Early View

Review

Which functional outcome measures can we use as a surrogate for exercise capacity during remote cardiopulmonary rehabilitation assessments? A rapid narrative review

Linzy Houchen-Wolloff, Enya Daynes, Amye Watt, Emma Chaplin, Nikki Gardiner, Sally Singh


This manuscript has recently been accepted for publication in the ERJ Open Research. It is published here in its accepted form prior to copyediting and typesetting by our production team. After these production processes are complete and the authors have approved the resulting proofs, the article will move to the latest issue of the ERJOR online.

Copyright ©ERS 2020. This article is open access and distributed under the terms of the Creative Commons Attribution Non-Commercial Licence 4.0.
Title: Which functional outcome measures can we use as a surrogate for exercise capacity during remote cardiopulmonary rehabilitation assessments? A rapid narrative review.

Authors: Linzy Houchen-Wolloff\textsuperscript{1,2,3*}, Enya Daynes \textsuperscript{1,2,3*}, Amye Watt\textsuperscript{1,2}, Emma Chaplin\textsuperscript{1,2}, Nikki Gardiner\textsuperscript{1,2}, Sally Singh\textsuperscript{1,2,3}.

*Joint first author- equal contribution.

1. Cardiopulmonary Rehabilitation, Glenfield Hospital, Leicester, UK.
2. Centre for Exercise and Rehabilitation Science, Leicester Biomedical Research Centre, Leicester, UK.
3. Department of Respiratory Sciences, University of Leicester, Leicester, UK.

Abstract

Introduction: The COVID-19 pandemic has seen many cardiopulmonary rehabilitation services delivering programmes remotely. One area of concern is how to assess exercise capacity when a supervised exercise test is not possible. The aim of this review was to examine the relationship between functional exercise tests with recommended exercise tests for cardiopulmonary rehabilitation. Methods: Rapid narrative review. Searches were conducted by 2 authors.

Participants: Adults, all long-term conditions.

Intervention: Any/ none.

Outcome: Duke activity status index (DASI), Sit to stand (STS 30 second, 1 minute and 5 repetition), short physical performance battery (SPPB), 4 metre gait speed (4MGS) or step test (Chester/ others) AND directly compared to one of the recommended exercise tests for cardiopulmonary rehabilitation: six minute walk test (6MWT), incremental shuttle walk test (ISWT) or cardiopulmonary exercise test (CPET) in terms of reporting agreement/ correlation.

Study design: primary research only, controlled trials or observational studies.

Results: 16 articles (249 screened). N=2271 patients. Overall there was weak-strong correlations for the included tests with a recommended exercise test r=0.38- 0.85. There were few reported issues with feasibility or safety of the tests. However all tests were supervised in a clinical setting. The test that correlated highest with field walking test was the 4MGS with the ISWT (r=0.78) and with the 6MWT (r=0.85).

Discussion: The 4MGS correlates highest with routine measures of exercise tolerance. However it may be difficult to standardise in a remote assessment or prescribe exercise
Clinicians should strive for face-to-face standardised exercise tests where possible to be able to guide exercise prescription.

**Introduction**

Patients with chronic respiratory and cardiac disease benefit from centre-based rehabilitation in the form of exercise and disease management education. The benefits of this intervention have been proven in a range of outcome measures including exercise tolerance, quality of life, mood and mortality (1;2). Despite the known benefits, uptake to rehabilitation is poor (3;4) and this problem is not unique to the United Kingdom (5). To this end, recently we have seen the emergence of new programme formats that aim to improve access and uptake, including tele-rehabilitation and other home-based models. However the evidence base for remotely delivered rehabilitation is not secure particularly in older patients with Chronic Obstructive Pulmonary Disease (6). Patient experience with remotely delivered healthcare is variable and will only be relevant for those who can access and engage with the technology (7;8). Evidence for tele-rehabilitation in the cardiac rehabilitation population is more convincing (9); perhaps due to patient demographics. All international guidelines state that an exercise programme should be individually prescribed and progressed, predicated on a measure of exercise capacity. Otherwise the programme is not recognised as Pulmonary/ Cardiac Rehabilitation [PR/ CR (10;11)]. The recent COVID-19 pandemic and ongoing social distancing measures have therefore had a profound effect on how rehabilitation services are delivered around the world (12). Many centres have temporally suspended face to face activity and assessments have largely ceased; this has posed a challenge to rehabilitation providers, who have been unable to individually assess and prescribe the essential exercise component of the programme. Added to this, rehabilitation staff may have been redeployed to acute areas, COVID research teams, or to work from home due to shielding.

Centres have risen to this challenge and turned to providing remote programmes in the form of home-based manuals, phone calls and virtual classes [tele-rehabilitation (13)]. However there is worry from the rehabilitation community that the effectiveness of rehabilitation may be diluted (12), particularly if no measure of exercise tolerance is completed at baseline.

Many functional exercise tests lend themselves to being conducted virtually in the home with minimal equipment [e.g. various sit to stand (STS) tests: e.g. 1 minute (14), 4 metre gate speed (4MGS: (15), short physical performance battery (SPPB: (16) and various step tests, e.g. Chester Step Test (17)]. Furthermore, the Duke Activity Status Index (DASI) questionnaire has been suggested as a proxy for maximal exercise tolerance and yields a METS (metabolic equivalents) value (18) that can be converted to heart rate (HR) training zones. There is no specific national/ international guidance on how to conduct virtual assessments (including exercise testing) in patients with cardiopulmonary disease; though locally developed protocols have been implemented. This is despite anxieties around safety;
in particular, cardiovascular stability and oxygen desaturation when the assessment is not conducted in person. Greenhalgh and colleagues (19) have recently completed a rapid review of remote exercise testing for oxygen desaturation in COVID patients. They concluded that the 1 minute sit to stand (1MSTS) and the 40-step test, validated in other lung diseases, could be considered and a 3% decrease in pulse oximetry would be cause for concern. Furthermore, they argue that the test should not be used outside of a supervised healthcare visit if saturation was <96%. However, what has not been explored in a systematic way is the correlation between recommended measures of exercise tolerance in cardiopulmonary disease [(e.g. Cardiopulmonary exercise test: CPET, the 6 minute walk test: 6MWT (20) or incremental shuttle walk test: ISWT (21)] and these functional alternatives. Future research in the wake of COVID is likely to examine how reliable, practical and safe these functional measures are for virtual use and how responsive the measures to interventions such as rehabilitation. Prior to this we need to scope the literature for how valid these alternatives are for exercise prescription (i.e. how do they relate to a recommended exercise tests?).

Aim: The aim of this rapid narrative review was to examine the relationship between functional exercise tests (in terms of correlation and/ or agreement) with measures of exercise tolerance, in patients with long term conditions. Although our population of interest is those with cardiorespiratory disease, there may be useful data from other long-term conditions which could influence the review findings.

Methods: The review took place in May 2020. Searches were conducted by lead authors LHW and ED, with a third reviewer (SS) to resolve any discrepancies. The review was developed and reported using PRISMA methodology (22).

Inclusion criteria:

Participants: Adults (>18 years), all long-term conditions.

Intervention: Any/ none.

Outcome- functional tests that can be performed in the home: DASI, STS (30 second, 1 minute and 5 repetition), SPPB, 4MGS, step test (Chester/ others) AND directly compared to one of the recommended standardised tests for exercise prescription: 6MWT, ISWT or CPET in terms of reporting agreement or correlation (10-11). See previous references for description of the test procedures. Secondary outcomes of interest were: feasibility, safety, equipment, reliability/ repeatability, and sensitivity.

Study design: primary research only, controlled trials or observational studies.

Exclusion criteria:

Participants: Paediatrics, not long-term conditions (e.g. surgical procedures).
**Outcome-functional tests of interest:** one of the 5 tests of interest not reported OR one of the 5 tests reported but not directly compared with a recommended exercise test for agreement/correlation. Tests that require additional equipment not routinely found in the patients’ home will be excluded.

**Study design:** reviews, single case studies, abstracts.

**Search terms for rapid review:** This was not a systematic review and as such, search terms were limited to:

SPPB Or “Short physical performance Battery” OR 4MGS OR “Four meter gait speed” OR STS OR “sit to stand” OR “step test” OR DASI OR “Duke Activity Status Index”

We then searched through the hits to apply our inclusion and exclusion criteria. Databases searched were: Medline and the Allied and Complimentary Medicine Database (AMED). There were no limits set on the language or time period of publication.

**Data Extraction and Risk of Bias Assessment:** Data extraction was performed by all authors manually and transcribed to an excel spreadsheet. A brief risk of bias assessment was conducted at the same time using the COSMIN taxonomy: Measurement Properties of Outcome Measurement Instruments (23). This is specifically designed for clinician-reported and performance-based outcome measures. For 8 categories the scale is poor, fair, good or excellent. We have converted this to an average rating low/moderate or high risk of bias.

**Results**

Figure 1 shows the PRISMA flow diagram of screened articles.

- Articles identified by database screening
  - n=249

- Abstracts included for full text screening
  - n=38

- Articles included in review
  - n= 16

Duplicates (n= 33) were removed. Reasons for exclusion were as follows: not one of the 5 tests of interest (n=31), no validity outcome (i.e. correlation/agreement with routine
exercise tests (n=73), not a long-term condition (n=82), paediatrics (n=4), wrong study design/ publication type (n=10).

**Primary outcome:**

Table 1 shows the included article details and average risk of bias score.

<table>
<thead>
<tr>
<th>First author Year (reference)</th>
<th>Outcome measure</th>
<th>N=</th>
<th>Disease</th>
<th>Agreement/ correlation to gold standard</th>
<th>Risk of bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Briand 2018 (24)</td>
<td>1 minute STS</td>
<td>107</td>
<td>ILD</td>
<td>1minute STS and 6MWT r=0.5</td>
<td>Moderate</td>
</tr>
<tr>
<td>Carter 2002 (25)</td>
<td>DASI</td>
<td>119</td>
<td>COPD</td>
<td>DASI and 6MWT r=0.53.</td>
<td>High</td>
</tr>
<tr>
<td>Coute 2017 (26)</td>
<td>DASI</td>
<td>100</td>
<td>CV</td>
<td>Predicted METs from DASI and maximum exercise METs r=0.38.</td>
<td>High</td>
</tr>
<tr>
<td>Crook 2017 (27)</td>
<td>1 minute STS</td>
<td>255 (2 studies combined)</td>
<td>COPD</td>
<td>1 minute STS and r=0.59 study 1 at baseline and r=0.67 at follow-up. r=0.64 study 2 at baseline and r=0.68 at follow-up.</td>
<td>Moderate</td>
</tr>
<tr>
<td>Di Thommazo-Luporini 2015 (28)</td>
<td>Step Test (6 minute)</td>
<td>56</td>
<td>Obesity</td>
<td>CPET (treadmill test VO₂ peak) r=0.56.</td>
<td>High</td>
</tr>
<tr>
<td>Jones 2013 (29)</td>
<td>5STS</td>
<td>475</td>
<td>COPD</td>
<td>ICC with ISWT -0.59.</td>
<td>Low</td>
</tr>
<tr>
<td>Kon 2013 (30)</td>
<td>4MGS</td>
<td>586</td>
<td>COPD</td>
<td>4MGS and ISWT r=0.78, 4MGS and ISWT % predicted 0.72.</td>
<td>Moderate</td>
</tr>
<tr>
<td>Matkovic 2017 (31)</td>
<td>4MGS</td>
<td>111</td>
<td>COPD</td>
<td>4MGS and 6MWT r=0.85.</td>
<td>Moderate</td>
</tr>
<tr>
<td>Mori 2019 (32)</td>
<td>SPPB</td>
<td>53</td>
<td>Charcot-Marie-Tooth</td>
<td>ICC with 6MWT 0.35.</td>
<td>High</td>
</tr>
<tr>
<td>Meriem 2015 (33)</td>
<td>1 minute STS</td>
<td>49</td>
<td>COPD</td>
<td>1 minute STS and 6MWT r = 0.47.</td>
<td>Moderate</td>
</tr>
<tr>
<td>Nolan 2018 (34)</td>
<td>4MGS</td>
<td>46</td>
<td>IPF</td>
<td>4MGS and 6MWT r=0.76.</td>
<td>Low</td>
</tr>
<tr>
<td>Ozalevli 2007 (35)</td>
<td>1 minute STS</td>
<td>53</td>
<td>COPD</td>
<td>1 minute STS and 6MWT r = 0.75.</td>
<td>Moderate</td>
</tr>
<tr>
<td>Reed 2020 (36)</td>
<td>DASI and Chester Step test</td>
<td>50</td>
<td>CV</td>
<td>Step test and CPET (treadmill test VO₂ peak) r=0.69, DASI and CPET r=0.38.</td>
<td>Moderate</td>
</tr>
<tr>
<td>Reyschler</td>
<td>1 minute STS</td>
<td>42</td>
<td>COPD</td>
<td>1 minute STS and 6MWT r=0.75.</td>
<td>Moderate</td>
</tr>
</tbody>
</table>
2018 (37)  |  STS  |  6MWT r= 0.72.  
Wilkinson 2018 (38)  |  1 minute STS and 5STS  |  41  |  CKD  |  5STS and ISWT r=0.55 time point 1 and r=0.74 at 6 weeks. 1minute STS r not reported.  |  Moderate  
Zhang 2018 (39)  |  5STS and 30 second STS  |  128  |  COPD  |  5STS and 6MWT r= -0.51, 30 second STS and 6MWT r=0.53.  |  Moderate


**Secondary outcomes (where reported):**

**Feasibility:**

There were 12 studies that explored feasibility of outcome measures (24-26, 28-31, 34, 36-39). These studies included the DASI, step tests, 5STS, and 4MGS. Reported completion of the tests ranged from 85-100%. The DASI was reported to be completed by 100% of participants in the studies. One study reported that 7% of patients failed to complete the 30STS (39). All studies explored the test in a hospital or outpatient setting under supervision and therefore they may not be feasible to perform in a home-based environment.

**Safety/ equipment:**

The equipment used in the listed tests includes a step, chair, stopwatch, 4 meter track and in some instances a heart rate/ oxygen saturation monitor. The step and chair heights, where described would be a standardised height between patients and between testing time points (e.g. pre and post intervention). Two studies stated that the haemodynamic stress was lower for sit to stand tests when compared to a 6MWT (33, 35). None of the included studies reported any serious adverse events when performing the tests. However the tests were performed in clinical settings under supervision.

**Reliability/ repeatability:**

7 studies assessed the reliability and repeatability of the tests (27, 29-30, 34, 37-39]). Inter-rater reliability was the most commonly reported measure using intraclass correlation coefficient (ICC). These were all moderate, good or excellent, range 0.68 [5STS (38)] and 0.99 [4MGS (30, 34), STS (27, 29)]. Test re-test ICCs were good where reported 0.93-0.97 (27, 29-30, 34). 2 studies reported limits of agreement between 2 testing sessions for the 4MGS: these were 0.09 to -0.09 (34) and 0.135 to -0.125 (30). One study examined the limits of agreement for the peak oxygen uptake (VO$_2$) predicted from the DASI questionnaire and Chester step test with the CPET (36). Both tests overestimated VO$_2$ when compared to the CPET: DASI by 3.6 millilitres per kilogram per minute (ml/kg/min) and the step test by 4.1 ml/kg/min. One study compared nadir oxygen saturation (SpO$_2$) in the 1minute STS test compared to the 6MWT: the agreement ICC was 0.77 (24).
Sensitivity:

Many of the studies were single time point/ not testing response to an intervention. However, the sensitivity of the tests was explored in 6 studies, generally following exercise-based rehabilitation (27, 29, 34, 36, 38-39). 1 study suggested an MCID for the 5STS in patients with COPD: 1.3-1.7 seconds (29) and the MID for this outcome was reported as 7.5 seconds in patients with CKD (38). For the 1 minute STS, one study found an MID of 3 reps in COPD (27) and 4 reps in CKD (38). Zhang et al. reported ROC curves and showed that the area under the curve (AUC) of the 5STS score in predicting poor 6 minute walk test distance (6MWD of <350 m) was 0.731, whereas the AUC of the 30STS score was 0.724, indicating that both tests are useful tools in predicting 6MWT performance (39). The 4MGS improved significantly following PR in patients with IPF [mean (95% CI) change: 0.16 (0.12–0.20) m/s), effect size 0.65 (34)]. In the cardiovascular study comparing DASI and Chester step test to the CPET, the change in DASI was 2.3±5 and for Chester step test the change was 1.9±5 ml/kg/min (36).

Discussion

This rapid review included 16 studies (of 2271 patients) to assess functional exercise tests compared to exercise tests in order to guide exercise prescription in patients with long term conditions. 13/16 studies were conducted in our population of interest (cardiopulmonary). This review was intended as a rapid review to guide practice in terms of exercise prescription. Searches were performed systematically and inclusion was assessed independently by two authors (ED/ LHW). However, search terms were limited and there were a large number of texts that were not included as they did not compare the agreement/correlation between the test of interest to a standard test, despite conducting a standard test in their study. Given the time scale of the rapid review authors were not contacted for additional data. This may affect the review conclusions. Since this review one study assessed the 1 minute STS with the 6MWT (r=0.82) and CPET (r=0.71), however this was in a small sample (n=15) of ILD patients with a high risk of bias (40). Therefore, this would not change the conclusion of this review. Furthermore, a rapid risk of bias assessment was performed by the authors; most studies (10 out of 16) were at moderate risk of bias which again may affect our findings.

It is imperative that cardiopulmonary rehabilitation is reinstated to a high standard as quickly as possible due to the importance of the intervention and the effects that shielding and isolation have on deconditioning (41). Overall, there was weak to strong correlations for the included tests (STS tests, step tests, DASI, 4MGS, SPPB) with recommended exercise tests (CPET, ISWT, 6MWT), r=0.38-0.85. The tests generally had good-excellent reliability and repeatability. The test that correlated highest with exercise tests was the 4MGS with the ISWT [r=0.78 (30)] and with the 6MWT [r=0.85 (31)]. This test was also shown to be sensitive to change following exercise rehabilitation in IPF [medium effect size of 0.65 (34)]. The study by Kon et al. reported 90% feasibility to perform the 4MGS in a home environment.
though this was taken from a previous study (42) and not their own which was performed in a hospital clinic (30). There were few reported issues with feasibility or safety in the included tests, only 1 study (39) reported that 7% of patients failed to complete the 30STS presumably due to fatigue. However all tests included in this review were conducted in a supervised setting. Therefore the feasibility and safety of these measures in the home warrant further consideration.

Risk assessment should be performed prior to virtual assessment via a subjective assessment, medical notes screening and/or a risk assessment proforma to identify patients who are safe to perform these tasks at home (43,44) and usual contraindications to exercise testing will remain (10;45) alongside local Standard Operating Procedures (SOPs). Considerations should be given to patients that have cardiovascular instability and exercise induced desaturation. Ideally patient's heart rate response and oxygen saturations would be measured during exercise testing and in the recovery period, this has implications for equipment provision if intended to be performed in the home. Monitoring of symptoms (i.e Borg breathlessness and perceived exertion scales) may help to guide exercise prescription, alongside the rise in wearable technology which would facilitate increased monitoring at home. Other equipment required is an important consideration and if performing tests at home it will be difficult to standardise the test due to chair or step height discrepancies, course length etc. Though within individual comparisons would be possible if the same height chair/ step/ length was used pre and post the intervention. Safety to avoids trips and falls is of particular concern with steps.

Virtual assessment also relies on technology provision and digital literacy which is known to be poor among patients with respiratory disease (8) and therefore non-technology based alternatives are important. Provision of devices and training may be required for patients prior to assessment. All studies included in this review performed the test in a hospital environment and therefore we do not know the implications of performing these tests at home. Furthermore, there will be local decisions that need to be made about the best virtual platform depending on what is considered safe in terms of data security and confidentiality.

This review excluded studies that required additional equipment; however there is evidence for the use of the 6 minute stepper test in respiratory disease. These studies explore feasibility, comparison to 6MWT and sensitivity to PR (46). We excluded 2 studies in cystic fibrosis [CF (47-48)]. Though these studies did compare the 1 minute STS to a standard (CPET), they looked at the relationship between cardiorespiratory variables and not the actual test results (e.g Watts vs STS strokes). Of note the agreement in VO₂ peak in these studies derived from STS and CPET was between r=0.59-0.98. This review included studies that explored long term conditions and as a result, a number of studies exploring the DASI in pre/ post-surgical patients were excluded. The DASI has previously been used in CR to prescribe target heart rate for exercise training and has been found to provide broadly
equivalent MET levels to a CPET (45). This may be of value given that it is a functional questionnaire and carries minimal safety and feasibility implications. Though the 4MGS performed the best when compared to recommended exercise tests, this does raise issues in how to prescribe exercise from this test as you can from the ISWT and 6MWT (44). Rehabilitation should include prescribed exercise and if it is not possible to derive exercise prescription, standard rehabilitation has not been delivered. This will have clinical implications for the patients, for service provision and for national audit reporting.

Based on the results of this review the 4MGS had the best correlation with exercise tests, however it may be difficult to standardise in a home environment/ during a virtual assessment. These findings are surprising given that the 4MGS is not a maximal test and patients can walk at their own speed. There is scope to perform the subjective assessment at home and an exercise test in a clinical or community setting. Practical considerations for infection prevention and how this might impact on a walking test [such as wearing a mask may impact validity; though there is some evidence that there is minimal alteration in peak ventilation when a mask is worn (50-51)]. With a potential second wave predicted for autumn/winter 2020 it is important that we futureproof rehabilitation to enable assessments and programmes without delay. Clinicians should strive for face to face services with standardised exercise tests where possible to be able to guide exercise prescription. However, the current situation provides an opportunity for services to evaluate the efficacy of providing interventions delivered in the home/virtually. Future research should examine how feasible, reliable and safe these functional tests are for virtual use and how responsive they are to rehabilitation delivered virtually.

**Funding:** nil.

**Trial registration:** n/a- review.

**Conflicts of interest:** nil.

**Reference List**


(10) BACPR. The BACPR Standards and Core Components for Cardiovascular Disease Prevention and Rehabilitation. 2017.


(47) Gruet N, Peyre-Tartaruga L, Mely L, Valiier J. The 1-Minute Sit-to-Stand Test in Adults With Cystic Fibrosis: Correlations With Cardiopulmonary Exercise Test, 6-Minute Walk Test, and Quadriceps Strength. Resp Care 2016 Dec;61(12):1620-1628.


