Objective: To perform a qualitative systematic overview of the measurement properties of the most commonly utilized walk tests in the cardiorespiratory domain: the 2-min walk test (2MWT), 6-min walk test (6MWT), 12-min walk test (12MWT), self-paced walk test (SPWT), and shuttle walk test (SWT).

Data sources: MEDLINE (1966 to January 2000) and CINAHL (1982 to December 1999) electronic databases were searched. Bibliographies of the retrieved articles were reviewed.

Study selection: Clinical trials and observational studies were included if they reported data on the validity, reliability, interpretability, or responsiveness of the 2MWT, 6MWT, 12MWT, SPWT, or SWT. Only studies conducted on patients with cardiac and/or respiratory involvement were included.

Results: Fifty-two studies examining measurement properties of the various walk tests were found: 5 studies on the 2MWT, 29 studies on the 6MWT, 13 studies on the 12MWT, 6 studies on the SPWT, and 4 studies on the SWT. Measurement properties were most strongly demonstrated for the 6MWT. Correlations of 6MWT distance and maximal oxygen consumption ranged from 0.51 to 0.90. A change in distance walked of at least 54 m was found to be clinically significant for the 6MWT. Reliability was shown to be optimized when the administration of walk tests was standardized and at least two practice walks were performed. Patients with increased likelihood of postoperative complications, hospitalization, and death were identified by analysis of distance walked.

Conclusions: Measurement properties of the 6MWT have been the most extensively researched and established. In addition, the 6MWT is easy to administer, better tolerated, and more reflective of activities of daily living than the other walk tests. Therefore, the 6MWT is currently the test of choice when using a functional walk test for clinical or research purposes.

(CHEST 2001; 119:256–270)

Key words: exercise test; heart disease; lung disease; systematic overview; walk test

Abbreviations: 2MWT = 2-min walk test; 6MWT = 6-min walk test; 12MWT = 12-min walk test; NS = not significant; NYHA = New York Heart Association; SPWT = self-paced walk test; SWT = shuttle walk test; \( \dot{V}O_2 \text{max} \) = maximal oxygen consumption; \( W \text{max} \) = maximum work capacity

Patients with cardiac and respiratory disease often present with limited activity levels and exercise capacity. Walk tests are typically administered as a means of evaluating functional status, monitoring treatment effectiveness, and establishing prognosis. A variety of walk tests exist, including time-based tests (eg, 2-min walk test [2MWT], 5-min walk test, 6-min walk test [6MWT], 9-min walk test, and 12-min walk test [12MWT]); fixed-distance tests (eg, 100 m, half-mile, and 2-km walk tests); and velocity-determined walk tests (eg, self-paced walk test...
[SPWT]); and controlled-pacing incremental tests (eg, incremental shuttle walk test [SWT]).

Functional walk tests are exercise tests that measure functional status or capacity, mainly the ability to undertake physically demanding activities of daily living. They are considered objective measures that provide a means of monitoring response to treatment. Compared to traditional laboratory indexes of exercise capacity such as cycle, treadmill, and step ergometry, walk tests require less technical expertise and equipment, making them inexpensive and easy to administer. More importantly, they employ an activity that individuals perform on a daily basis (ie, walking).

In 1976, McGavin and colleagues modified the run test of Cooper to an indoor 12-min walk format for the assessment of exercise tolerance in patients with chronic bronchitis. In 1982, Butland and colleagues explored the possibility of using walking tests of shorter duration (ie, 2 min and 6 min) in similar populations. These time-based tests (ie, 2MWT, 6MWT, and 12MWT) are ideally conducted in an enclosed quiet corridor. Patients are instructed to walk from end to end, covering as much ground as possible in the allotted time period. Distance walked in the specified time period is recorded.

In contrast to the fixed-time walk tests, Bassey and colleagues developed the SPWT, a functional walk test that usually consists of walking a specified distance three times in response to standard instructions to go “rather slowly,” “at a normal pace,” “neither fast nor slow,” and “rather fast, but without over-exerting yourself.” Average speed of walking is calculated for each pace.

In 1992, Singh and colleagues developed the incremental SWT, a walk test based on the 20-m shuttle run test, to assess functional capacity in patients with COPD. The incremental SWT requires patients to walk at increasing speeds up and down a 10-m course. Speed of walking is increased every minute (by 0.17 m/s) and is controlled by audio signals played from a tape cassette. There are 12 speeds or levels, with the number of shuttles (ie, 10-m course) increasing with each level. The test is terminated when the patient becomes too breathless to maintain the required speed or if the patient fails to complete a shuttle in the time allowed.

More recently, Revill and colleagues developed the endurance SWT, a walk test for the assessment of endurance capacity for individuals with COPD. This SWT was designed to complement the incremental SWT using the same 10-m course and audio signal. While the incremental SWT measures maximal capacity, the endurance SWT examines the ability to use that capacity. The intensity (ie, speed of walking) used for the endurance SWT is related to a percentage of each patient’s maximal performance assessed by the incremental SWT.

This article is a qualitative systematic review that examines the measurement properties of functional walk tests. Systematic reviews are “conceivable summaries of the best available evidence” that address a specifically defined clinical question. It is important to make the distinction between a meta-analysis and a systematic review. A meta-analysis is a statistical approach used to combine data and can be used as part of a systematic review. Since no statistical methods were used to pool the data in this review, we have referred to this review as a “qualitative systematic review.” The purpose of this article is to systematically examine the literature of the validity, reliability, interpretability, and responsiveness of the most commonly utilized walk tests in the cardiorespiratory domain: the 2MWT, 6MWT, 12MWT, SPWT, and SWT. Recommendations on the optimal application of these functional walk tests in the clinical and research settings will be made, and potential areas for future research will be outlined.

Materials and Methods

Literature Search

A computer search of the English-language medical literature using the databases MEDLINE (1966 to January 2000) and CINAHL (1982 to December 1999) was conducted. Combinations of the following key words were used: exercise test, walk, walking, heart diseases, lung diseases, and cardiopulmonary. A further manual search examined the key references provided in each article, and potential articles were retrieved. Investigations published only in abstract form were not included, and no attempt was made to contact experts in the field to uncover unpublished material.

Inclusion and Exclusion Criteria

We used the following criteria in selecting studies for this review. The target population consisted of patients with cardiorespiratory disease (ie, cardiac disease, respiratory disease, or both). Studies dealing with elderly subjects were also included. Clinical trials and observational studies were selected if they reported data on measurement properties (ie, validity, reliability, responsiveness, or interpretability) of the 2MWT, 6MWT, 12MWT, SPWT, or SWT. We included two types of studies: ones that specifically set out to examine measurement properties of these walk tests, and clinical trials that examined the effectiveness of interventions but also reported on some properties of these tests. Only data related to the measurement properties were included in this review.

Using the above criteria, the search printouts from MEDLINE and CINAHL were reviewed and a decision was made on which articles to retrieve. Retrieved articles were reviewed; if they fulfilled the inclusion criteria, data were extracted and tabulated.

Assessment of Studies

We could not locate a validity assessment scale to assess the quality of trials dealing with measurement properties of a tool or
instrument. Therefore, no formal instrument was used to examine methodologic quality of the articles. In the assessment of studies, interpretation of the strength of correlations was based on a grading scheme used by Lacasse and colleagues.22,23 Specifically, coefficients of correlations ranging from 0 to 0.20 were considered negligible, 0.21 to 0.35 were weak, 0.36 to 0.50 were moderate, and > 0.50 were strong.

Data Extraction and Synthesis

One reviewer extracted and tabulated information from the original articles selected for inclusion in this review. The tabulated information included the following: (1) the patient population, (2) the specific measure, (3) the measurement property, and (4) the associated results. When there was difficulty in interpreting the results, a second reviewer reviewed the studies.

With respect to the measurement properties of interest, we were specifically interested in validity, reliability, responsiveness, and interpretability. Validity refers to the extent that a test measures what it intends to measure.24,25 Both criterion validity (the extent that a measure is related to other well-established criteria in the same realm24,25) and construct validity (the “theory behind the test”25) are typically demonstrated by showing that correlations with other tests or instruments are in the direction and magnitude that one would expect if the measure under study is working the way it should.26 Reliability, both interrater and intrarater, refers to the consistency of measurement with repeated applications.27 Responsiveness is the ability of a measure to detect clinically important change.24,27 Finally, interpretability refers to the meaning of the test score.27 A score is considered interpretable when it tells whether the difference between two scores (ie, between individuals for a discriminative test or over time for an evaluative measure) is negligible, small, moderate, or large.25 If one knows the clinically minimal important difference (ie, the smallest difference in score that patients’ view as important), interpretation of scores is much easier.28

Results

Literature Search

More than 120 articles were identified through the electronic database and manual searches. Of these, 52 were clinical trials or observational studies that reported on the measurement properties of the 2MWT, 6MWT, 12MWT, SPWT or SWT; the others were reviews or only reported the existence or use of the selected or other walk tests. Table 1 summarizes the evidence for the different measurement properties of the selected walk tests.

2MWT

Five studies on the 2MWT were found (Table 2). Of these, two studies addressed validity,12,29 three studies evaluated reliability,30–32 and three studies assessed responsiveness,29–31

Two studies validated the 2MWT as a similar measure of exercise tolerance as the 6MWT and 12MWT in patients with chronic respiratory disease.12,29 Both Butland and colleagues12 and Bernstein and colleagues29 found strong correlations with distance walked in the respective time intervals; however, Bernstein and colleagues29 further demonstrated that 2MWT distance was moderately to strongly associated with measures of oxygen consumption (ie, some correlations were strong while others were of moderate strength).

One study specifically investigated the effect of encouragement and time on the reliability of the 2MWT.31 In addition to establishing the presence of a learning/training effect with repeated testing, Guyatt and colleagues31 found that encouragement improved distance walked, while the time of day had no effect.

Only one study assessed measurement properties of the 2MWT in each of the pediatric and elderly populations. Upton and colleagues30 evaluated the 2MWT as a measure of exercise tolerance in children with cystic fibrosis. They found the test to be reliable by showing no significant difference in distance walked on repeated testing (ie, two tests). They also determined that the distance walked was more responsive to treatment than peak expiratory flow rate in those with near-normal respiratory function. In the frail elderly subjects, Connelly and colleagues32 reported that intrarater and interrater reliability of the 2MWT to be good to high.

Table 1—Summary of the Evidence for the Measurement Properties of the Selected Walk Tests

<table>
<thead>
<tr>
<th>Walk Tests</th>
<th>Validity</th>
<th>Reliability</th>
<th>Interpretability</th>
<th>Responsiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>2MWT</td>
<td>+ (COPD)</td>
<td>+ (COPD, CHF, CF, Elderly)</td>
<td>0</td>
<td>+ (COPD)</td>
</tr>
<tr>
<td>6MWT</td>
<td>+ (COPD, CHF, pacemaker, ESLD, surgical, PAD, elderly)</td>
<td>+ (COPD, CHF, pacemaker, ESLD, PAD, elderly)</td>
<td>+ (COPD, CHF)</td>
<td>+ (COPD, CHF, pacemaker)</td>
</tr>
<tr>
<td>12MWT</td>
<td>+ (COPD)</td>
<td>+ (COPD)</td>
<td>0</td>
<td>+ (COPD)</td>
</tr>
<tr>
<td>SPWT</td>
<td>+ (CHF, elderly)</td>
<td>+ (Elderly)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Incremental SWT</td>
<td>+ (COPD, pacemaker)</td>
<td>+ (Pacemaker)</td>
<td>0</td>
<td>– (COPD)</td>
</tr>
<tr>
<td>Endurance SWT</td>
<td>0</td>
<td>+ (COPD)</td>
<td>0</td>
<td>+ (COPD)</td>
</tr>
</tbody>
</table>

*Items in parentheses denote population in which measurement property was examined; + = evidence to support; − = evidence to refute; 0 = no evidence; CHF = congestive heart failure; CF = cystic fibrosis; ESLD = end-stage liver disease; PAD = peripheral arterial disease.
Thirty-one studies on the 6MWT were found (Tables 3–5). Of these, 28 studies investigated validity,7,12,29,33–57 13 studies evaluated reliability,7,31,34,36,39,40,44,46,51,52,55,57,58 4 studies determined responsiveness,29,31,38,55 and 2 studies commented on interpretability.29,37 Two studies59,60 did not investigate or substantiate any measurement properties of the 6MWT; however, these studies established normal values and reference equations for the prediction of 6MWT distance for healthy adults aged 40 to 85 years.

The 6MWT has been studied among several different populations (Tables 3–5), including patients with COPD (n = 10), heart failure (n = 13), pacemakers (n = 3), peripheral arterial disease (n = 1), and surgical (n = 2) and pediatric patients (n = 1). The majority of investigations have focused on validating the test by correlating distance walked to several other reference criteria.

In patients with COPD, several studies have examined validity of the 6MWT by correlating distance walked to maximal oxygen consumption (VO₂max), pulmonary function tests, and measures of function and dyspnea (Table 3). Distance walked has been found to strongly correlate with VO₂max and maximum work capacity (Wmax) measured during cycle ergometry.7,29,36,43 and to moderately to strongly correlate with measures of function.7,36 Correlation with spirometry lacks agreement with some studies finding strong relationships42,43 and others weak.29 Correlation to measures of dyspnea have not yielded consistent results,7,36,42,43 and correlation with the Chronic Respiratory Disease Questionnaire has been reported to be weak.43

In patients with heart failure, strong correlations between 6MWT distance and exercise ergometry and VO₂max were observed by several researchers.7,34–36,40,51 and moderate to strong associations with the New York Heart Association (NYHA) functional classifications, oxygen cost diagram, Specific Activity Scale, Chronic Heart Failure Questionnaire, and the Rand Instrument.7,30,55 Furthermore, several studies have shown that 6MWT distance can discriminate between NYHA classification levels.35,40,54 Results of other studies in the same population have shown that distance walked on the 6MWT (≤ 300 m) can identify those with increased likelihood of death or hospitalization within a time frame ranging from 3 months to 1 year.33,34,47,56

<table>
<thead>
<tr>
<th>Source</th>
<th>Population</th>
<th>Measurement Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Butland et al12</td>
<td>Adults with a variety of respiratory diseases; predominately chronic airflow limitation</td>
<td>Validity: Distance walked correlates strongly with 6MWT (r = 0.89) and 12MWT (r = 0.96) distances.</td>
</tr>
<tr>
<td>Guyatt et al13</td>
<td>Adults with chronic airflow limitation or chronic heart failure or both</td>
<td>Reliability: Trend observed that simple encouragement improved distance walked (NS). Time of day of testing had no effect on distance walked. Within-subject variability similar with and without encouragement. Distance walked improved on first two walks compared to last four (p &lt; 0.0001). Responsiveness: Less responsive (0.90) than 6MWT (0.74) with respect to ratio of within-person standard deviation to treatment effect.</td>
</tr>
<tr>
<td>Upton et al30</td>
<td>Children with cystic fibrosis</td>
<td>Reliability: No significant difference on repeated testing; mean coefficient of variation, 2.6%. Responsiveness: More responsive to inpatient treatment than PEFR.</td>
</tr>
<tr>
<td>Bernstein et al32</td>
<td>Elderly men with COPD</td>
<td>Validity: Using 2-min intervals of the 12MWT, distance walked in 2 min correlated strongly with distance walked in 6 min (r = 0.95) and 12 min (r = 0.94). Distance walked correlated strongly with VO₂/kg (r = 0.53); moderately with VO₂max (r = 0.45); and negligibly with spirometric values (r = 0.04–0.13). Responsiveness: Change in distance walked strongly correlated with changes in VO₂/kg (r = 0.53) and VO₂max (r = 0.53)</td>
</tr>
<tr>
<td>Connelly et al32</td>
<td>Frail elderly</td>
<td>Reliability: Interrater, ICC: 0.93–0.95. Intranater, ICC: 0.82–0.89.</td>
</tr>
</tbody>
</table>

*ICC = intraclass correlation coefficient; PEFR = peak expiratory flow rate; VO₂ = oxygen consumption.
<table>
<thead>
<tr>
<th>Source</th>
<th>Population</th>
<th>Measurement Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Butland et al</td>
<td>Adults with stable chronic respiratory disability</td>
<td>Validity: Distance walked correlated strongly with 12MWT ($r = 0.96$) and 2MWT ($r = 0.99$) distances.</td>
</tr>
<tr>
<td>Guyatt et al</td>
<td>Adults with chronic airflow limitation or chronic heart failure or both</td>
<td>Reliability: Simple encouragement improved distance walked (mean 30.5 m, $p &lt; 0.02$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Time of day of testing had no effect on distance walked</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Within subject variability similar with and without encouragement with respect to test-retest (first two walks compared to last four), distance walked improved ($p &lt; 0.0001$).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Responsiveness:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reported to be more responsive (0.74) with respect to ratio of within-person standard deviation to treatment effect.</td>
</tr>
<tr>
<td>Guyatt et al</td>
<td>Adults with chronic heart failure or chronic lung disease</td>
<td>Validity: Distance walked strongly correlated with cycle ergometer test results ($r = 0.58$, $p &lt; 0.0001$) and negatively and moderately correlated with functional status as determined by NYHA criteria ($r = -0.45$, $p = 0.06$) and the SAS ($r = -0.47$, $p = 0.001$).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reliability: Distance walked plateaued during walks 3–6 ($p &lt; 0.0001$). Within-person standard deviation of subject’s mean distance walked was &lt; 6%.</td>
</tr>
<tr>
<td>Guyatt et al</td>
<td>Adults with chronic heart failure or chronic lung disease</td>
<td>Validity: Distance walked strongly correlated with cycle ergometer test results ($r = 0.58$, $p &lt; 0.0001$) and moderately to strongly with four functional status questionnaires (Rand Instrument, BDI, Oxygen Cost Diagram, and SAS) ($r = 0.47–0.59$, $p &lt; 0.0001$).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reliability: ICC, 0.91–0.92. Within-person standard deviation was 22.52 m, with a coefficient of variation of 0.05 for walks 3–6 (vs within-person SD, 29.8; coefficient of variation, 0.07 for all six walks).</td>
</tr>
<tr>
<td>Mak et al</td>
<td>Adults with COPD or severe asthma</td>
<td>Validity: Distance walked significantly ($p &lt; 0.0001$) and strongly correlated with DLco ($r = 0.68$), PEF ($r = 0.55$), and FEV1 ($r = 0.53$); negatively and strongly correlated with breathlessness rating on the MRC scale ($r = -0.52$); moderately correlated with FVC ($r = 0.48$); negatively and weakly correlated with perceived breathlessness as measured by VAS ($r = -0.35$) and RPE measured by the Borg scale ($r = -0.30$); and did not correlate with $SaO2$.</td>
</tr>
<tr>
<td>Wijkstra et al</td>
<td>Adults with COPD</td>
<td>Validity: Distance walked strongly correlated with Winax as assessed by bicycle ergometry ($r = 0.81$, $p &lt; 0.01$), spirometric values and MIP ($r = 0.30–0.58$) and DLco ($r = 0.62$); negatively and moderately correlated with dyspnea at rest measured by the Borg scale ($r = -0.41$, $p &lt; 0.01$); and negligibly-weakly ($r = -0.03$ to $-0.25$) with quality of life, as measured by the CRQ (fatigue, emotion, and mastery domains).</td>
</tr>
<tr>
<td>Bernstein et al</td>
<td>Elderly men with moderate COPD</td>
<td>Validity: Distance walked strongly correlated to $Vo2_{max}$ ($r = 0.51$) and $Vo2_{Ag}$ ($r = 0.67$) as assessed by bicycle ergometry and distance walked in 2 min ($r = 0.95$) and 12 min ($r = 0.97$), but negligibly to weakly correlated with spirometry values ($r = 0.05–0.24$).</td>
</tr>
<tr>
<td>Roomi et al</td>
<td>Elderly subjects with COPD</td>
<td>Validity: Distance walked significantly and strongly correlated with Guyatt dyspnea score ($r = 0.65$, $p = 0.01$).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reliability: Mean distance difference on repeated tests = 0.63 m ($p = 0.94$). Coefficient of repeatability = 63.0.</td>
</tr>
<tr>
<td>Szekely et al</td>
<td>Volume reduction surgery patients secondary to severe COPD</td>
<td>Validity: Preoperative distance walked weakly to moderately correlated with length of hospitalization ($r = 0.32–0.40$, $p &lt; 0.05$). Preoperative distance walked of &lt; 200 m specific (84%, $p &lt; 0.004$) in death prediction.</td>
</tr>
<tr>
<td>Redelmeier et al</td>
<td>Adults with stable COPD</td>
<td>Validity: Distance walked strongly correlated with patients’ rating of their walking ability relative to other patients ($r = 0.59$; 95% CI, 0.54 to 0.63). Interpretability: Distance walked needs to differ by 54 m (95% CI, 37 to 71 m) for the average patient to stop rating themselves as “about the same” and to start rating themselves as “a little bit better” or “a little bit worse” in rating their walking ability relative to others with the same disease.</td>
</tr>
</tbody>
</table>

*BDI = Baseline Dyspnea Index; CI = confidence interval; CRQ = Chronic Respiratory Disease Questionnaire; DLco = diffusing capacity of the lung for carbon monoxide; MIP = maximum inspiratory pressure; MRC = Medical Research Council; PEF = peak expiratory flow; RPE = rating of perceived exertion; $SaO2$ = arterial oxygen saturation; SAS = Specific Activity Scale; VAS = visual analogue scale; see Table 2 legend for other abbreviations.

†Denotes mixed sample of subjects with COPD and/or heart failure.
Table 4—Studies of Measurement Properties of the 6MWT in Patients with Heart Failure*

<table>
<thead>
<tr>
<th>Source</th>
<th>Population</th>
<th>Measurement Properties</th>
</tr>
</thead>
</table>
| Lipkin et al⁵⁵           | Adults with stable chronic heart failure (NYHA class II to III) | Validity:  
  Distance walked related to VO₂max curvilinearly (large variance in those with low vs high VO₂max).  
  All patients considered walk test more representative of daily physical activity than treadmill test.  
  Able to distinguish between normal subjects, NYHA class II, and class III heart failure patients (p < 0.0003) based on distance walked.  |
| Riley et al⁵⁶            | Adults with chronic heart failure (NYHA class II to IV) | Validity:  
  Distance walked strongly correlated with VO₂max measured during treadmill exercise (r = 0.95, p < 0.0001) and peak VO₂ measured during the walk test (r = 0.90, p < 0.0001).  
  Able to distinguish between NYHA class II, III, and IV patients based on distance walked.  
  Reliability:  
  Distance walked increased from test 1 to test 2, but no significant difference was seen from test 2 to test 3.  |
| Bittner et al⁵⁷          | Adults with CHF (SOLVD registry)                    | Validity:  
  Patients in the lowest performance levels (distance walked < 300 m) had a significantly greater chance of dying (10.23% vs 2.99%, p = 0.01), of being hospitalized in general (40.91% vs 19.90%, p = 0.0027), and of being hospitalized for heart failure (22.16% vs 1.99%, p < 0.001) within subsequent year.  
  Compared with those who walked at least 450 m, patients who walked < 300 m had a 3.7-fold risk of dying (95% CI, 1.44-9.33) and those who walked between 300 m and 374.9 m had a 2.78-fold risk (95% CI, 1.09-7.11) within the subsequent year.  
  Interpretability:  
  Each decrement in distance walked of 120 m resulted in a 160% increase in hospitalization for CHF during the 1-yr follow-up period.  |
| Cahalin et al⁵⁸          | Adults with chronic heart failure undergoing cardiac transplant evaluation | Validity:  
  Distance walked strongly correlated with VO₂max (r = 0.64, r² = 0.41, p < 0.0001).  
  In a multivariate analysis of patient characteristics, resting hemodynamics and 6MWT distance, distance walked was the strongest predictor of VO₂max.  
  Distance walked < 300 m predicted an increased likelihood of death or pretransplant hospital admission within 6 mo (40% vs 12%, p = 0.04) but did not predict long-term overall or event-free survival.  
  Reliability: ICC, 0.96.  |
| Peeters and Mets⁵⁹       | Elderly subjects with chronic heart failure (NYHA class II to III) | Validity:  
  Significant difference in distance walked between NYHA class II and III patients and between controls and between control subjects and class III patients (p < 0.001); difference between control subjects and class II patients is NS.  |
| Milligan et al⁶⁰         | Phase 1 cardiac rehabilitation inpatients with left ventricular dysfunction | Validity:  
  Patients who survived > 3 mo walked significantly further than those who survived < 3 mo (5.36 vs 2.35, p = 0.002) according to the following scale:  
  1: nonambulatory; 2: 10–100 feet; 3: 101–500 feet; 4: 501–1,000 feet; 5: > 2,000 feet.  
  Distance walked related (value not given) to 1-yr mortality (p < 0.03) but not to risk of sudden death.  |
| Woo et al⁶¹               | Adults with advanced heart failure (NYHA class III to IV) | Validity:  
  Distance walked not significantly associated with heart rate variability (noninvasive measure of autonomic tone).  
  Distance walked related (value not given) to 1-yr mortality (p < 0.03) but not to risk of sudden death.  |
| Roul et al⁶²              | Adults with heart failure                           | Validity:  
  Distance walked strongly correlated to VO₂max in patients who walked < 300 m (r = 0.65, r² = 0.42, p = 0.001).  
  No significant difference in distance walked between patients who died or were hospitalized for heart failure and those who survived event free. Using ROC curves and survival curve analysis, subjects walking < 300 m tended to have worse outcomes.  
  Reliability: ICC, 0.82.  |
| O’Keeffe et al⁶³          | Frail elderly subjects with heart failure           | Validity:  
  Strong baseline correlation between distance walked and total CHQ score: r = -0.79; with dyspnea dimension of CHQ: r = -0.58.  
  Reliability: ICC, 0.91.  
  Responsiveness:  
  Responsiveness coefficient, 1.73.  
  Effect size for detecting subjective heart failure deterioration (2.13) greater than for detecting improvement (0.85).  
  Change in distance walked strongly correlated with change in total CHQ score (r = 0.70), dyspnea dimension (r = 0.60), fatigue dimension (r = 0.58), and global rating of change (r = 0.78); moderately correlated with emotion dimension of CHQ (r = 0.47).  |
| Opaich et al⁶⁴            | Adults with chronic CHF                             | Reliability:  
  Results (ie, distance walked) of tests performed the same day 30 min apart are equivalent to results of tests performed on 2 consecutive days.  |
| Lucas et al⁶⁵             | Adults with advanced heart failure                  | Validity:  
  Strong correlation (r = 0.57) with peak VO₂ when all subjects included; weak correlation (r = 0.28) when only subjects with peak VO₂ between 10 to 20 mL/kg/min included.  
  6MWT distance did not predict survival.  |

*CHF = congestive heart failure; CHQ = Chronic Heart Failure Questionnaire; ROC = receiver operating characteristic; SOLVD = studies of left ventricular dysfunction; see Tables 2 and 3 for other abbreviations.
Table 5—Other Studies of Measurement Properties of the 6MWT*

<table>
<thead>
<tr>
<th>Source</th>
<th>Population</th>
<th>Validity</th>
<th>Measurement Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rozkovec et al38</td>
<td>Adults with pacemakers</td>
<td>Distance walked directly related to $O_2$ cost (<em>ie</em> as per $O_2$ cost diagram) (p &lt; 0.05).</td>
<td>Significant difference in distance walked between pacing rates (50, 70, and 90 beats/min).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Improvement in breathlessness over preceding 2 wk.</td>
<td>Significantly associated with an increase in walking distance.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Distance walked strongly correlated with cycle ergometry (r = 0.74).</td>
<td>All patients stated walk test replicated daily physical activity more accurately than cycle ergometry.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No significant difference in distance walked found between repeated walks.</td>
<td>Reliability: (\text{NPV} = 0.95)</td>
</tr>
<tr>
<td>Langenfeld et al39</td>
<td>Adults with pacemakers</td>
<td>Distance walked strongly correlated with cycle ergometry (r = 0.74).</td>
<td>All patients stated walk test replicated daily physical activity more accurately than cycle ergometry.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No significant difference in distance walked found between repeated walks.</td>
<td>Reliability: (\text{NPV} = 0.95)</td>
</tr>
<tr>
<td>Provenier and Jordaens53</td>
<td>Adults with pacemakers</td>
<td>Distance walked significantly greater in the VVIR set pacemaker than in the fixed VVI of 60 beats/min (21.9 m; 95% CI, 3.5–40.3 m) and in VVI 85 than in VVI 60 (14.7 m; 95% CI, 0.6–28.9 m).</td>
<td></td>
</tr>
<tr>
<td>Cahalin et al44</td>
<td>Adults with end-stage lung disease (transplant candidates)</td>
<td>Distance walked strong predictor of $V_{O_2}\text{max} (r = 0.73, r^2 = 0.54, p &lt; 0.0001)$.</td>
<td>(\text{NPV} = 0.95) (\text{PPV} = 0.95) (\text{VVI} = \text{fixed rate response; VVIR} = \text{optimal rate response; SF} = \text{short form; see Tables 2, 3 for other abbreviations.}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No significant difference found between estimated and observed $V_{O_2}\text{max}$ using prediction equation based on 6MWT distance.</td>
<td>Reliability: (\text{NPV} = 0.95)</td>
</tr>
<tr>
<td>Kadikar et al49</td>
<td>Adults with end-stage lung disease</td>
<td>Findings suggestive of predictive validity.</td>
<td>Distance walked (&lt; 400 \text{ m} \text{ predictive of death with sensitivity, 0.80; specificity, 0.49; PPV, 0.27; and NPV, 0.91.} \text{ ICC reported to be 0.99; however, no indication was given in methods that reliability was tested.}</td>
</tr>
<tr>
<td>Nixon et al45</td>
<td>Children with end-stage cardiac or pulmonary disease (awaiting transplantation)</td>
<td>Distance walked correlated with $V_{O_2}\text{max} (r = 0.70, p &lt; 0.05)$ and physical work capacity (r = 0.64, p &lt; 0.05), but was not significantly related to indexes of pulmonary function (r = 0.15–0.26).</td>
<td></td>
</tr>
<tr>
<td>Holden et al41</td>
<td>Pulmonary resection patients secondary to bronchogenic carcinoma</td>
<td>Those who had successful surgical outcomes walked further than those who died within 90 d of surgery.</td>
<td>(\text{NPV} = 0.95) (\text{PPV} = 0.95) (\text{VVI} = \text{fixed rate response; VVIR} = \text{optimal rate response; SF} = \text{short form; see Tables 2, 3 for other abbreviations.}</td>
</tr>
<tr>
<td>Montgomery and Gardner52</td>
<td>Adults with intermittent claudication secondary to peripheral arterial disease</td>
<td>Distance walked strongly correlated with the Ankle/Brachial Index (r = 0.55, p &lt; 0.001) and moderately correlated with $V_{O_2}\text{max} (r = 0.37, p = 0.01)$.</td>
<td>Reliability: (\text{NPV} = 0.95) (\text{PPV} = 0.95) (\text{VVI} = \text{fixed rate response; VVIR} = \text{optimal rate response; SF} = \text{short form; see Tables 2, 3 for other abbreviations.}</td>
</tr>
<tr>
<td>Harada et al37</td>
<td>Older adults</td>
<td>Distance walked significantly greater for active vs inactive subjects (p &lt; 0.0001).</td>
<td>Distance walked correlated strongly with lower body strength (ie, chair stands) (r = 0.67), standing balance (r = 0.52), self-reported physical functioning (SF-36 Health Survey) (r = 0.55) and gait speed (r = −0.73); moderately with general health perceptions (r = 0.39) and negligibly with body mass index (r = −0.07).</td>
</tr>
</tbody>
</table>

*NPV = negative predictive value; PPV = positive predictive value; VVI = fixed rate response; VVIR = optimal rate response; SF = short form; see Tables 2, 3 for other abbreviations.
study found no prognostic value for survival of the 6MWT in patients with advanced heart failure. Finally, no significant correlation between 6MWT distance and heart rate variability has been observed.48

In patients with pacemakers, distance walked has been determined to be significantly related to oxygen cost (as measured by oxygen cost diagram38) and exercise ergometry performance.39 Other investigations have shown that distance walked in 6 min is able to discriminate between pacemodes (i.e., fixed rate or optimal response) and rates.38,53

In the context of preoperative assessment, Holden and colleagues41 and Szekely and colleagues50 showed that distance walked was predictive in regards to the probability of successful surgical outcomes in patients undergoing pulmonary resection and volume reduction surgery, respectively. Another study49 determined that distances walked of < 400 m and 300 m were useful indicators of when to list and prioritize patients for lung transplantation, respectively.

Studies of reliability have concentrated on three main areas: the effect of encouragement (in patients with chronic heart and lung disease), serial testing (in patients with chronic lung disease, heart failure, pacemakers and peripheral arterial occlusive disease), and time of testing (in patients with chronic heart and lung disease). Guyatt and colleagues31 found that while the time of day of testing had no significant effect on distance walked, differential encouragement produced variable results. Investigations of repeated testing have shown generally comparable findings, with total distance walked being variable on the initial two walks and establishing consistency on the third.7,31,36,39,40

One study was specifically devoted to interpretability of the 6MWT. Redelmeier and colleagues37 evaluated patients with COPD and established that a minimal change in walking distance of 54 m is clinically significant, in that it translates into a noticeable change in functional status.

Finally, studies of responsiveness have shown that improvement in 6MWT distance is related to diminished breathlessness in pacemaker patients,38 improvement in quality of life for elderly patients with heart failure,55 and that changes in distance walked correlate with changes in V02 max in patients with COPD.29

12MWT

Thirteen studies on the 12MWT were found (Table 6). Of these, 12 studies examined validity9,10,12,20,61–68 5 studies evaluated reliability,9,10,63,64,66 and 4 studies commented on responsiveness.29,63,68,69 Only one study65 had patients perform the test on a treadmill.

Adults with respiratory diseases were the only patient group with published literature on the 12MWT. Several studies have correlated distance walked in 12 min with measures of functional status and V02 max or Wmax with generally consistent findings. For instance, several investigators42,62,63,66 have found moderate to strong correlations with subjective assessments of function, while others40,29,61,64 have found strong correlations with measurements taken during exercise ergometry. As noted for the 6MWT, studies evaluating the relationship between distance walked in 12 min and measures of pulmonary function have produced variable results. For example, correlations between 12MWT distance and FEV1 have ranged from not significant (NS),9,10 to negligible,63,64,66 to moderate,62,66 and to strong.61,65 One study68 found that 12MWT performance (i.e., distance walked) after outpatient pulmonary rehabilitation was the most influential predictor of mortality for individuals with severe COPD when compared to FEV1, arterial blood gases, weight, quality-of-life scores, comorbidity, and oxygen and medication requirements.

The only study to evaluate the 12MWT in the context of preoperative assessment was by Bagg.67 The results of this investigation suggested that the test was not discriminative in regards to the risk of occurrence of postoperative complications in patients undergoing pulmonary resection.

Studies of reliability have shown generally similar findings, with distance walked being variable on the initial two walks and establishing consistency on the third.9,10,64,66

Studies of responsiveness have found the 12MWT to be sensitive to changes in exercise capacity28,69 and to strongly correlate with changes in assessments of breathlessness63 but that changes in distance walked not to be related to long-term survival in patients with COPD.68

SPWT

Six studies on the SPWT were found (Table 7). Of these, five studies evaluated validity33,70–72 and two studies assessed reliability.13,70 One study74 did not investigate or substantiate any measurement properties of the SPWT; however, this study examined the effects of age on choice of walking speed, and derived equations that allow the calculation of age-adjusted normal values for walking speeds at each of the three test paces.

Only one study71 used a population with cardiopulmonary disease. Sensitivity of the SPWT to the presence of congestive heart failure in cardiac patients as well its correlation with another mode of exercise testing was demonstrated.
**Table 6—Studies of Measurement Properties of the 12MWT**

<table>
<thead>
<tr>
<th>Source</th>
<th>Population</th>
<th>Measurement Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>McGavin et al&lt;sup&gt;30&lt;/sup&gt;</td>
<td>Adult men with chronic bronchitis</td>
<td>Validity: Distances walked significantly correlated with VO&lt;sub&gt;2max&lt;/sub&gt; (r = 0.52, p &lt; 0.01). VC, FVC (r = 0.53, p &lt; 0.01), TLC, FEV&lt;sub&gt;1&lt;/sub&gt;, FVC (r = 0.41, p &lt; 0.05), NS with FEV&lt;sub&gt;1&lt;/sub&gt; (r = 0.28, p &gt; 0.05). Reliability: Reported to be reliable if performed twice, however, data used to support this conclusion are not clearly presented.</td>
</tr>
<tr>
<td>McGavin et al&lt;sup&gt;62&lt;/sup&gt;</td>
<td>Adults with respiratory disease (airway obstruction or infiltrative diseases)</td>
<td>Validity: Distance walked significantly (p &lt; 0.01) and strongly correlated with FVC (r = 0.52–0.64), oxygen-cost diagram (r = 0.60–0.65), and RPE (r = 0.49 to 0.74) in both diseases; strongly correlated with DLCO (r = 0.63) and moderately correlated with FEV&lt;sub&gt;1&lt;/sub&gt; (r = 0.44) in those with infiltrative disease.</td>
</tr>
<tr>
<td>Mungall and Haynesworth&lt;sup&gt;1&lt;/sup&gt;</td>
<td>Adult men with chronic bronchitis</td>
<td>Validity: Distance walked strongly correlated with DLCO (r = 0.67, p &lt; 0.01); inversely and strongly correlated with ventilatory response to an increase in oxygen uptake (r = 0.77, p &lt; 0.01); NS with FVC, FEV&lt;sub&gt;1&lt;/sub&gt;, or TLC. Reliability: Distance walked on test 3 significantly (p &lt; 0.05) better than on tests 1 and 2; no significant change after test 3. Coefficient of variation ± 4.2% after test 3.</td>
</tr>
<tr>
<td>Cockcroft et al&lt;sup&gt;68&lt;/sup&gt;</td>
<td>Predominantly (87%) adults with severe COPD</td>
<td>Responsiveness: The most sensitive index of change in functional exercise tolerance after rehabilitation when compared to spirometry, treadmill ergometry, and subjective indexes.</td>
</tr>
<tr>
<td>Alison and Anderson&lt;sup&gt;61&lt;/sup&gt;</td>
<td>Adults with COPD</td>
<td>Validity: Distance walked strongly correlated with W&lt;sub&gt;max&lt;/sub&gt; (r = 0.69, p &lt; 0.001), FEV&lt;sub&gt;1&lt;/sub&gt; (r = 0.62, p &lt; 0.001), and VC (r = 0.65, p &lt; 0.001). VO&lt;sub&gt;2max&lt;/sub&gt; measured during 12MWT did not differ significantly from VO&lt;sub&gt;2max&lt;/sub&gt; measured on the bicycle ergometer.</td>
</tr>
<tr>
<td>O’Reilly et al&lt;sup&gt;52&lt;/sup&gt;</td>
<td>Adult men with COPD</td>
<td>Validity: Distance walked moderately strongly correlated with assessments of breathlessness (r = 0.50–0.70, p &lt; 0.001). Residual relationship: Mean variation of 3.1% when performed twice on same day. 9.1% when performed 2 wk apart.</td>
</tr>
<tr>
<td>Butland et al&lt;sup&gt;47&lt;/sup&gt;</td>
<td>Adults with chronic respiratory disability</td>
<td>Validity: Largest variance (29.6 m) when compared to the 2MWT (23.4 m) and 6MWT (20.0 m) distances.</td>
</tr>
<tr>
<td>Bagg&lt;sup&gt;67&lt;/sup&gt;</td>
<td>Lung resection patients secondary to carcinoma of the bronchus</td>
<td>Validity: No significant difference (p &lt; 0.05) in distance walked between patients who did and did not suffer postoperative complications. While FFT provided significant separation between groups, no further discriminating power was observed with distance walked.</td>
</tr>
<tr>
<td>Swinburn et al&lt;sup&gt;64&lt;/sup&gt;</td>
<td>Adults with severe COPD</td>
<td>Validity: Distance walked strongly correlated with performance on cycle ergometry (r = 0.51, p &lt; 0.001) and step ergometry (r&lt;sup&gt;2&lt;/sup&gt; = 0.52, p &lt; 0.01), but not with FEV&lt;sub&gt;1&lt;/sub&gt; (r&lt;sup&gt;2&lt;/sup&gt; = 0.13) or FVC (r&lt;sup&gt;2&lt;/sup&gt; = 0.17). Reliability: Significant (p &lt; 0.01) increase in distance walked between tests 1 and 4; however, increments between successive attempts tended to decrease.</td>
</tr>
<tr>
<td>Dekhuyzen et al&lt;sup&gt;51&lt;/sup&gt;</td>
<td>Outpatient adults with COPD</td>
<td>Validity: Distance walked strongly correlated with FEV&lt;sub&gt;1&lt;/sub&gt; (r = 0.62, p = 0.001); moderately correlated with FVC, % predicted [r = 0.49, p = 0.001]; FEV&lt;sub&gt;1&lt;/sub&gt;, % predicted [r = 0.43, p = 0.001]. PaCo&lt;sub&gt;2&lt;/sub&gt; (p = 0.44, p = 0.01), and weakly correlated with DLCO (r = 0.34, p = 0.05).</td>
</tr>
<tr>
<td>Bernstein et al&lt;sup&gt;70&lt;/sup&gt;</td>
<td>Elderly men with moderate COPD</td>
<td>Validity: Distance walked correlated strongly to VO&lt;sub&gt;2max&lt;/sub&gt;/kg (r = 0.65) and moderately to VO&lt;sub&gt;2max&lt;/sub&gt; as assessed by bicycle ergometry, correlated negligibly with spirometry values (r = 0.12–0.26). Distance walked strongly correlated with that in 2 min and 6 min (r = 0.94–0.97). Responsiveness: Changes in VO&lt;sub&gt;2max&lt;/sub&gt; more closely related to changes in 12MWT distance (r = 0.72) than to changes in shorter-duration walk test distances (r = 0.53–0.63).</td>
</tr>
<tr>
<td>Gerardi et al&lt;sup&gt;68&lt;/sup&gt;</td>
<td>Predominantly (57%) adults with severe COPD</td>
<td>Validity: Prehabilitation distance walked negligibly correlated with FEV&lt;sub&gt;1&lt;/sub&gt; (r = 0.19, p = 0.03) and weakly correlated with total CRQ score (r = 0.23, p = 0.01). Posthabilitation pulmonary rehabilitation distance walked most significant variable related to prognostic compared to FEV&lt;sub&gt;1&lt;/sub&gt;, ABG, weight, CRQ score, comorbidity, oxygen requirements, and medication requirements. Patients with posthabilitation distance walked &lt; 750 m had 68% 3-yr survival, those with distances walked &gt; 750 m had 92% 3-yr survival. Responsiveness: Change in VO&lt;sub&gt;2max&lt;/sub&gt; more closely related to changes in 12MWT distance (r = 0.72) than to changes in shorter-duration walk test distances (r = 0.53–0.63).</td>
</tr>
<tr>
<td>Larson et al&lt;sup&gt;66&lt;/sup&gt;</td>
<td>Adults with moderate-severe COPD</td>
<td>Validity: Walk distance correlated strongly with MIP (r = 0.52); moderately with FEV&lt;sub&gt;1&lt;/sub&gt;, % predicted (r = 0.40), and moderately and negatively with the total SIF (r = 0.37), the physical dimension of the SIP (r = 0.45), and exercise-related breathlessness as measured by ATS-DLD breathlessness scale (r = 0.49). Reliability: Distance walked increased over first 3 tests (p &lt; 0.001); test-retest reliability, r = 0.98 for tests 3 and 4.</td>
</tr>
</tbody>
</table>

*ABG = arterial blood gases; ATS-DLD = American Thoracic Society-Division of Lung Disease; FVC = inspiratory vital capacity; SIP = sickness impact profile; VC = vital capacity; TLC = total lung capacity; Ve = minute ventilation; see Tables 2, 3 for other abbreviations. Performs on treadmill.
Comparing different populations, Bassey and colleagues\(^ {13}\) examined performance (i.e., speed of walking) of young and elderly men on the SPWT to bicycle ergometry. In addition to finding a strong correlation, the walk test was found to be reliable and sensitive to age differences. Similarly, Cunningham and colleagues\(^ {70}\) investigated the association between four self-selected speeds of walking and age and aerobic power measured on a treadmill. Speed of self-selected walking paces was associated with cardiovascular fitness, independent of age. Furthermore, the reliability of the test was established in the 55- to 66-year age range. More recently, Cunningham and colleagues\(^ {72}\) showed that independent elderly individuals as compared to dependent ones (i.e., living independently compared to living in supervised rest homes) chose to walk at a significantly greater speed at all three paces.

Finally, Piotrowski and Cole\(^ {73}\) used a modified version of the test (i.e., a timed walk over a central 10 m of a longer path at the subject's choice of pace) and demonstrated its ability to detect falls in the elderly.

SWT

Four studies on the SWT were found (Table 8). Of these, three studies examined validity\(^ {14,16,75}\) and two studies evaluated reliability\(^ {14,75}\) of the incremental SWT. One study examined validity, reliability, and responsiveness of the endurance SWT.\(^ {17}\)

The incremental SWT was originally developed and examined by Singh and colleagues\(^ {14}\) as a functional capacity measure in patients with COPD. In this study, both a 10-level and a 12-level version of the test were investigated for feasibility and reproducibility. In addition to noticing a training effect with the 12-level test, distance walked (i.e., number of completed shuttles) was significantly related to distance walked in 6 min. Further study\(^ {16}\) has revealed strong correlations between SWT distance and \(\dot{V}O_2\text{max}\), both measured during treadmill testing and during the walk test itself. In a completely different population, Payne and Skehan\(^ {75}\) were able to demonstrate the validity and reliability of the 12-level incremental SWT in patients with dual-chamber pacemakers.

*HR = heart rate; see Table 2 for other abbreviation.

<table>
<thead>
<tr>
<th>Source</th>
<th>Population</th>
<th>Measurement Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bassey et al(^ {13})</td>
<td>Young (19–21 yr old) and elderly (64–66 yr old) men</td>
<td>Validity: Standardized HR significantly and strongly correlated with those undergoing cycle ergometry ((r = 0.79; p &lt; 0.001)). As a performance index, speed of walking reported to be sensitive to age differences.</td>
</tr>
<tr>
<td>Cunningham et al(^ {70})</td>
<td>Healthy men aged 19 to 60 yr</td>
<td>Validity: When age, weight, height, and fatness held constant, (\dot{V}O_2\text{max}) (treadmill test) significantly ((p &lt; 0.01)) related to speed of walking at the &quot;normal,&quot; &quot;rather fast,&quot; and &quot;fast as you can&quot; paces.</td>
</tr>
<tr>
<td>Ajayi and Balogun(^ {71})</td>
<td>Cardiac patients with and without heart failure</td>
<td>Validity: Walk time correlated significantly and strongly with Bruce Protocol treadmill time in those with cardiac failure ((r^2 = 0.91, p &lt; 0.004)). Able to distinguish between cardiac patients with and without heart failure with regard to walking time, speed, and distance ((p &lt; 0.001)).</td>
</tr>
<tr>
<td>Cunningham et al(^ {72})</td>
<td>Independent (i.e., living independently) and dependent (i.e., living in supervised rest homes) elderly</td>
<td>Validity: Choice of normal walking speed significant determinant of the Incapacity Index (i.e., the slower the normal walking speed, the less likely that the individual can remain independent). Able to distinguish between independent and dependent elderly with regard to speed of walking at each pace.</td>
</tr>
<tr>
<td>Piotrowski and Cole(^ {73})</td>
<td>Elderly subjects representing a variety of abilities and living arrangements</td>
<td>Validity: Modified version able to discriminate between elderly subjects on the basis of previously reported falls.</td>
</tr>
</tbody>
</table>
The endurance SWT was recently developed and examined by Revill and colleagues\textsuperscript{17} as a constant-workload walk test to assess endurance in patients with COPD. This study found that the most suitable endurance intensity in terms of duration and patient acceptability was 85% maximal intensity, as assessed by the incremental SWT. The endurance SWT showed good repeatability after one practice walk and was more sensitive to change following pulmonary rehabilitation than the incremental SWT.\textsuperscript{17}

**DISCUSSION**

The measurement of functional status and capacity has become an integral component of evaluating the impact of an intervention and determining prognosis in patients with cardiorespiratory disease. This qualitative systematic review has examined the measurement properties of the most commonly used functional walk tests in the cardiorespiratory domain. Most of the literature found evaluating measurement properties of functional walk tests pertained to the 6MWT and 12MWT. Few studies have been conducted on the SPWT and SWT.

In addition to considering the measurement properties of these tests, thought must also be given to feasibility, ease of administration, and patient tolerance when recommending a specific test for clinical or research purposes. When comparing timed-based walk tests, the 6MWT presents with a number of advantages. The 6MWT is better tolerated by patients with respiratory disease than the 12MWT.\textsuperscript{12} It

<table>
<thead>
<tr>
<th>Source</th>
<th>Population</th>
<th>Measurement Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singh et al\textsuperscript{14}</td>
<td>Adults with COPD</td>
<td>Validity: 12-level incremental SWT distance significantly related to 6MWT distance but graded cardiovascular response observed with SWT only. HR on the 12-level incremental SWT strongly correlated with HR on the 6MWT ($r = 0.76$); however, maximal HR attained with SWT significantly higher.</td>
</tr>
<tr>
<td>Singh et al\textsuperscript{16}</td>
<td>Adults with chronic airflow limitation</td>
<td>Validity: Distance walked strongly correlated with $\dot{V}O_{2\text{max}}$ during treadmill testing ($r = 0.58$, $p &lt; 0.05$) and during the incremental SWT ($r = 0.81$, $p &lt; 0.05$). Distance walked correlated ($r = 0.36$, $p &lt; 0.05$) with FEV\textsubscript{1} values. Consistent and incremental increase in oxygen consumption and ventilation in response to the increasing intensity.</td>
</tr>
<tr>
<td>Payne and Skehan\textsuperscript{75}</td>
<td>Adults with dual chamber pacemakers</td>
<td>Validity: Incremental SWT able to distinguish between patients with optimal rate response and fixed-rate pacemaker settings ($p &lt; 0.003$).</td>
</tr>
<tr>
<td>Revill et al\textsuperscript{17}</td>
<td>Adults with COPD</td>
<td>Validity: Comparing the endurance SWT to treadmill endurance tests, patients tended to walk for longer on the treadmill at 75%, 85%, and 95% maximal intensity, but differences for HR, dyspnea, and RPE at end of tests were NS. Reliability: Endurance SWT has good repeatability after one practice walk when performed at 85% maximal intensity (ie, strong relation [$r = 0.905$] between tests 2 and 3 with no significant difference between them). Responsiveness: No significant difference in distance walked in either the incremental or endurance SWT following a 5-wk control period prior to rehabilitation. Following 7 wk of pulmonary rehabilitation, endurance SWT duration increased by 160% (effect size = 2.9), whereas incremental SWT distance increased by only 32% (effect size = 0.41).</td>
</tr>
</tbody>
</table>

*See Tables 3, 7 for abbreviations.*
is also more reliable and responsive than the 2MWT, and more reflective of the requirements of activities of daily living. When comparing the 6MWT to the SPWT and SWT, the evidence on the measurement properties of the latter two tests is limited. In addition, the SWT is more difficult to administer, requires more equipment, and may be less reflective of activities of daily living than the 6MWT. Given these advantages and supported measurement properties, the 6MWT should be the test of choice for patients with cardiorespiratory disease.

The majority of the studies that examined validity attempted to correlate walk test performance to traditional indexes of exercise capacity (ie, cycle or treadmill ergometry), $V_o_{2}\max$, and or spirometric measures (Tables 2–8). The use of these measures as appropriate reference criteria for functional status and capacity is questionable. While spirometry is considered an indicator of pulmonary function, pulmonary function alone is not a valid gauge of functional capacity. Similarly, because most activities of daily living represent exertion at submaximal exercise capacity, $V_o_{2}\max$ is not a fitting measure of functional capacity. Furthermore, training effects can be seen in submaximal performance with limited change in maximal capacity.

Part of the difficulty inherent in validating these walk tests as measures of functional capacity is the ambiguity and confusion in the definition of the concept. While some authors define functional capacity as the ability to undertake physically demanding activities of daily living, others use broader definitions encompassing the areas of physical, mental, and social functioning. Furthermore, the terms functional status, functional capacity, functional ability, and functional tolerance have all been used interchangeably in the literature. However, consistent among these definitions and terminology is that behavior is related to daily life. As such, validation of functional walk tests should use criterion measures that incorporate this notion. For example, functional status questionnaires, tests of submaximal exercise capacity, and measures corresponding to activities of daily living may be appropriate options.

Limited literature exists on the interpretability of walk tests. In essence, only one study was found that directly dealt with this property. In this study, a minimal change in walking distance of 54 m for the 6MWT was found to represent a clinically significant change in functional status. This information is useful for examining the results of clinical trials. For example, Redelmeier and colleagues reviewed the literature on the use of walk tests for measuring the effectiveness of treatments for patients with COPD and found that 68% of the studies finding statistically significant results reported differences in 6-min walking distance that were < 54 m. Similarly, a meta-analysis revealed that the best estimate of the effect of respiratory rehabilitation for individuals with COPD on distance walked in 6 min was 56 m, just 2 m greater than the minimal clinically important difference. However, it is important to consider the patient population that Redelmeier and colleagues studied to arrive at their findings (ie, patients with stable COPD; average 6MWT distance, 371 m). A change in distance walked of < 54 m could be a clinically significant change for a patient population that is more severely disabled (for instance, able to walk only 100 m in 6 min). Established normal values and reference equations for the 6MWT distance will potentially allow clinicians to determine percent predicted values.

Several variables have been identified in the literature as potential influences on walk test performance. The use of differential encouragement can affect test results, and therefore necessitates that the use of encouragement with walk test administration be standardized with respect to type and timing. In the studies reviewed, standardized encouragement consisted of specific statements given by the test administrator (eg, “keep up the good work” and “you are doing well”) every 30 s. The magnitude of the effect of encouragement on distance walked found by Guyatt and colleagues (ie, 30.5 m) was similar to that reported in studies claiming to show beneficial effects of treatment interventions, and approached the minimal clinically important difference found for the 6MWT by Redelmeier and colleagues. Furthermore, training and learning effects with repeated testing have been identified in several studies.

Similarly, numerous studies have recognized that the 6MWT has been shown to be discriminative for cardiac patients (heart failure, pacemakers), in patients undergoing pulmonary resection, and in patients awaiting lung transplantation. In contrast, the only study on the 12MWT found that it was not discriminative in regards to the risk of occurrence of postoperative complications in patients undergoing pulmonary resection. The opposing performance of the 6MWT and 12MWT as preoperative assessors in similar populations (ie, pulmonary resection) may
be due to differences in sample size, criteria for successful outcome, and preoperative health status. For the SPWT and SWT, the findings indicate some discriminative ability in patients with cardiac disease and the elderly (SPWT only).

Controversy exists concerning whether particular walk tests (ie, fixed-time walk tests) should be performed on a treadmill or along a corridor. While some studies have found significant differences in energy cost and performance between floor and treadmill walking, others have found them to be equivalent measures. Nevertheless, inherent advantages have been advocated for both modes of testing: treadmill testing allows easy concurrent measurement of physiologic data, such as gas exchange and cardiac status (ie, ECG), whereas corridor walking is easier to implement, requires minimal equipment, and is more closely replicates everyday walking.

This review was systematic rather than narrative in nature. Hunt and McKibbon identified several determinants of quality of a systematic review. Determinants that were met by our review include the following: (1) a focused question, (2) the unlikelihood that relevant high-quality studies were missed, (3) appropriate and clearly defined inclusion and exclusion criteria, and (4) results that will help in the management of patients. The determinants that were not met were (1) having two reviewers extract the data, and (2) using a validity scale to assess the methodologic quality of the studies. Although such a scale exists to assess the quality of the reports of trials, no such tool could be found to assess the quality of studies that examine the measurement properties of tools or instruments. Consideration must also be given to the grading scheme used for interpreting the strength of the correlations when validating the various walk tests. Although the scheme employed in this article has been used in similar reviews, controversy does exist concerning what values constitute weak, moderate, or strong correlations.

When interpreting results of the various studies, consideration must be given to the methodologic quality of the investigations. Numerous studies used small sample sizes and therefore may have lacked power to detect significant differences. Other studies used different methodologies with respect to the administration of the tests. For example, some studies employed standardized encouragement during the walk test, others used encouragement "as needed," and some did not mention the use of encouragement at all. Inconsistencies also existed in regards to where testers placed themselves during the test (ie, stood still, walked behind, or walked beside patients) and the length of corridor used. These variables (ie, differential encouragement and pacing) have the potential to effect performance and therefore suggest caution when making generalizations or comparing study results.

This review revealed the need for additional studies in various areas. Independent evaluation of predictive performance threshold values identified in several studies is required before these values can be used clinically with confidence. Likewise, further validation of the SWT and SPWT in patients with cardiorespiratory disease is required to ensure that these tests are appropriate measures of functional capacity. As stated previously, additional studies are needed on the interpretability of the 6MWT and other functional walk tests in order to assess whether change in test performance relates to clinical and functional changes. Identification of percent change vs absolute values is suggested, so that the values apply to patients with different degrees of disability. Corridor length and indoor vs outdoor walk test administration should also be evaluated to determine if these factors alter performance.

**Conclusion**

The 2MWT, 6MWT, 12MWT, SPWT, and SWT are functional walk tests that are used in the assessment of patients with cardiorespiratory disease. Measurement properties of the 6MWT have been the most extensively researched and established. In addition, the 6MWT is easy to administer, better tolerated, and more reflective of activities of daily living than the other walk tests. Therefore, the 6MWT is currently the test of choice when using a functional walk test for clinical or research purposes. The 6MWT should be administered in an indoor corridor free of distractions, with use of standardized encouragement and at least two practice trials before measurements are recorded. No recommendation can be made on the optimal corridor length, as no investigations have evaluated this variable.

**References**


268 Reviews
11 Cooper KH. A means of assessing maximal oxygen intake: correlation between field and treadmill testing. JAMA 1968; 203:201–204
21 Moher D, Jadad AR, Klassen TP. Guides for reading and interpreting systematic reviews: III. How did the authors synthesize the data and make their conclusions? Arch Pediatr Adolesc Med 1996; 152:915–920
22 Lacasse Y, Wong E, Guyatt G. A systematic overview of the measurement properties of the chronic respiratory questionnaire. Can Respir J 1997; 4:131–139
25 Currier DP. Elements of research in physical therapy. Baltimore, MD: Williams & Wilkins, 1990; 160–177
47 Milligan NP, Havey J, Dossa A. Using a 6-min walk test to predict outcomes in patients with left ventricular dysfunction. Rehabil Nurs 1997; 22:177–181
67 Bagg LR. The 12-min walking distance: its use in the pre-operative assessment of patients with bronchial carcinoma.