Systematic Review Article

Title: A systematic review of the six-minute walk test in outpatient cardiac rehabilitation: validity, reliability, and responsiveness.

Authors: R. Nicole Bellet, BPhty, MPH\textsuperscript{1,2,3}, Lewis Adams BSc PhD\textsuperscript{2,3}, Norman R. Morris BAppSc Physio, PhD\textsuperscript{2,3}.

\textsuperscript{1} Physiotherapy Department, The Prince Charles Hospital, Metro North Health Service District, Brisbane, Queensland, Australia
\textsuperscript{2} School of Physiotherapy and Exercise Science, Griffith University, Gold Coast, Australia
\textsuperscript{3} Griffith Health Institute, Griffith University, Gold Coast, Australia

Key words: six-minute walk test, cardiac rehabilitation, reliability, validity, responsiveness

Word Count: 250 words (Structured Abstract) 3048 words (Text)

References: 36

Figures: 2

Box: 2

Tables: 4

Source(s) of support: Grant from Queensland Health Community Rehabilitation Workforce Project: Community Rehabilitation Research Scheme: letter dated 24/6/07.

Acknowledgements: We would like to acknowledge Leanne Bisset MPhty PhD for her expert advice in the review of the manuscript and Ian Yang MBBS (Hons), PhD, FRACP, Grad Dip Clin Epid for his expert advice and assistance with the meta-analysis.

Competing interests: No commercial party having a direct financial interest in the results of the research supporting this article has or will confer a benefit on the authors or on any organization with which the authors are associated.
Title Page

Systematic Review Article

Title: A systematic review of the six-minute walk test in outpatient cardiac rehabilitation: validity, reliability, and responsiveness.

Key words: six-minute walk test, cardiac rehabilitation, reliability, validity, responsiveness

Word Count: 250 words (Structured Abstract)
3048 words (Text)

References: 36

Figures: 2

Box: 2

Tables: 4
Structured Abstract (250 words)

**Background:** The six-minute-walk-test (6MWT) is a common outcome measurement in cardiac rehabilitation (CR); however, a search of the literature found no guidelines established for the use of the 6MWT in CR.

**Objectives:** Systematically review the validity, reliability, and responsiveness of the 6MWT in CR.

**Data Sources:** OvidMEDLINE, SPORTdiscus, EMBASE, CINAHL, Cochrane Reviews, Cochrane Clinical Trials between January 1948 and April 2011.

**Eligibility Criteria:** Studies using 6MWTs in subjects with coronary artery disease undergoing outpatient-CR were included. Non-English-language articles were excluded.

**Study appraisal and methods:** Quantitative and qualitative analyses were conducted including, methodology quality assessment, meta-analysis and assessment against level of evidence criteria.

**Results:** Fifteen articles met inclusion criteria. One high quality study was identified for reliability, six for validity and eleven for responsiveness. Meta-analysis indicated strong evidence that the 6MWT was responsive to change in clinical status following CR with an estimated mean difference (95% CI, p-value) in 6MWD of 60.43m (54.57, 66.30, <0.00001).

Qualitative analysis indicated moderate evidence for the repeatability of the 6MWT in the CR population, for a 2-8% learning effect between repeated-6MWTs, for a relationship between peak heart rate during 6MWT and during cycle-exercise at ventilatory threshold and for a moderate to high correlation between the 6MWD and maximum-metabolic-equivalents achieved on symptom-limited-exercise-test.

**Limitations:** Few studies assessed similar aspects of validity for the 6MWT.
**Conclusion:** Strong evidence suggests that the 6MWT is responsive to clinical change during CR outcome measurement. Intra- and inter-tester reliability of the 6MWT and its validity in the CR population requires more research.

**Background**

Measurement of oxygen consumption, during cardiopulmonary exercise testing, represents the gold standard for determining baseline functional capacity, training intensity, cardiovascular risk, and for the evaluation of training outcomes in the cardiac rehabilitation (CR) population (1). However, the six-minute-walk-test (6MWT), is frequently recommended (2) (3) (4) to estimate functional exercise capacity in CR rather than subjecting patients to an exercise stress test. The primary outcome of the 6MWT, the 6-minute walk distance (6MWD), is used both to prescribe exercise training intensity (5) and as an outcome measure for CR (6).

Despite common usage of the 6MWT in CR, guidelines for the use of the 6MWT in this population were not found in a search of PubMed and CINAHL. Instead CR researchers utilize guidelines developed for pulmonary patients (6) (7) (8) (9) (10) (11) (12) (13, 14). It is unknown if the 6MWT is valid and reliable in the CR population and what changes in 6MWD would be expected following CR intervention. Repeated 6MWTs are recommended in the pulmonary (15, 16) (17) (18) and chronic heart failure populations (19) (20), although, a recent study suggests repeated 6MWTs are unnecessary in patients with heart failure who walk less than 300m on an initial 6MWT (21).

Therefore, we conducted a systematic review to examine the validity, reliability and responsiveness of the 6MWT in the outpatient CR population.

Data Sources:
The first author conducted searches of OvidMEDLINE (January 1948 to March 2011), CINAHL and SPORTdiscus (January 1997 to April 2011), EMBASE (January 1980 to April 2011), Cochrane Reviews (current), and Cochrane Clinical Trials (current) databases using the search terms in Box 1. The 'date of publication' limitation varied between the databases bases upon library access availability.

Eligibility Criteria:
Trials using 6MWTs in subjects with coronary artery disease undergoing outpatient-CR were included. We included clinical trials and observational studies that described repeated 6MWTs; that compared 6MWD to established reference tests; and that examined 6MWD pre- and post-CR. We excluded trials not available as full-text articles, not in the English language, and in populations other than outpatient-CR.

Study Appraisal and synthesis methods:
The first author screened titles and abstracts of the identified articles for duplicates, and adherence to inclusion and exclusion criteria. The reference lists of the included articles were scanned for potentially relevant studies. The first author extracted and tabulated the data from the included articles under the categories of reliability, validity, and responsiveness and the third author confirmed this process.

To assess the quality of the extracted articles for reliability, validity and responsiveness, Brink and Louw’s thirteen questions (22), were combined with additional criteria on responsiveness and reliability developed by Jerosch-Herold (23) and May et al (24) (25), capitalising on the strengths of each tool. Questions were adapted to ensure a yes/no response(23) (24). Studies

scoring over 60% positive responses for methodology criteria relevant to the study type were considered high in quality (24) while studies scoring less than 40% were assessed as low quality. In scoring the quality of the reliability studies, we also omitted the criteria for intra- and inter-tester reliability, as the methodology design of both reliability studies, did not include these comparisons. The first and third authors, blinded to one another, undertook assessment of each article against these criteria as shown in Table 1. Differences in opinion between the two assessors were resolved through discussion and consensus.

Qualitative analysis was based upon the established level of evidence criteria in Box 2 (24) (26) and quantitative analysis was performed using Review Manager (RevMan)\textsuperscript{a}. To enable meta-analysis of 6MWD responsiveness, the data was standardised by converting scores to mean differences and standard errors. Data from subject subgroups were included in the meta-analysis where whole of cohort figures were not available. To further explore responsiveness of the 6MWD in CR, we calculated percentage change and the effect size (27) for the subjects (and sub-groups) of the studies based on the mean and standard deviations for 6MWDs provided for the groups. We recognise that the percentage change calculated is a grouped measure and may not accurately reflect the true percentage change for the cohorts. Further, we examined the data for evidence of the ability of the 6MWD to discriminate between subjects based upon physiological factors.

**Results:**

The search yielded 175 acceptable articles. Figure 1 outlines the flow of article selection for analysis and the reasons for exclusions. The size of study cohorts varied, with large subject numbers in retrospective studies (6) (28) (29) (30) and smaller numbers in prospective trials (7).
One article described a prospective study and made comparisons retrospectively with data from patient files (10).

Table 1 shows the quality appraisal of the included articles. Some studies were assessed under more than one category, i.e. reliability, validity or responsiveness, and the methodology quality of these studies varied depending upon assessment against the appraisal criteria required for each study type. One high quality study was identified for reliability, six for validity and eleven for responsiveness.

**Reliability:**
Qualitative analysis of the reliability data was conducted as only two studies were identified, precluding quantitative analysis. These two studies, shown in Table 2, examined the repeatability of the 6MWT in CR (7) (10) without addressing either intra-tester or inter-tester reliability (23). One study (10), assessed as high in quality, demonstrated a strong test-retest reliability (Intraclass correlation = 0.97) between repeated 6MWDs however, a statistical difference was found for repeated tests between 6MWD1 and 6MWD2 (7, 10) and between 6MWD2 and 6MWD3 (10). These differences were 3% to 8% between the 6MWD1 and 6MWD2 (7) (10) and of 2% between 6MWD 2 and 6MWD3 (10).

**Validity:**
Qualitative analysis of the validity data was conducted as the variety of reference tests used for comparison with the 6MWT prevented quantitative analysis. Eight articles assessed criterion-related validity of the 6MWT; however, the predictive, diagnostic and prognostic validity of this test was not evaluated in any of the studies identified. While only two studies indicated intent to assess the validity of the 6MWT, or clinical observations during the 6MWT (10) (12), six articles

used the 6MWT as part of patient assessment in CR (6) (7) (8) (9) (11) (31). Studies compared the 6MWT with a range of reference tests, the most common of which were measurements taken during a symptom-limited-exercise-test (7) (8) (9) (10) (11) (12) and the physical function subscales of quality-of-life-questionnaires (6) (10, 31). Pearson’s correlation coefficient was reported in many of these studies; however the 95% confidence limits for these coefficients were not reported.

Table 3 shows the details and limitations of the study designs with the comparisons and associated statistical analyses performed. The correlation of the 6MWD to maximum power, oxygen uptake and maximum metabolic equivalents during symptom-limited-exercise-test was moderate to high (r=.56 to .93) in the four articles reporting this data. The highest correlation was reported in a study we assessed to be of moderate quality (11) examining maximum power. The second highest correlation (r= 0.69, p<0.001) was reported in a high quality study (10) which found a linear relationship between maximum metabolic equivalents achieved at a symptom-limited-exercise-test (cycle or treadmill). Both these studies used the best of repeated 6MWDs, the former the best of four (11) and the latter the best of three (10). All other studies reported on comparisons with a single 6MWD.

Two trials examined peak heart rate in symptom-limited-exercise-tests and in 6MWTs, one moderate quality study reported a moderate correlation (r=.64, p <.009) (11) and the other, assessed as high in quality (8) reported that 6MWT peak heart rate was 78%, SD=6%, of peak heart rate during symptom-limited-exercise-tests. One high (8) and one moderate (7) quality study reported a non-significant difference for heart rate assessed at ventilatory threshold and during the 6MWT. The latter study also reported a correlation of r=.56 between peak oxygen uptake during symptom-limited-exercise-tests and the 6MWD (7). Heart rate recovery was found

to be similar following symptom-limited-exercise-tests and 6MWTs but this report failed to perform correlation analysis (12). A moderate quality study (11) reported a moderate correlation (r=.52, p < .038) between blood pressure responses during symptom-limited-exercise-tests and 6MWTs.

Three moderate to high quality studies used quality-of-life-questionnaires as reference tests for validation of the 6MWT. Two studies reported moderate correlations (r=.54 to .62, p<.001) (31) (10) while the third study found no relationships with health and function domains (6).

**Responsiveness:**

The ability of the 6MWT to detect change in clinical status over time and difference between subject variables was explored in 12 studies. Four studies compared different cohorts based upon sex, age, left ventricular ejection fraction and initial 6MWD (5) (28) (30) (33). One study reported an effect size for pre-CR to post-CR (6). Eleven studies reported data for pre-CR, post-CR 6MWDs and standard deviations (SD) or mean difference and SDs for inclusion in the quantitative analysis (see Table 4).

We performed meta-analysis of the data for 6MWD responsiveness using random effects analysis due to a high heterogeneity (60%) in the data. Figure 2 reports the grouped estimate of the mean difference in 6MWD pre-CR to post-CR of 60.43m, with 95% confidence intervals of 54.57m and 66.30m, p<.00001. Table 4 also shows the values we calculated for effect sizes and percentage change for the cohorts in italics. Effect sizes ranged from 0.40 to 4.28 (mean = .96; median =0.65) and the percentage change post-CR ranged from 10% to 28%. Pre-CR and post-CR 6MWD ranged from 301 to 489m and 377 to 554.5m, respectively.

Discussion:

The 6MWT is a common outcome measure in CR, however, to our knowledge, we are the first to systematically review the reliability, validity and responsiveness of the 6MWT in this population. Utilizing meta-analysis, we found strong evidence that the 6MWT is responsive to change in clinical status following CR and report a grouped estimate of the mean difference in 6MWD following CR of 60.43m with 95% confidence interval of 54.57 to 66.30m. We found moderate qualitative evidence for the repeatability of the 6MWT in the CR population, with moderate to high correlation between the 6MWD and maximum metabolic equivalents achieved during symptom-limited-exercise-test and for a relationship between peak heart rate during the 6MWT and at ventilatory threshold during cycle-exercise. Further, there is moderate evidence that a learning effect occurs between repeated-6MWTs.

Reliability:

Examining reliability, we found moderate evidence for the repeatability of the 6MWT in the CR population (2 trials, 103 subjects), demonstrating a 2% to 8% change with test-retest with a .97 ICC. Repeatability was assessed with 2-repeated (7), or 3-repeated (10) 6MWTs, however, the timing of the repeated 6MWTs was not stated clearly in either study, and may have been separated several days (7, 10). We found no evidence for inter-tester or intra-tester reliability of this test in CR (zero trials, zero subjects).

Validity:

Assessing the evidence for validity, we found moderate evidence of a moderate to high correlation between the 6MWD and maximum-metabolic-equivalents achieved on symptom-limited-exercise-test (2 trials, 175 subjects) and moderate evidence for a relationship between peak heart rate during 6MWT and at ventilatory threshold during cycle-exercise.

peak heart rate responses during 6MWTs and during exercise-tests at ventilatory threshold (2 trials, 55 subjects). We found limited evidence for comparisons with all other aspects of the exercise-tests, including, peak heart rate responses during symptom-limited-exercise-tests and 6MWT (2 trials, 40 subjects), 6MWD and maximum power on symptom-limited-exercise-test (1 trial, 10 subjects); blood pressure responses in 6MWTs and those in symptom-limited-exercise-tests (1 trial, 10 subjects); 6MWD and oxygen uptake during 6MWT and during symptom-limited-exercise-tests (1 trial, 25 subjects).

We found conflicting evidence for the relationship between 6MWD and quality-of-life-meaures with moderate correlations reported in two trials (113 subjects) and no relationship found in another trial (630 subjects). The difference in results could relate to differences in subject numbers, the timing of comparisons made or to the tools used. Verrill (6) compared pre- and post-CR quality-of-life-meaures and 6MWD while one study compared Short-Form-36 to pre-CR 6MWD alone (10) and the other study, does not describe the timing of the comparison (31). The studies reporting moderate correlations used the Short-Form-36-Health-Survey-physical-score (31) (10) while the other used subscales of the Ferrans-and-Powers'-Quality-of-Life-Index-Cardiac-Version-III (6). Others have reported greater body weight loss and improvement in Short-Form-36-Health-Survey-physical-score in subjects with a 23% or greater increase in 6MWD post-CR when compared to those with 12% or less improvement (28). However, while this could be a factor the two studies reporting 6MWD changes both demonstrated a 15% improvement (31) (6).

**Responsiveness:**

Strong evidence of the ability of the 6MWT to demonstrate change in clinical status was indicated. We report a mean difference in 6MWD following CR intervention (12 trials, 2487

subjects) of 60.43m with 95% confidence interval of 54.57 to 66.30m and a percentage or relative increase of 10% to 28% (3) (5) (6) (8) (12) (28) (29) (30) (31) (32) (33). Importantly, this relative increase was 1 to 14 times greater than the increase that would be expected by test-retest familiarization alone (7) (10). The mean difference in 6MWD reported in our analysis is higher than the minimal-clinically-important-difference of 25 to 27m reported for 6MWD in subjects with coronary artery disease (9) and subjects with chronic obstructive airways disease (33) and may reflect a good response to CR training. To our knowledge, we are the first to report 6MWDs pre-CR and post-CR across a number of studies, providing a range of expected pre-CR, post-CR 6MWDs for comparative purposes. Moreover, we have calculated effect sizes to inform relative comparisons across programs. The largest effect size of 4.28 was reported in a subgroup (28) demonstrating a narrow standard deviation compared to other groups (see Table 4 and Figure 2).

We explored the high heterogeneity of the data reported in our review and found that when data from two studies (9) (28) and Maniar’s aged-less-than-65-years subgroup (30) were removed from the analysis the heterogeneity reduced to zero. These studies demonstrated a high mean difference in 6MWD pre-CR to post-CR (69.3 to 73.9m) and it is possible that differences in exercise dose between studies are the cause of this heterogeneity. Two of these studies based training prescription on symptom-limited–exercise–tests and had high exercise doses, one reporting 1.5 hours (9) training sessions and another of up to 12 months training (28), however, Gremeaux’s earlier article (8) reports on 2-hour training sessions and removing this earlier article from the analysis increased heterogeneity. Another contributing factor could be the potential to improve. Maniar’s aged-less-than-65-years subgroup (30) had higher BMI than all other study cohorts and poorer waist circumference, diastolic blood pressure, glucose control, and lipid profiles than Maniar’s aged-greater-than-65-years subgroup. It is possible that exercise Bellet RN, et al. The 6-minute walk test in outpatient cardiac rehabilitation: validity, reliability and responsiveness—a systematic review. Physiotherapy (2012), doi:10.1016/j.physio.2011.11.003
training in this subgroup had greater ability to improve 6MWD than in other groups. It is interesting however, that removing the Tallaj’s subgroup with left-ventricular-ejection-fraction-less-than-40% (5), this being the group that demonstrated the greatest improvement in 6MWD, reduced the mean 6MWD difference by approximately 1m without affecting the 60% heterogeneity. This could be due to the small numbers in this subgroup. Therefore, the reasons for the heterogeneity reported remain inconclusive.

We found limited evidence for the ability of 6MWD to discriminate between subjects in CR based upon sex (1 trial, 630 subjects) and age of the subjects (1 trial, 685 subjects). Further, sex and age (29) were reported as independent predictors of change in 6MWD following CR in a multivariate linear regression analysis. These findings are consistent with reported effects of sex, weight, height and age on 6MWD in linear regression models for healthy adults (34) (35). One study (29) (156 subjects, 49 women) reported symptoms of dyspnoea and angina were more common in women than men (48% versus 24%, p = 0.007).

Limited evidence was found to support a relationship between 6MWD improvement post-CR and initial 6MWD (1 trial, 425 subjects) or left-ventricular-function (1 trial, 179 subjects). Our review excluded articles that examined heart failure subjects specifically; however, one study we included grouped CR subjects based upon degree of left-ventricular-ejection-fraction and examined the differences between these groups (5). Subjects with left-ventricular-ejection-fraction-less-than-40% increased 6MWD post-CR by 74m (26%) and had an initial 6MWD similar to another subgroup aged-75-years-or-older (286m,SD=88.7m and 283, SD=88m, respectively) (30). It could be that subjects recording low initial 6MWDs have greater room for improvement than those walking greater distances and that a “ceiling effect” may occur for 6MWD in those CR subjects who have greater initial cardiovascular fitness (28).

Limitations:

Although an extensive search of the databases was conducted it could be that by refining our search question to ensure capture of studies in the outpatient CR population and to exclude those in other populations, that we missed relevant articles. Further, in reviewing the the titles in the reference lists of retrieved full-text articles for the terms ‘6MWT’, ‘6MWD’ and ‘outpatient CR’ we may have excluded relevant articles. We did not contact authors to seek missing data or to request access to unpublished studies. In addition, our adapted quality assessment tool for reliability, validity and responsiveness may not be valid and could have led us to misrepresent the quality of the articles assessed.

A major limitation of the studies reviewed was that many examined validity using cycle-exercise-tests as the gold standard reference test. An improved capacity in walking may not result in an improved capacity in cycling due to specificity of training (36). The use of a treadmill-exercise-test as a reference test may have improved the comparisons with the 6MWT.

In addition, the studies did not consistently note the elapsed time between the reference tests and the 6MWT, the time from cardiac event to assessment, and left-ventricular-ejection-fraction for subjects. While we excluded studies in the heart failure population, we acknowledge that many subjects within the included studies will have had a degree of heart failure and the presence of heart failure may have affected subjects’ response to exercise training and 6MWD.
Conclusion and implications of key findings

The implication of our review is that the 6MWT has been shown to be suitable for pre- and post-CR outcome assessment, despite a learning effect of 2-8% with repeated tests. To allow program effectiveness comparisons we have reported an estimated change in 6MWD following CR of 60.4m with a .65 median effect size.

The evidence for validity against symptom-limited and ventilatory threshold exercise-tests and against quality-of-life-measurement remains inconclusive. We recommend future research examines intra-rater and inter-tester reliability of the 6MWT, further assesses the validity of the 6MWT against treadmill-exercise-tests and explores the prognostic and predictive value of 6MWDs in the CR population.
Suppliers

References:


Box 1: Search question
(six min* walk or six-min* walk or 6 min* walk or 6-min* walk or 6MW* or 6-MW*) and (cardiac rehabilitation or coronary artery disease or coronary heart disease or heart valve) and (valid* or reliab* or distance or value or predict*) not (heart Failure) not (lupus) not (multiple sclerosis) not (chronic obstructive pulmonary disease or COPD or respiratory disease).

Box 2: Levels of evidence

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strong</td>
<td>Consistent findings from three or more high-quality studies</td>
</tr>
<tr>
<td>Moderate</td>
<td>Consistent findings from at least one high-quality study and a number of low-quality studies</td>
</tr>
<tr>
<td>Limited studies</td>
<td>Consistent findings in one or more low-quality studies</td>
</tr>
<tr>
<td>Conflicting</td>
<td>Inconsistent findings irrespective of study quality</td>
</tr>
<tr>
<td>No evidence</td>
<td>No studies found</td>
</tr>
</tbody>
</table>
Potentially relevant citations identified:  N = 175

35 Duplicates

124 abstracts excluded:
85 - Not Outpatient CR population
36 - 6MWT not main outcome measure
2 - Not in English language

16 full manuscripts retrieved for detailed evaluation

3 articles included from reference lists

4 articles excluded: Not outpatient-CR population

15 full and relevant articles included in review

Figure 2: Forest plot of the mean difference in reported 6MWD pre-CR to post-CR
Table 1: Critical appraisal of articles for Reliability, Validity and Responsiveness

<table>
<thead>
<tr>
<th>Questions</th>
<th>1</th>
<th>2</th>
<th>3a</th>
<th>4b</th>
<th>5b</th>
<th>6b</th>
<th>7a</th>
<th>8b</th>
<th>9b</th>
<th>10</th>
<th>11b</th>
<th>12</th>
<th>13</th>
<th>14c</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>% Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reliability Studies scored out of 10</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gayda 2004</td>
<td>Y</td>
<td>N</td>
<td>n/a</td>
<td>N/A</td>
<td>N/A</td>
<td>N</td>
<td>n/a</td>
<td>N</td>
<td>n/a</td>
<td>N</td>
<td>Y</td>
<td>n/a</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>n/a</td>
<td>N</td>
</tr>
<tr>
<td>Hamilton 2000</td>
<td>Y</td>
<td>N</td>
<td>n/a</td>
<td>N/A</td>
<td>N/A</td>
<td>N</td>
<td>n/a</td>
<td>N</td>
<td>n/a</td>
<td>Y</td>
<td>N/a</td>
<td>Y</td>
<td>N/a</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>n/a</td>
</tr>
<tr>
<td><strong>Validity Studies scored out of 12</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ades 2003</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>Y</td>
<td>n/a</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>n/a</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Gayda 2004</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>N</td>
<td>n/a</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>n/a</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>58</td>
</tr>
<tr>
<td>Gremeaux 2011</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>Y</td>
<td>n/a</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>n/a</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>100</td>
</tr>
<tr>
<td>Gremeaux 2009</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>Y</td>
<td>n/a</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>a</td>
<td>N</td>
<td>n/a</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>67</td>
</tr>
<tr>
<td>Hamilton 2000</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>N</td>
<td>n/a</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>n/a</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>83</td>
</tr>
<tr>
<td>Kristjánsson 2000</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>Y</td>
<td>n/a</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>a</td>
<td>N</td>
<td>n/a</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Roberts 2006</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>Y</td>
<td>n/a</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>n/a</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>92</td>
</tr>
<tr>
<td>Verrill 2003</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>n/a</td>
<td>n/a</td>
<td>N</td>
<td>n/a</td>
<td>Y</td>
<td>n/a</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>n/a</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td><strong>Responsiveness Studies scored out of 9</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ades 2003</td>
<td>Y</td>
<td>N</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>Y</td>
<td>n/a</td>
<td>N</td>
<td>n/a</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Araya-Ramirez 2010</td>
<td>Y</td>
<td>N</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>Y</td>
<td>n/a</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Bittner 2000</td>
<td>Y</td>
<td>N</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>Y</td>
<td>n/a</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Gremeaux 2011</td>
<td>Y</td>
<td>Y</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>Y</td>
<td>n/a</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Gremeaux 2009</td>
<td>Y</td>
<td>N</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>Y</td>
<td>n/a</td>
<td>a</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Hung 2004</td>
<td>Y</td>
<td>N</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>N</td>
<td>n/a</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Maniar 2009</td>
<td>Y</td>
<td>Y</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>N</td>
<td>n/a</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Rejeski 2002</td>
<td>Y</td>
<td>N</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>Y</td>
<td>n/a</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Roberts 2006</td>
<td>Y</td>
<td>Y</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>Y</td>
<td>n/a</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Tallaj 2001</td>
<td>Y</td>
<td>N</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>Y</td>
<td>n/a</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Verrill 2003</td>
<td>Y</td>
<td>N</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>Y</td>
<td>n/a</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Wright 2001</td>
<td>Y</td>
<td>Y</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>Y</td>
<td>n/a</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
</tbody>
</table>

No, N; Yes, Y, not applicable, n/a; grey shaded n/a, criteria not applicable for the study type; grey shaded N/A, inter- and intra-tester criteria not applicable to the study design

a Items relevant to validity; b Items relevant to reliability; c Items relevant to responsiveness; Items unmarked were relevant to all study types.

Questions (adapted from Brink and Louw (22) Jerosch-Herold (23) and May et al (24, 25): 1, If human subjects were used, did the authors give a detailed description of the sample of subjects used to perform the (index) test?; 2, Did the authors clarify the qualification, or competence of the rater(s) who performed the (index) test?; 3, Was the reference standard explained?; 4, If interrater reliability was tested, were Bellet RN, et al. The 6-minute walk test in outpatient cardiac rehabilitation: validity, reliability and responsiveness—a systematic review. Physiotherapy (2012), doi:10.1016/j.physio.2011.11.003

26
raters blinded to the findings of other raters?; 5, If intrarater reliability was tested, were raters blinded to their own prior findings of the test under evaluation?; 6, Was the order of examination varied?; 7, If human subjects were used, was the time period between the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests?; 8, Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time interval between repeated measures?; 9, Was the reference standard independent of the index test?; 10, Was the execution of the (index) test described in sufficient detail to permit replication of the test?; 11, Was the execution of the reference standard described in sufficient detail to permit its replication?; 12, Were withdrawals from the study explained?; 13, Were the statistical methods appropriate for the purpose of the study? 14, Does the instrument capture clinical change?; 15, Were subjects selected randomly or consecutively?; 16, Were the number of subjects >50 or was a sample size calculation provided?; 17, Did subjects give consent prior to testing?
Table 2: Reliability of six-minute walk test prior to participation in cardiac rehabilitation

<table>
<thead>
<tr>
<th>Reference</th>
<th>Sample Size</th>
<th>Protocol</th>
<th>Test 1 (SD)</th>
<th>Test 2 (SD)</th>
<th>Test 3 (SD)</th>
<th>% change Test 1 - Test 2</th>
<th>% change Test 2 - Test 3</th>
<th>% change Test 1 - Test 3</th>
<th>ICC</th>
<th>Study Limitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gayda 2004</td>
<td>9</td>
<td>Repeated 6MWTs =2, during CR.</td>
<td>462(78)</td>
<td>498(58)(^a)</td>
<td>n/a</td>
<td>8</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>Unknown timing of 2(^{nd}) 6MWT</td>
</tr>
<tr>
<td>Hamilton 2000</td>
<td>94</td>
<td>Repeated 6MWT =3, pre-CR on non-consecutive days.</td>
<td>521(112)(^b)</td>
<td>539(114)(^b)</td>
<td>550(122)</td>
<td>3</td>
<td>2</td>
<td>6</td>
<td>.97(^c)</td>
<td>n/a</td>
</tr>
</tbody>
</table>

CR Cardiac Rehabilitation
ICC Intraclass correlation
NS non-significant
n/a not applicable
SD Standard Deviation
\(^a\) P<.002, Test 1 v Test 2
\(^b\) P < 0.001, Test 1 v Test 2, Test 2 vs. Test 3
\(^c\) ICC for Test 1 to Test 2, Test 2 to test 3
Table 3: Validity of six-minute walk test in cardiac rehabilitation population

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study Design</th>
<th>Time of Testing</th>
<th>Sex (n)</th>
<th>Age (SD) Years</th>
<th>Diagnosis</th>
<th>6MWT protocol</th>
<th>Standard Test for Comparison</th>
<th>Comparison / Correlation</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ades 2003</td>
<td>Random controlled trial: 6 months training in CR, 3 times/wk.</td>
<td>Pre &amp; post CR</td>
<td>Men(0) Women(33) n=19 in intervention group</td>
<td>73 (6)</td>
<td>CAD</td>
<td>Cress 1996</td>
<td>SF-36 questions on 'walk a mile' 'walk a block'</td>
<td>6MWD, r=0.54</td>
<td>.0001</td>
</tr>
<tr>
<td>Gayda 2004</td>
<td>Validity Study: Expired gas analysis during SLET &amp; 6MWT</td>
<td>During CR</td>
<td>Men (22) Women (3)</td>
<td>60 (10)</td>
<td>CAD</td>
<td>Guyatt 1985</td>
<td>6MWT peak oxygen uptake</td>
<td>6MWD, r= .58</td>
<td>n/a</td>
</tr>
<tr>
<td>Gremeaux 2011</td>
<td>Prospective study: using distribution and anchoring-based methods. Training 8-weeks, 3 times/wk</td>
<td>Pre &amp; post CR</td>
<td>Men (77) Women (4)</td>
<td>58 (8)</td>
<td>CAD</td>
<td>ATS 2002</td>
<td>SLET Maximal METs on treadmill</td>
<td>6MWD, r=.59</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>Gremeaux 2009</td>
<td>Random cross over design: SLET followed by 200m fast-walk Test or 6MWT. 6-weeks training, 3 times/wk. 10 subjects performed SLET with gas exchange to determine VT</td>
<td>Pre &amp; post CR</td>
<td>Men (28) Women (2)</td>
<td>52 (8)</td>
<td>CAD</td>
<td>ATS 2002</td>
<td>SLET peak heart rate, before and after CR. = 78% (SD = 6%) of SLET peak heart rate before and after CR 6MWT peak heart rate</td>
<td>n/a</td>
<td>NS</td>
</tr>
<tr>
<td>Hamilton 2000</td>
<td>Validity and Reliability study: in CR phase II and III.</td>
<td>Pre &amp; post CR</td>
<td>Men (61) Women (33)</td>
<td>63 (10)</td>
<td>CAD, NYHA Class: 1=78% 2=19% 3=2%</td>
<td>Steele 1996 &amp; 1995</td>
<td>DASI Physical Function subscale of SF-36 SLET Maximum METs on treadmill or cycle</td>
<td>6MWD, r=.502 6MWD, r=.624 6MWD, r=.687</td>
<td>&lt;.001  &lt;.001 &lt;.001</td>
</tr>
<tr>
<td>Kristjánsdóttir 2000</td>
<td>Validity Study: Comparison of SLET to best of four 6MWTs at 3-months post-surgery</td>
<td>During CR</td>
<td>Men (6) Women (4)</td>
<td>76 (6)</td>
<td>CAD Post Cardiac surgery</td>
<td>Guyatt 1984 &amp; 1985</td>
<td>SLET maximum power SLET maximum heart rate SLET maximum systolic blood pressure</td>
<td>6MWD, r=.93 6MWT maximum heart rate, r=.64 6MWT maximum systolic blood pressure, r=.52</td>
<td>&lt;.001 &lt;.009 &lt;.038</td>
</tr>
</tbody>
</table>


30
<table>
<thead>
<tr>
<th>Study</th>
<th>Design Type</th>
<th>Pre &amp; Post CR</th>
<th>Sample Size</th>
<th>Duration</th>
<th>Inclusion Criteria</th>
<th>Outcome Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roberts 2006</td>
<td>Longitudinal cohort design. 6MWT and SLET pre and post 8-week CR</td>
<td>Men (20)</td>
<td>62(9)</td>
<td>CAD, LVEF &gt;30%, NYHA class 1 - 3</td>
<td>ATs 2002 SLET Heart rate recovery significantly improved in each 30-second interval of a 2-minute recovery period</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Women (10)</td>
<td></td>
<td></td>
<td></td>
<td>6MWT Heart rate recovery significantly improved in each 30-second interval of a 2-minute recovery period</td>
</tr>
<tr>
<td>Verrill 2003</td>
<td>Non-experimental, prospective, comparative in multiple CR programs across Nth Carolina</td>
<td>Men (424)</td>
<td>61(10)</td>
<td></td>
<td>CAD 4% CCF 1% Valve Disease Hamilton 2000 Subscales of the Ferrans and QOLI, including the Health and Function domain</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Women (206)</td>
<td></td>
<td></td>
<td></td>
<td>6MWD = No relationships were found with any domain</td>
</tr>
</tbody>
</table>

American College of Sports Medicine, ACSM; American Thoracic Society, ATS; Cardiac Rehabilitation, CR; Classification, Class; Congestive Cardiac Failure, CCF; Coronary artery disease, CAD; Duke Activity Status Index, DASI; Left Ventricular Ejection Fraction, LVEF; Metabolic equivalents, METs; Not significant, NS; New York Heart Association, NYHA; Powers’ Quality of Life Index-Cardiac Version III QOLI; Six-minute walk test, 6MWT; Six-minute walk test distance, 6MWD; Short Form 36 Health Survey, SF-36; Symptom limited Exercise Test, SLET; Ventilatory Threshold, VT; Times per week, x/wk; Pearson’s Correlation coefficient, r.

Table 4: Reported responsiveness of the 6MWD Pre-CR and Post-CR

<table>
<thead>
<tr>
<th>Study</th>
<th>Subgroups</th>
<th>Total (N)</th>
<th>Men</th>
<th>Women</th>
<th>Age</th>
<th>SD</th>
<th>6MW</th>
<th>CR</th>
<th>SD</th>
<th>Post-CR</th>
<th>SD</th>
<th>Change</th>
<th>SD</th>
<th>% increase</th>
<th>SD</th>
<th>Effect Size</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ades 2003</td>
<td>Resistance group:</td>
<td>19</td>
<td>0</td>
<td>19</td>
<td>73</td>
<td>6</td>
<td>309</td>
<td>101</td>
<td>355</td>
<td>100</td>
<td>43</td>
<td>51</td>
<td>15</td>
<td>n/a</td>
<td>0.45</td>
<td>ANOVA</td>
<td>&lt;.005</td>
</tr>
<tr>
<td>Araya-Ramirez 2010</td>
<td>Entire cohort:</td>
<td>425</td>
<td>293</td>
<td>132</td>
<td>62</td>
<td>12</td>
<td>399</td>
<td>87</td>
<td>472</td>
<td>97</td>
<td>74</td>
<td>n/a</td>
<td>20</td>
<td>16.3</td>
<td>0.84</td>
<td>Paired t-test</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>1st Tertile:</td>
<td>142</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>301</td>
<td>64</td>
<td>383</td>
<td>89</td>
<td>n/a</td>
<td>n/a</td>
<td>28</td>
<td>20</td>
<td>1.28</td>
<td>Paired t-test</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>2nd Tertile:</td>
<td>141</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>410</td>
<td>18</td>
<td>487</td>
<td>57</td>
<td>n/a</td>
<td>n/a</td>
<td>19</td>
<td>14</td>
<td>4.28</td>
<td>Paired t-test</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>3rd Tertile:</td>
<td>142</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>485</td>
<td>37</td>
<td>548</td>
<td>58</td>
<td>n/a</td>
<td>n/a</td>
<td>13</td>
<td>10</td>
<td>1.7</td>
<td>Paired t-test</td>
<td>n/a</td>
</tr>
<tr>
<td>Bittner 2000</td>
<td></td>
<td>156</td>
<td>107</td>
<td>49</td>
<td>60</td>
<td>11</td>
<td>321</td>
<td>102</td>
<td>n/a</td>
<td>47</td>
<td>68</td>
<td>15</td>
<td>n/a</td>
<td>0.53</td>
<td>Paired t-test</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>Gremeaux 2011</td>
<td></td>
<td>81</td>
<td>77</td>
<td>4</td>
<td>58</td>
<td>9</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>73</td>
<td>57</td>
<td>16</td>
<td>12.2</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>Gremeaux 2009</td>
<td></td>
<td>30</td>
<td>28</td>
<td>2</td>
<td>52</td>
<td>9</td>
<td>490</td>
<td>33</td>
<td>552</td>
<td>76</td>
<td>63</td>
<td>56</td>
<td>13</td>
<td>n/a</td>
<td>1.9</td>
<td>Not stated</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Hung 2004</td>
<td>Aerobic training:</td>
<td>9</td>
<td>0</td>
<td>9</td>
<td>70</td>
<td>6</td>
<td>418</td>
<td>74</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>10</td>
<td>n/a</td>
<td>0.56</td>
<td>ANOVA</td>
<td>&lt;.05</td>
</tr>
<tr>
<td></td>
<td>Aerobic &amp; Strength training:</td>
<td>9</td>
<td>0</td>
<td>9</td>
<td>71</td>
<td>7</td>
<td>438</td>
<td>110</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>10</td>
<td>n/a</td>
<td>0.40</td>
<td>ANOVA</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>Maniar 2009</td>
<td>&lt;65 years: b</td>
<td>388</td>
<td>286</td>
<td>102</td>
<td>54</td>
<td>7</td>
<td>360</td>
<td>97</td>
<td>n/a</td>
<td>69</td>
<td>74</td>
<td>19</td>
<td>n/a</td>
<td>0.71</td>
<td>Not stated</td>
<td>&lt;.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>≥65 years:</td>
<td>297</td>
<td>201</td>
<td>96</td>
<td>71</td>
<td>5</td>
<td>317</td>
<td>98</td>
<td>n/a</td>
<td>53</td>
<td>58</td>
<td>17</td>
<td>n/a</td>
<td>0.54</td>
<td>Not stated</td>
<td>&lt;.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>65-74 years: d</td>
<td>226</td>
<td>156</td>
<td>70</td>
<td>69</td>
<td>3</td>
<td>328</td>
<td>99</td>
<td>n/a</td>
<td>55</td>
<td>60</td>
<td>17</td>
<td>n/a</td>
<td>0.56</td>
<td>Not stated</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>75+ years:</td>
<td>71</td>
<td>45</td>
<td>26</td>
<td>79</td>
<td>3</td>
<td>283</td>
<td>88</td>
<td>n/a</td>
<td>48</td>
<td>50</td>
<td>17</td>
<td>n/a</td>
<td>0.55</td>
<td>Not stated</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>Rejeski 2002</td>
<td>Women:</td>
<td>129</td>
<td>75</td>
<td>72</td>
<td>65</td>
<td>8</td>
<td>456</td>
<td>76</td>
<td>514</td>
<td>82</td>
<td>57</td>
<td>48</td>
<td>13</td>
<td>n/a</td>
<td>0.76</td>
<td>t-test</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Men:</td>
<td>72</td>
<td>0</td>
<td>72</td>
<td>n/a</td>
<td>n/a</td>
<td>426</td>
<td>67</td>
<td>486</td>
<td>75</td>
<td>60</td>
<td>39</td>
<td>14</td>
<td>n/a</td>
<td>0.91</td>
<td>t-test</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Roberts 2006</td>
<td></td>
<td>30</td>
<td>20</td>
<td>10</td>
<td>62</td>
<td>9</td>
<td>486</td>
<td>114</td>
<td>553</td>
<td>112</td>
<td>n/a</td>
<td>n/a</td>
<td>14</td>
<td>n/a</td>
<td>0.59</td>
<td>Paired t-test</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Tallaj 2001</td>
<td>LVEF &lt; 40%; e f</td>
<td>50</td>
<td>33</td>
<td>17</td>
<td>61</td>
<td>10</td>
<td>286</td>
<td>89</td>
<td>n/a</td>
<td>74</td>
<td>70</td>
<td>26</td>
<td>n/a</td>
<td>0.84</td>
<td>Paired t-test</td>
<td>&lt;.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LVEF ≥40%:</td>
<td>129</td>
<td>84</td>
<td>45</td>
<td>61</td>
<td>11</td>
<td>329</td>
<td>99</td>
<td>n/a</td>
<td>50</td>
<td>63</td>
<td>15</td>
<td>n/a</td>
<td>0.5</td>
<td>Paired t-test</td>
<td>&lt;.01</td>
<td></td>
</tr>
<tr>
<td>Verrill 2003</td>
<td>Women:</td>
<td>206</td>
<td>0</td>
<td>206</td>
<td>n/a</td>
<td>n/a</td>
<td>328</td>
<td>80</td>
<td>379</td>
<td>79</td>
<td>n/a</td>
<td>n/a</td>
<td>15</td>
<td>n/a</td>
<td>0.63</td>
<td>Not stated</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Men:</td>
<td>424</td>
<td>424</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>386</td>
<td>90</td>
<td>444</td>
<td>92</td>
<td>n/a</td>
<td>n/a</td>
<td>15</td>
<td>n/a</td>
<td>0.65</td>
<td>Not stated</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Wright 2001</td>
<td></td>
<td>159</td>
<td>103</td>
<td>56</td>
<td>62</td>
<td>9</td>
<td>315</td>
<td>76</td>
<td>377</td>
<td>79</td>
<td>62</td>
<td>53</td>
<td>20</td>
<td>n/a</td>
<td>0.82</td>
<td>Paired t-test</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Cardiac Rehabilitation, CR; Meters, m; Standard Deviation, SD, Not Significant, NS, Not applicable, n/a; six-minute walk test distance, 6MWD. Age is in mean years; Pre-CR, Post-CR and Change in 6MWDs are mean distances in meters (6MWDs, where applicable were converted from feet to meters); P-Values are for Pre-CR to Post-CR comparisons of 6MWD. Italics indicate data calculated for this review.


32
Effect Size was calculated by the authors using published data within the study and Cohen’s formula, except Verrill (2003) where effect size was reported.

p<.001 between groups for 6MWD in subjects aged < 65 years versus subjects aged ≥ 65 years at pre-CR assessment, t-test

p<.01 between groups for 6MWD in subjects aged < 65 years versus subjects aged ≥ 65 years at post-CR assessment, t-test

p<.001 between groups for 6MWD in subjects aged 65-74 years versus subjects aged 75+ years at pre-CR assessment, t-test

p<.008 between groups for 6MWD in subjects with LVEF < 40% versus subjects with LVEF ≥40% at pre-CR assessment, unpaired t-test

p<.07 between groups for 6MWD in subjects with LVEF < 40% versus subjects with LVEF ≥40% at post-CR assessment, unpaired t-test

p<.0001 between groups for 6MWD in men versus women, at pre-CR assessment and at post-CR assessment, ANOVA for repeated measures