mild: 15 Moderate: 23 Severe or very severe: 30 | Not stated |
| Peruzza et al.     | Italy, N=60 | To investigate impact of COPD on QoL and functional status in elderly using anthropometric measurements, lung function tests, exercise-tolerance evaluation, and multidimensional assessment | Trial procedures completed during regular checkup visits as outpatients at geriatric day hospital | Inclusion: COPD >65 years of age Exclusion: underweight (BMI <18.5) or obese (BMI >30), ischemic heart disease, experienced changes in medication in preceding 30 days, hospital admission in preceding 6 weeks | Age: 74.5±5.8 Male: 60 Female: 0 BMI: 25.1±3.8 | As per ERS criteria | FEV₁ L: 1.1±0.5 FEV₁ % predicted: 48.1±18.3 |
| Rejeski et al.     | US, N=209  | To assess test–retest reliability of performance tests                                  | Assessments scheduled within 2 weeks | Inclusion: COPD, 55–80 years old, self-reported disability attributed to breathlessness when performing daily activities, prior or current history of smoking, FEV₁/FVC ≥70%, FEV₁ >20% predicted Exclusion: not stated | Age: 67.2±6 Male: 117 Female: 92 BMI: not stated | ATS Stage I (mild): 134 Stage II (moderate): 55 Stage III (severe): 20 | FEV₁ L: 1.57±0.58 FEV₁ % predicted: 57.1±17 |
| Rosa et al.        | Brazil, N=24 | Cross-sectional, descriptive study to evaluate applicability of incremental (shuttle) walk test compared with encouraged 6MWT in COPD patients | Assessments carried out over 1 day | Inclusion: consecutive COPD patients from pulmonary rehabilitation center with PaO₂ 55 mmHg or SpO₂ 92% (at rest and on room air), at least 6 weeks of clinical stability and satisfactory ability to walk unaided Exclusion: SpO₂ of 80% during exercise, with other pulmonary diseases, heart diseases | Age: 67.8±7.5 Male: 17 Female: 7 BMI: 24.2±4.2 | GOLD Stage I (mild): 2 Stage II (moderate): 7 Stage III (severe): 12 Stage IV: 3 | FEV₁ % predicted: 48.6±21 |
| Study | Country | N | Study Design | Inclusion Criteria | Exclusion Criteria | Patients | Measurements |
|-------|---------|---|--------------|--------------------|-------------------|----------|--------------|
| Sun et al | Taiwan | 200 | Prospective cross-sectional study to investigate association between asymptomatic PAD and walking endurance, measured by the 6MWT in patients with COPD | Patients over 4 years of age | Cardiac insufficiency, or other comorbidities considered uncontrolled or significant, and those presenting with formal contraindications to performing exercise tests | | |
| Wegner et al | Germany | 62 | To examine relationship between exercise capacity, as assessed by standardized 6MWT, lung-function parameters, and clinical ratings of dyspnea | Patients assessed over 5 days | | | |
| Wijkstra et al | the Netherlands | 40 | To investigate relative contribution of lung function, maximal inspiratory pressure, dyspnea, and QoL to performance in walking-distance and bicycle ergometer tests in patients with COPD using measurements of lung function, maximal inspiratory pressure, QoL, dyspnea assessment, and exercise capacity | Patients assessed over 2 days after admittance to hospital in stable condition | | | |

**Note:** Values presented are mean ± standard deviation unless otherwise stated.

**Abbreviations:** ATD, antitrypsin deficiency; ATS, American Thoracic Society; BMI, body-mass index; BODE, BMI, airflow obstruction, dyspnea and exercise-capacity index; ERS, European Respiratory Society; FEV₁, forced expiratory volume in 1 second; FVC, forced vital capacity; GOLD, Global Initiative for Chronic Obstructive Lung Disease; HRQoL, health-related quality of life; ISwT, incremental shuttle walk tests; IQR, interquartile range; IVC, inspiratory slow vital capacity; LVRS, lung volume-reduction surgery; 6MWT, six-minute walk test; MI, myocardial infarction; ND, not described; QoL, quality of life; PAD, peripheral arterial disease; PaO₂, partial pressure of arterial oxygen; PR, pulmonary rehabilitation; RER, respiratory exchange ratio; RV, residual volume; SGRQ, St George’s Respiratory Questionnaire; SpO₂, arterial oxygen saturation; TLC, total lung capacity; VC, vital capacity.
Table 2 Correlations between exercise test outcomes and selected quality-of-life PRO measures

|                 | 6MWT | 12MWT | ISWT | ESWT | ICET (\(\text{VO}_{2}\)) | ICET (\(\text{W}_{\text{max}}\)) | ECET (t) | TT (\(\text{VO}_{2}\)) |
|-----------------|------|-------|------|------|--------------------------|-----------------------------|---------|----------------------|
| **SGRQ**       |      |       |      |      |                          |                             |         |                      |
| r               | -0.26<sup>7</sup> 1.218 | -0.26<sup>9</sup> 1.218 | -0.29<sup>4</sup> 365 | -0.39<sup>1</sup> 36 | -0.23<sup>7</sup> 1.218 | -0.29<sup>4</sup> 365 | -0.23<sup>7</sup> 1.218 | -0.23<sup>7</sup> 1.218 |
| \(\rho\)       | -0.27<sup>4</sup> 146 | -0.27<sup>4</sup> 146 | -0.36<sup>4</sup> 29 | -0.49<sup>2</sup> 29 | -0.54<sup>3</sup> 29 | -0.54<sup>3</sup> 29 | -0.54<sup>3</sup> 29 | -0.54<sup>3</sup> 29 |
| **SGRQ**       |      |       |      |      |                          |                             |         |                      |
| r               | -0.35<sup>7</sup> 1.218 | -0.67<sup>4</sup> 29 | -0.31<sup>7</sup> 1.218 | -0.31<sup>7</sup> 1.218 | -0.31<sup>7</sup> 1.218 | -0.31<sup>7</sup> 1.218 | -0.31<sup>7</sup> 1.218 | -0.31<sup>7</sup> 1.218 |
| \(\rho\)       | -0.68<sup>1</sup> 36 | -0.68<sup>1</sup> 36 | -0.51<sup>2</sup> 36 | -0.62<sup>3</sup> 36 | -0.62<sup>3</sup> 36 | -0.62<sup>3</sup> 36 | -0.62<sup>3</sup> 36 | -0.62<sup>3</sup> 36 |
| **SGRQ**       |      |       |      |      |                          |                             |         |                      |
| r               | -0.22<sup>7</sup> 1.218 | -0.53<sup>4</sup> 29 | -0.27<sup>1</sup> 1.218 | -0.27<sup>1</sup> 1.218 | -0.27<sup>1</sup> 1.218 | -0.27<sup>1</sup> 1.218 | -0.27<sup>1</sup> 1.218 | -0.27<sup>1</sup> 1.218 |
| \(\rho\)       | -0.35<sup>2</sup> 36 | -0.35<sup>2</sup> 36 | -0.48<sup>2</sup> 29 | -0.51<sup>3</sup> 29 | -0.51<sup>3</sup> 29 | -0.51<sup>3</sup> 29 | -0.51<sup>3</sup> 29 | -0.51<sup>3</sup> 29 |
| **SGRQ**       |      |       |      |      |                          |                             |         |                      |
| r               | -0.03<sup>7</sup> 1.218 | -0.35<sup>9</sup> 29 | -0.03<sup>7</sup> 1.218 | -0.03<sup>7</sup> 1.218 | -0.03<sup>7</sup> 1.218 | -0.03<sup>7</sup> 1.218 | -0.03<sup>7</sup> 1.218 | -0.03<sup>7</sup> 1.218 |
| \(\rho\)       | -0.34<sup>4</sup> 29 | -0.34<sup>4</sup> 29 | -0.44<sup>2</sup> 29 | -0.44<sup>2</sup> 29 | -0.44<sup>2</sup> 29 | -0.44<sup>2</sup> 29 | -0.44<sup>2</sup> 29 | -0.44<sup>2</sup> 29 |
| **SF-36 physical** |      |       |      |      |                          |                             |         |                      |
| r               | 0.67<sup>4</sup> 63 | 0.18<sup>7</sup> 1.218 | 0.18<sup>7</sup> 1.218 | 0.18<sup>7</sup> 1.218 | 0.18<sup>7</sup> 1.218 | 0.18<sup>7</sup> 1.218 | 0.18<sup>7</sup> 1.218 | 0.18<sup>7</sup> 1.218 |
| \(\rho\)       | 0.53<sup>4</sup> 29 | 0.53<sup>4</sup> 29 | 0.53<sup>4</sup> 29 | 0.53<sup>4</sup> 29 | 0.53<sup>4</sup> 29 | 0.53<sup>4</sup> 29 | 0.53<sup>4</sup> 29 | 0.53<sup>4</sup> 29 |
| **SF-36 mental** |      |       |      |      |                          |                             |         |                      |
| r               | 0.17<sup>8</sup> 1.218 | 0.17<sup>8</sup> 1.218 | 0.17<sup>8</sup> 1.218 | 0.17<sup>8</sup> 1.218 | 0.17<sup>8</sup> 1.218 | 0.17<sup>8</sup> 1.218 | 0.17<sup>8</sup> 1.218 | 0.17<sup>8</sup> 1.218 |
| \(\rho\)       | 0.34<sup>4</sup> 29 | 0.34<sup>4</sup> 29 | 0.34<sup>4</sup> 29 | 0.34<sup>4</sup> 29 | 0.34<sup>4</sup> 29 | 0.34<sup>4</sup> 29 | 0.34<sup>4</sup> 29 | 0.34<sup>4</sup> 29 |
| **CRQ**        |      |       |      |      |                          |                             |         |                      |
| r               | 0.41<sup>4</sup> 62 | 0.41<sup>4</sup> 62 | 0.41<sup>4</sup> 62 | 0.41<sup>4</sup> 62 | 0.41<sup>4</sup> 62 | 0.41<sup>4</sup> 62 | 0.41<sup>4</sup> 62 | 0.41<sup>4</sup> 62 |
| **CRQ**        |      |       |      |      |                          |                             |         |                      |
| r               | 0.25<sup>4</sup> 209 | 0.25<sup>4</sup> 209 | 0.25<sup>4</sup> 209 | 0.25<sup>4</sup> 209 | 0.25<sup>4</sup> 209 | 0.25<sup>4</sup> 209 | 0.25<sup>4</sup> 209 | 0.25<sup>4</sup> 209 |

(Continued)
found a weak correlation between the SGRQ symptom and the ICET $W_{\text{max}}$ ($r=-0.20$).\(^{37}\)

St George’s Respiratory Questionnaire symptom domain

Very limited data were found concerning associations between the SGRQ symptom domain score ($\text{SGRQ}_{\text{symptom}}$) and exercise test outcomes (Table 2). Five studies\(^{37,46,50,53,55,60}\) assessing the association between SGRQ symptom and 6MWT distance reported nonsignificant relationships. In the two studies that reported significant correlations between SGRQ symptom and 6MWT, very weak negative ($-0.03$)\(^{37}\) and weak negative ($-0.35$)\(^{59}\) correlations were found. Two studies\(^{45,55}\) of three\(^{37,45,55}\) examining ISWT found no correlation with SGRQ symptom. One article reported weak ($\rho=-0.34$) and moderate ($\rho=-0.44$) correlations between the SGRQ symptom and the ISWT and TT, respectively.\(^{53}\)

St George’s Respiratory Questionnaire for COPD

One study publication reported on the relationship between the 6MWT and the disease-specific SGRQ for COPD (SGRQ-c); however, this study examined risk factors, and reported only that worse (higher) SGRQ-c scores were an independent predictor of lower 6MWT distances.\(^{62}\)

Other quality-of-life instruments

Too few publications reporting associations between exercise test outcomes and the SF-36, CRQ, or EQ-5D PROs were found to enable any meaningful assessment of their relationships.

### Correlations between exercise test outcomes and self-reported breathlessness

Associations between the BDI and exercise test outcomes were reported for six studies (Table 3; Figure 3).\(^{36,44,48,52,55,56}\) Four articles noted an association between the 6MWT and BDI score: three publications reported moderate–very strong Pearson’s correlations ($r=0.47$–$0.86$),\(^{44,48,56}\) with another reporting a moderate Spearman’s correlation ($\rho=0.49$).\(^{55}\) Only one study assessing this relationship found no significant correlation.\(^{52}\) Two additional publications reported moderate–strong Pearson’s correlations between the BDI and ISWT ($-0.46$ and $0.76$), with another reporting no significant correlation.\(^{53}\) The positive correlations demonstrated in these studies show that PRO

---

**Table 2 (continued)**

|       | 6MWT | 12MWT | ISWT | ESWT | ICET $\left(\text{VO}_2\right)$ | ICET $\left(W_{\text{max}}\right)$ | ECET $\left(t\right)$ | TT $\left(\text{VO}_2\right)$ |
|-------|------|-------|------|------|-------------------------------|-----------------------------|----------------------|-----------------------------|
|       | Corr | n     | Corr | n    | Corr                          | n                           | Corr                   | n                           |
| CRQ_disease |      |       |      |      |                               |                             |                       |                             |
| $r$   | 0.25$^{34}$ | 209   | NS$^{37}$ | 40    |                               |                             |                       | NS                          |

Notes: *Not stipulated whether $r$ or $\rho$ incorrectly reported as 0.267 in the citation (JP de Torres confirmed via email that –0.267 was the correct correlation coefficient, which has been reported to two decimal places here). Shaded areas indicate that no data were found for the relevant association.

Abbreviations: 6MWT, 6-minute walk test; 12MWT, 12-minute walk test; Corr, correlation; CRQ, Chronic Respiratory Disease Questionnaire; ECET, endurance cycle ergometer test; ESWT, endurance shuttle walk test; ICET, incremental cycle ergometer test; ISWT, incremental shuttle walk test; NS, no significance (reported); VO$_2$, oxygen consumption; PRO, patient-reported outcome; SF-36, 36-item Short-Form Health Survey; SGRQ, St George’s Respiratory Questionnaire; t, time; TT, treadmill test; $W_{\text{max}}$, highest workload achieved.

---

**Figure 2** Pearson’s correlations in studies reporting significant associations between exercise test outcomes and SGRQ symptom. Notes: Numbers in parentheses refer to studies reporting significant correlations/total number of studies reporting Pearson’s correlations. Negative correlations indicate reduced SGRQ symptom scores with improvements in exercise test performance. Lower SGRQ scores indicate improved health status.

Abbreviations: 6MWT, 6-minute walk test; 12MWT, 12-minute walk test; ECET, endurance cycle ergometer test; ESWT, endurance shuttle walk test; ICET, incremental cycle ergometer test; ISWT, incremental shuttle walk test; peak VO$_2$, peak rate of oxygen consumption; SGRQ symptom, St George’s Respiratory Questionnaire total score; TT, treadmill test; $W_{\text{max}}$, highest workload achieved.
Correlations between exercise test outcomes and selected patient-reported breathlessness measures

|       | 6MWT | 12MWT | ISWT | ESWT | ICET (VO₂) | ICET (W_max) | ECET (t) | TT (VO₂) |
|-------|------|-------|------|------|------------|--------------|----------|----------|
|       | r    | n    | corr | n    | corr       | n            | corr     | n        |
| BDI   |      |      |      |      |            |              |          |          |
| r     | (0.47, 0.65)⁴⁴ | 143 |       | −0.46³⁶ | 64 |
|       | 0.86⁴⁰ | 20 |       | 0.76⁴⁰ | 20 |
|       | 0.54⁴⁴ | 62 |       |           |              |
|       | NS⁵¹ | 68 |       |           |              |
| ρ     | 0.49⁵⁵ | 24 |       |           |              |
| OCD   |      |      |      |      |            |              |          |          |
| r     | 0.4⁴⁰ | 27 |       | 0.496⁵¹ | 10 | 0.33⁴⁶ | 64 |
|       | (0.49, 0.34)⁴⁴ | 143 |       |           |              |
|       | 0.52⁴⁴ | 62 |       |           |              |
| ρ     | 0.66⁴⁰ | 36 |       |           |              |
|       | NS⁵¹ | 50 |       |           |              |
| MRC/mMRC |      |      |      |      |            |              |          |          |
| r     | −0.66⁷⁷ | 200 |       | Correlated | 50 | (r not stated)⁵⁵ |      |          |
|       | −0.51³¹ | 60 |       |           |              |
|       | −0.52⁴¹ | 365 |       |           |              |
|       | −0.63³⁶ (reported as negative in Table 2, but positive in text) | 62 |       |           |              |
|       | −0.7³⁷ | 26 |       |           |              |
|       | NS³⁹ | 68 |       |           |              |
| ρ     | −0.51³⁹ | 50 |       |           |              |
|       | −0.39³⁸ | 50 |       |           |              |

Notes: Parentheses enclose results of different subgroups within the same study; shaded areas indicate that no data were found for the relevant association.

Abbreviations: 6MWT, 6-minute walk test; 12MWT, 12-minute walk test; BDI, Baseline Dyspnea Index; Corr, correlation; ECET, endurance cycle ergometer test; ESWT, endurance shuttle walk test; ICET, incremental cycle ergometer test; ISWT, incremental shuttle walk test; mMRC, modified Medical Research Council (dyspnea scale); NS, no significance (reported); OCD, oxygen-cost diagram; VO₂, oxygen consumption; t, test; TT, treadmill test; W_max, highest workload achieved.

Scores increase as exercise performance improves, indicating improved health status.

Associations between OCD and exercise test outcomes were reported in seven articles (Table 3; Figure 3). Of these, three publications reported weak–moderate Pearson’s correlations between the 6MWT and the OCD, ranging from 0.34 to 0.52, with a further study finding a strong Spearman’s correlation (ρ=0.66). Only one study found no significant correlation between the 6MWT and the OCD. Single articles noted weak–moderate correlations for the 12MWT (r=0.50) and the ISWT (r=0.33), and a further paper reported strong Spearman’s correlations of 0.61 and 0.74 for the ICET, when peak VO₂ and W_max were used for the main outcome measure, respectively. The same study also found moderate (ρ=0.59) correlations between the OCD and the ECET. Of six studies reporting Pearson’s correlations between the 6MWT and the MRC/mMRC scale, five reported r-values ranging from −0.51 to −0.7. However, one of these studies published Pearson’s correlations between the 6MWT and the MRC/mMRC scale described r² rather than r. For the basis of this analysis, this study was included under the assumption that a Pearson’s correlation was used and r² was a typographical error. One study reported no significant correlation (Table 3; Figure 3), and another two studies reported a significant Spearman’s correlation between MRC/mMRC and the 6MWT (−0.51⁹ and −0.39⁸). The negative correlations demonstrated in these studies show that PRO scores decrease as exercise performance improves, indicating improved health status.

Discussion

This systematic review has shown that there are limited studies available reporting on the correlations between exercise test outcomes and PROs. Of these, the body of evidence describing a relationship between these outcome measures is even smaller. The most commonly reported association was between SGRQ total outcomes and distance covered in the 6MWT. Though typically significant, these correlations were generally weak–moderate. The available evidence also showed correlations between the SGRQ total and SGRQ activity.


and the ISWT, ICET (W_{max}), ECET, and TT, which tended to be moderate–strong. No studies were found that assessed associations between exercise test and SGRQ outcomes (including all subscales) for 12MWT and ESWT. However, all of these correlations must be considered in the context of substantial heterogeneity in study design, disease severity, and sample size.

The majority of studies investigating associations between exercise test outcomes and self-reported breathlessness scores, such as the BDI, OCD, and mMRC, have used the 6MWT. The PROs of these tended to exhibit at least moderate correlations with exercise test outcomes. In particular, the BDI was generally reported to have moderate–very strong Pearson’s correlation with the 6MWT in three of four studies, with a further study reporting a moderate Spearman’s correlation for this relationship.

Among included studies, there was a wide range of study designs and patient cohorts. As high-quality evidence is limited in this field, observational studies were combined with the results of randomized controlled trials, with the risk of affording similar weight to their interpretation. It is thus possible that significant associations could be underrecognized, owing to a type II statistical reporting error, a factor that must be considered when designing or interpreting studies to assess these exercise tests. Many of these studies reported nonsignificant correlations. Given the number of studies with small patient numbers that did report significant correlations, sample size is unlikely to be a factor in the nonsignificant correlations demonstrated by some studies. However, substantial heterogeneity was observed among some of the patient populations in terms of baseline demographics.

There are also inconsistencies in the way in which results were reported, with some study publications reporting incongruous negative/positive correlations compared with others describing the same relationship. It is unclear whether this reflects differences in the way in which the instruments were used or the way in which the statistical tests were applied between instruments and exercise test outcomes. We have presented such values as reported in the source articles. Additionally, the inclusion criteria and COPD severity are often not clearly stated in the articles included in this review. There is thus a risk that the patients in the studies included were not a broadly homogeneous group. Therefore, we would recommend that future studies clearly state inclusion criteria and the clinical rationale for diagnosis whenever possible.

Although not included in the review, consideration should also be given to the correlations demonstrated in studies that fell outside the selected search criteria. Several studies that were not included here showed significant relationships between exercise and breathlessness or PROs, either demonstrating correlations with new tools or using correlation coefficients other than Pearson’s or Spearman’s. For example, associations between exercise capacity and the i-BODE index (body-mass index, airflow obstruction, dyspnea, exercise-capacity index [exercise-capacity measured using ISWT]), the Functional Assessment of Chronic Illness Therapy – fatigue (FACIT-F) instrument, the London Chest Activity of Daily Living (LCADL) scale,
and the McGill Pain Questionnaire (MPQ)\textsuperscript{66} were identified in patients with COPD. Moreover, correlations between the 6MWT and the COPD assessment test (CAT)\textsuperscript{67} and between the desaturation:distance ratio (DDR) and Borg scale\textsuperscript{68} have been identified. The criteria chosen for this review were based on well-established PROs, rather than those that have been more recently developed. However, many of these more novel tests are increasing in popularity, and as such, further investigation into their correlation with breathlessness or HRQoL PROs will be useful to determine their relevance in predicting prognosis in patients with COPD.

Despite the noted limitations, three further recommendations can be drawn directly from these findings. First, one of our included studies reported correlations between two PRO measures (SGRQ\textsubscript{total} and OCD) and two common outcome measures for the ICET: peak V\textsubscript{O\textsubscript{2}} and W\textsubscript{max}. In both, the PRO was more closely correlated with the ICET W\textsubscript{max} than the ICET peak V\textsubscript{O\textsubscript{2}}. This finding complements that of another systematic review conducted by these authors\textsuperscript{69} which showed that exercise test outcomes tended to be more closely correlated with W\textsubscript{max} as a measure of lung function than with peak V\textsubscript{O\textsubscript{2}}. It thus seems prudent to suggest that when investigating ICET, researchers report W\textsubscript{max} values as a priority.

Second, as a combination of the SF-36 and EQ-5D is the instrument of choice for assessing HRQoL of several health technology-assessment bodies, the lack of studies reporting associations between this outcome measure and exercise tests is of note. Clinical trial designers might usefully consider this in assessing the most appropriate clinical trial end point. Third, the results of this review suggest that exercise tests are not interchangeable, and as such, comparisons of different exercise tests in response to interventions are inappropriate.

In conclusion, these findings indicate that only limited evidence is available to support an association between exercise test outcomes and HRQoL and breathlessness PROs in patients with COPD. The evidence that does exist suggests a very weak–moderate negative correlation between the 6MWT and the SGRQ. Both moderate–strong positive and negative correlations between 6MWT outcomes and breathlessness were observed. It has not been possible to assess other tests adequately, though the limited data available suggest that the ISWT, ICET, ECET, and TT may be more closely associated with the SGRQ (and HRQoL outcomes) than the 6MWT. Recent guidelines on the diagnosis and treatment of COPD indicate that the assessment of disease severity is improved by using functional criteria, such as exercise capacity.\textsuperscript{4,70,71} However, the current evidence suggests that no single exercise test accurately reflects HRQoL or breathlessness in patients with COPD. Therefore, despite the paucity of data for some tests, it may be justified to conclude that these tests assess features not measured by these HRQoL outcomes. Consequently, a composite measurement assessing several factors reflective of COPD, such as the BODE index, which evaluates a surrogate of nutritional state (body-mass index), airflow obstruction (FEV\textsubscript{1}), dyspnea (mMRC), and exercise capacity (6MWT), may be a more accurate measure of COPD severity and prognosis.\textsuperscript{72} The BODE index predicts the requirement for hospitalizations among patients with COPD better than either FEV\textsubscript{1} or classic GOLD staging.\textsuperscript{73} It thus seems reasonable to surmise that individual PROs may have limited prognostic ability in patients with COPD, and should be supported by additional measurements wherever possible.

**Acknowledgments**

The authors would like to thank Martin Bell, Iain Fotheringham, and Sarah Cockle for their contribution to the original study, and Jelle Spoorendonk, Weiwei Xu, and Janita Balradi (Pharmerit International) for conducting the updated systematic literature review. Editorial support (in the form of writing assistance, assembling tables and figures, collating author comments, grammatical editing, and referencing) was provided by Rachael Baylie, PhD at Fishawack Indicia Ltd, UK, and was funded by GSK. This study was funded by GSK.

**Author contributions**

All authors contributed to the conception and design of the study, analysis and interpretation of data, and revision of the manuscript, and approved the final version of the manuscript.

**Disclosure**

YSP, JHR, EL, and MD are current employees of GlaxoSmithKline and hold stocks in GlaxoSmithKline. SJS was involved with the development of the incremental shuttle walk test, and has served on advisory boards for GlaxoSmithKline. SJS was part funded by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care East Midlands (CLAHRC EM). Support was also provided by the NIHR Leicester Respiratory Biomedical Research Unit. The views expressed are those of the authors, and not necessarily those of the National Health Service (NHS), the NIHR, or the Department of Health. The work presented here, including
the conduct of the study, data analysis, and interpretation, was funded by GSK (HO-12-12583).

References

1. Lopez AD, Shibuya K, Rao C, et al. Chronic obstructive pulmonary disease: current burden and future projections. Eur Respir J. 2006;27(2):397–412.
2. Halbert RJ, Natoli JL, Gano A, Badamgarav E, Buist AS, Mannino DM. Global burden of COPD: systematic review and meta-analysis. Eur Respir J. 2006;28(3):523–532.
3. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med. 2006;3(11):e442.
4. Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global Strategy for the Diagnosis, Management, and Prevention of COPD. Bethesda (MD): GOLD; 2014.
5. Gossel HR, Lencer NH, Franssen FM, Van Der Vusse GJ, Schols AM. Striking similarities in systemic factors contributing to decreased exercise capacity in patients with severe chronic heart failure or COPD. Chest. 2003;123(5):1416–1424.
6. O’Donnell DE, Gehke KB. Activity restriction in mild COPD: a challenging clinical problem. Int J Chron Obstruct Pulmon Dis. 2014;9:577–588.
7. Jones P, Miravitles M, van der Molen T, Kulich K. Beyond FEV1, in COPD: a review of patient-reported outcomes and their measurement. Int J Chron Obstruct Pulmon Dis. 2012;7:697–709.
8. Cooper CB. The connection between chronic obstructive pulmonary disease symptoms and hyperinflation and its impact on exercise and function. Am J Med. 2006;119(10 Suppl 1):21–31.
9. Jones PW. Issues concerning health-related quality of life in COPD. Chest. 1995;107(5 Suppl):187S–193S.
10. Mahler DA, Harver A. A factor analysis of dyspnea ratings, respiratory muscle strength, and lung function in patients with chronic obstructive pulmonary disease. Am Rev Respir Dis. 1992;145(2 Pt 1):467–470.
11. Butland RJ, Pang J, Gross ER, Woodcock AA, Geddes DM. Two-, six-, and 12-minute walking tests in respiratory disease. Br Med J (Clin Res Ed). 1982;284(6329):1607–1608.
12. McGavin CR, Gupta SP, McHardy GJ. Twelve minute walking test for assessing disability in chronic bronchitis. Br Med J. 1976;1(6013):822–823.
13. Singh SJ, Morgan MD, Scott S, Walters D, Hardman AE. Development of a shuttle walking test of disability in patients with chronic airways obstruction. Thorax. 1992;47(12):1019–1024.
14. Revill SM, Morgan MD, Singh SJ, Williams J, Hardman AE. The endurance shuttle walk: a new field test for the assessment of endurance capacity in chronic obstructive pulmonary disease. Thorax. 1999;54(3):213–222.
15. Lu Y, Nyunt MS, Gwee X, et al. Life event stress and chronic obstructive pulmonary disease (COPD): associations with mental well-being and quality of life in a population-based study. BMJ Open. 2012;2(6):e001674.
16. Garrod R, Marshall J, Barley E, Jones PW. Predictors of success and failure in pulmonary rehabilitation. Eur Respir J. 2006;27(4):788–794.
17. Jones PW, Quirk FH, Baveystock CM, Littlejohns P. A self-complete measure of health status for chronic airflow limitation: the St. George’s Respiratory Questionnaire. Am Rev Respir Dis. 1992;145(6):1321–1327.
18. Wijkstra PJ, Van Altena R, Kraan J, Otten V, Postma DS, Koeter GH. Quality of life in patients with chronic obstructive pulmonary disease improves after rehabilitation at home. Eur Respir J. 1994;7(2):269–273.
19. Fotheringham I, Meakin G, Punekar YS, Riley JH, Cockle SM, Singh SJ. Comparison of laboratory- and field-based exercise tests for COPD: a systematic review. Int J Chron Obstruct Pulmon Dis. 2015;10:625–643.
20. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ. 2009;339:b2535.
43. Dowson LJ, Newall C, Guest PJ, Hill SL, Stockley RA. Exercise capacity predicts health status in alpha-1-antitrypsin deficiency. *Am J Respir Crit Care Med*. 2001;163(4):936–941.

44. Eakin EG, Kaplan RM, Ries AL, Sassi-Damburon DE. Patients’ self-reports of dyspnea: An important and independent outcome in chronic obstructive pulmonary disease. *Ann Behav Med*. 1996;18(2):87–90.

45. Emmett MI, Arnardottir HR, Hallin R, Lindberg E, Janson C. Walking distance is a predictor of exacerbations in patients with chronic obstructive pulmonary disease. *Respir Med*. 2007;101(5):1037–1040.

46. Heijdra YF, Pinto-Plata VM, Kenney LA, Rassulo J, Celli BR. Cough and phlegm are important predictors of health status in smokers without COPD. *Chest*. 2002;121(5):1427–1433.

47. Hillman CM, Heinicke EL, Hii JW, Cecins NM, Jenkins SC, Eastwood PR. Relationship between body composition, peripheral muscle strength and functional exercise capacity in patients with severe chronic obstructive pulmonary disease. *Intern Med J*. 2012;42(5):578–581.

48. Hodgev VA, Allman OL, Marinov BI, Kostianev SV, Mandulova PV. Cardiovascular and dyspnea response to six-minute and shuttle walk tests in COPD patients. *Folia Med (Plovdiv)*. 2003;45(3):26–33.

49. Mangueira NM, Viega IL, Mangueira MA, Pinheiro AN, Costa MR. Correlation between clinical parameters and health-related quality of life in women with COPD. *J Bras Pneumol*. 2009;35(3):248–255.

50. Oga T, Nishimura K, Tsukino M, Hajiro T, Ikeda A, Mishima M. Relationship between different indices of exercise capacity and clinical measures in patients with chronic obstructive pulmonary disease. *Heart Lung*. 2002;31(5):374–381.

51. O’Reilly JF, Shaylor JM, Fromings KM, Harrison BD. The use of the 12 minute walking test in assessing the effect of oral steroid therapy in patients with chronic airways obstruction. *Br J Dis Chest*. 1982;76(4):374–382.

52. Pelegrino NR, Lucheta PA, Sanchez FF, Faganello MM, Ferrari R, de Godoy I. Influence of lean body mass on cardiopulmonary repercussions during the six-minute walk test in patients with COPD. *J Bras Pneumol*. 2009;35(1):20–26.

53. Peruzzo S, Sergi G, Vianello A, et al. Chronic obstructive pulmonary disease (COPD) in elderly subjects: impact on functional status and quality of life. *Respir Med*. 2003;97(6):612–617.

54. Rejeski WJ, Foley KO, Woodard CM, Zaccaro DJ, Berry MJ. Evaluating and understanding performance testing in COPD patients. *J Cardiopulm Rehabil*. 2000;20(2):79–88.

55. Rosa FW, Camelier A, Mayer A, Jardim JR. Evaluating physical capacity in patients with chronic obstructive pulmonary disease: comparing the shuttle walk test with the encouraged 6-minute walk test. *J Bras Pneumol*. 2006;32(2):106–113.

56. Wegner RE, Jorres RA, Kirsten DK, Magnussen H. Factor analysis of exercise capacity, dyspnea ratings and lung function in patients with severe COPD. *Eur Respir J*. 1994;7(4):725–729.

57. Wijkstra PJ, TenVerget EM, van der Mark TW, et al. Relation of lung function, maximal inspiratory pressure, dyspnea, and quality of life with exercise capacity in patients with chronic obstructive pulmonary disease. *Thorax*. 1994;49(5):468–472.

58. Kaplan RM, Ries AL, Reilly J, Mohsenifar Z. Measurement of health-related quality of life in the national emphysema treatment trial. *Chest*. 2004;126(3):781–789.

59. Agrawal SR, Joshi R, Jain A. Correlation of severity of chronic obstructive pulmonary disease with health-related quality of life and six-minute walk test in a rural hospital of central India. *Lung India*. 2015;32(3):233–240.

60. Bavrasad MB, Shariati A, Eidani E, Latifi M. Relationship between exercise capacity and clinical measures in patients with chronic obstructive pulmonary disease. *Biomed Pharmacol J*. 2015;8(1):1–7.

61. Sun KS, Lin MS, Chen YJ, Chen YY, Chen SC, Chen W. Is asymptomatic peripheral arterial disease associated with walking endurance in patients with COPD? *Int J Chron Obstruct Pulmon Dis*. 2015;10:1487–1492.

62. Spruit MA, Watkins ML, Edwards LD, et al. Determinants of poor 6-min walking distance in patients with COPD: the ECLIPSE cohort. *Respir Med*. 2010;104(6):849–857.

63. Utku ES, Ergun PS, Kaymaz DY, et al. i-BODE for evaluating the efficiency of pulmonary rehabilitation in patients with COPD. *Eur Respir J*. 2013;42 Suppl 57:P2230.

64. Alahmari AD, Kowlessar BS, Patel AR, et al. Physical activity and exercise capacity in patients with moderate COPD exacerbations. *Eur Respir J*. 2016;48(2):340–349.

65. Garcia-Polo C, Alazar-Narvarete B, Ruiz-Iturriaga LA, et al. Factors associated with high healthcare resource utilisation among COPD patients. *Respir Med*. 2012;106(12):1734–1742.

66. Ghanbari BH, Garland SJ, Road JD, Reid WD. Pain and physical performance in people with COPD. *Respir Med*. 2013;107(11):1692–1699.

67. Akor AA, Obaseki D, Adeniyi B, et al. COPD assessment test in patients with COPD in Africa: a preliminary survey. *Am J Respir Crit Care Med*. 2014;189:A2951.

68. Iijii N, Kanazawa H, Yoshikawa T, Hirata K. Application of a new parameter in the 6-minute walk test for manifold analysis of exercise capacity in patients with COPD. *Int J Chron Obstruct Pulmon Dis*. 2014;9:1235–1240.

69. Bell M, Fotheringham I, Punekar YS, Riley JH, Cockle S, Singh SJ. Systematic review of the association between laboratory- and field based exercise tests and lung function in patients with chronic obstructive pulmonary disease. *Chronic Obstr Pulm Dis*. 2015;2(3):321–342.

70. National Institute for Health and Care Excellence. *Management of Chronic Obstructive Pulmonary Disease in Adults in Primary and Secondary Care*. Partial update. London: NICE; 2010.

71. American Thoracic Society/European Respiratory Society Task Force. *Standards for the Diagnosis and Management of Patients with COPD*. Version 1.2. New York: American Thoracic Society; 2005.

72. Celli BR, Cote CG, Marin JM, et al. The body-mass index, airflow obstruction, dyspnea, and exercise capacity index in chronic obstructive pulmonary disease. *N Engl J Med*. 2004;350(10):1005–1012.

73. Ong KC, Earnest A, Lu SJ. A multidimensional grading system (BODE index) as predictor of hospitalization for COPD. *Chest*. 2005;128(6):3810–3816.

74. [No authors listed]. Rationale and design of the national emphysema treatment trial: a prospective randomized trial of lung volume reduction surgery. *Chest*. 1999;116(6):1750–1761.

75. Fishman A, Martinez F, Naunheim K, et al. A randomized trial comparing lung-volume-reduction surgery with medical therapy for severe emphysema. *N Engl J Med*. 2003;348(21):2059–2073.

76. Sassi-Damburon DE, Eakin EG, Ries AL, Kaplan RM. Treatment of dyspnea in COPD: a controlled clinical trial of dyspnea management strategies. *Chest*. 1995;107(3):724–729.