Validation of a 2 Minute Step Test for Assessing Functional Improvement

F. Haas¹,², G. Sweeney², A. Pierre², T. Plusch², J. Whiteson¹

¹Department of Rehabilitation Medicine, NYU School of Medicine, New York, NY, USA
²The Joan and Joel Smilow Cardiac Rehabilitation and Prevention Center of the Rusk Institute of Rehabilitation Medicine, NYU Langone Medical Center, New York, NY, USA

Email: haasf01@med.nyu.edu

Abstract

Introduction: Although many cardiopulmonary patients require home care services, the 6-minute walk test (6MWT) is unusable in most dwellings for objectively evaluating exercise capacity because it requires a 20 - 30 meter hallway. To meet this need, we developed a 2-minute step test (2MST): stepping up and down an 8 inch step for 2 minutes (1 step = bilateral step up + step down). Purpose: Evaluate the statistical validity, reproducibility, and sensitivity of the 2MST in assessing exercise capacity. Method: We compared the heart rate, oxygen saturation and perceived exertion obtained during performance of 2MST with those obtained during the 6MWT. Results: Comparing 2MST and 6MWT in 158 subjects for validity, r = 0.925 (P < 0.0001). They were statistically equivalent in heart rate (98 ± 4 and 94 ± 4 bpm), oxygen saturation (96 ± 0.5 and 95 ± 0.6%), rating of perceived exertion (14.6 ± 0.4 and 13.4 ± 0.4), and blood pressure (130 ± 4/62 ± 3 and 128 ± 3/62 ± 2 mmHg), (mean ± SE, 2MST and 6MWT, respectively). Reproducibility revealed small learning effects: 8% and 5% for the 2MST and 6MWT, respectively (P < 0.001). Sensitivity of the 2MST was high, based on the number of steps achieved by inpatients (27 ± 13 steps, mean ± SD), outpatients (48 ± 14 steps), and healthy subjects (64 ± 18 steps) (P < 0.001), and by the outpatients before (36 ± 11 steps) and after (42 ± 14 steps) rehabilitation (P < 0.001). Conclusion: The 2MST is valid, reproducible, sensitive, safe, well-tolerated, and is a suitable substitute for the 6MWT.

Keywords

6 Minute Walk Test, Step Test, Exercise Capacity

1. Introduction

Section 3021 of the Patient Protection and Affordable Care Act (PPACA), aka
“the health care reform bill,” establishes the Center for Medicare and Medicaid Innovation (CMI) [1]. Its purpose “is to test innovative payment and service delivery models to reduce program expenditures while preserving or enhancing the quality of care furnished to individuals.” Many of the 18 specific models that this legislation has earmarked for CMI to evaluate involve increased home-based care, and one in particular targets “promoting greater efficiencies and timely access to outpatient services (such as outpatient physical therapy services).”

Many patients with chronic cardiopulmonary problems require home care services, including physical therapy, after discharge from acute hospitalization or inpatient cardiopulmonary rehabilitation units. Critical to providing the most effective home care for these patients is the ability to conduct a meaningful assessment of exercise ability, first at the initiation of care, then periodically to evaluate progress. Exercise ability is defined as “a patient’s ability to undertake physically taxing activities encountered in everyday life” [2].

The single existing tool for home-based assessment of exercise ability is the 6-minute walk test (6MWT), which was developed for use in the professional rehabilitation setting and required a hallway that was 30 meters in length. This self-paced sub-maximal exercise evaluation tool is a validated test with good reproducibility that is well-accepted by patients, easily administered, requires little equipment, and has been used to assess functional status in a variety of patient populations [3] [4] [5] [6].

The 6MWT has also been used as an outcome measure for patients participating in cardiopulmonary rehabilitation [7] [8] [9] [10] but it has serious limitations in this context. Although the 6MWT is currently advocated for evaluating functional exercise capacity regardless of the setting, in reality it is not feasible in most apartments, homes, and medical offices because of its requirement for a hall that is 30 meters long [11]. This long-standing need for a practical, effective exercise evaluation tool that is independent of the dimensions of physical space takes on even more urgent importance with the new health care mandate for improving both efficiency of, and access to, home care.

A step-test appears to hold potential as an exercise evaluation tool that will meet this need. Although one had not yet been designed that is appropriate for assessing a population with cardiopulmonary limitations, the history of step tests in estimating exercise capacity over the past several decade’s points in a promising direction. Stepping was first used for exercise evaluation in the early 1940s. Developed by Brouha et al. in the Harvard Fatigue Laboratories during WWII [12], the Harvard Step Test, as it became known, requires a 20-inch step and a stepping frequency of 30 steps per minute (one step every two seconds) for 5 minutes, or until exhaustion. Because the test was designed for healthy young men, however, it is impractical or irrelevant for older or impaired subjects.

Although step tests developed after the Harvard Step Test have been widely used to assess cardiorespiratory fitness [13] [14] [15] [16], they all suffer from similar shortcomings. They require a pre-set cadence that is inappropriate for subjects who are older or impaired, and/or their step height—ranging from 13” -
—is too difficult for patients who are short or have neuromuscular or skeletal problems. As one typical example, Shephard et al. [15] developed the Canadian Home Fitness Test, a self-administered step test that requires subjects to step at an age- and sex-specific rhythm using a series of two 8" steps (16” total height) because that is typical of what is found in most North American homes [17]. Bolton et al. [18] developed a non-cadenced (self-paced) stair climbing test in which patients walked up a maximum of five flights (total of 127 steps). Once the patient stopped, number of steps and climbing time were recorded. They found a significant correlation between steps walked and PFT variables. When this test was evaluated in COPD patients [19], the number of steps walked correlated significantly with Ve and VO2. Dal Corso et al. [20] illustrated the use of a single-stage, self-paced, 6-minute step test (6MST) for patients with interstitial lung disease that used a 20 cm (7.9”) step height. There were no significant differences between repeated tests in the total number of steps climbed or the rate of climbing. In addition, both parameters were related to peak aerobic capacity and breathlessness in daily life. Seynnes [21] developed a 4-step test showing stair-climbing to be highly correlated with the 6MWT in a population of frail elders.

Stepping is particularly attractive for the current purpose because it requires a minimum of space and technical expertise, especially when used in conjunction with pulseoximetry [22]. In addition, stair height (i.e., the riser) is dictated by local building code and varies only minimally from state to state. In many states this is 20.3 cm; in most of the others it is either 19.7 cm or 21 cm. In New York State, for example, the great majority of buildings constructed after 1945 have a mandated stair height between 20.3 - 21 cm (2008 NYS Stair Code: 311.5.3.1) [17]. We thus developed a self-paced 2-minute step test (2MST), based on a step height of 8”, that can be used in constrained spaces and is a suitable alternative to the 6MWT. We chose a 2-minute duration in place of a longer cadenced step test because clinical experience suggests that a briefer self-paced step test improves patient compliance among lower-level cardiac and pulmonary patients undergoing rehabilitation.

The study compares this self-paced, single-stage, single-step 2MST with the 6MWT to assess its validity, reproducibility, and sensitivity and thus determine its value as a surrogate for the 6MWT when the latter cannot be performed according to the prescribed standards [11].

2. Methods

The project was approved by the medical center’s IRB and the data were collected at the NYU Langone Medical Center’s Department of Cardiopulmonary Rehabilitation.

Participants: Subject participating in the study provided informed consent. Inpatients and outpatients who were participating in cardiopulmonary rehabilitation performed a 2MST and a 6MWT as part of their routine functional evaluation. Healthy subjects were volunteers from the rehabilitation staff, friends, and
families, and provided informed consent. Subjects with significant neuromuscular or skeletal impairment were excluded from testing.

**Data collection:** The following variables were collected: gender, age, diagnosis, heart rate (HR), O2 saturation (O2 sat), systolic and diastolic blood pressures (BPs, BPd), and rating of perceived exertion (RPE) [22].

**Procedures:** With inpatients, the 2MST and 6MWT were done within the first week following their transfer to the inpatient rehabilitation unit. The order was randomized for each patient. Inpatient data were collected from chart reviews. Patients performed the 2MST and 6MWT twice each during the first week of their program, then once each at the end of their rehabilitation program. Healthy volunteers performed two 2MWT and two 6MWT within one week. In the 2-day repeat design for outpatients and healthy volunteers, on Day One the 2MST and a 6MWT were carried out in random order, with a 30-minute rest period between tests. On Day Two, the two tests were carried out in the reverse order. Outpatients performed the two post-program tests in randomized order within the first week of their enrollment.

The 6 MWT was performed in hallway adjacent to the gym and the 2MST was performed on a step located in the rehabilitation gym. It consisted of stepping up and down one 8”-high step as rapidly as possible for 2 minutes. One step = one bilateral step up + one bilateral step down. Prior to beginning a test, subjects stood quietly for 2 minutes while HR, O2 sat, BP, and RPE were recorded. The same variables were assessed again immediately at test termination. The 6MWT was performed in accord with ATS/ACCP recommendations [11]. For the 2MST test, the instructions given subjects were adapted from the instructions for the 6MWT. Prior to beginning the test, subjects were told:

“**You are now going to do a 2-minute step test. The object of this test is to step up and down as many times as you can in 2 minutes. You can hold the rail for balance, but do not pull yourself up while holding the rail. You may get out of breath or feel tired. You are permitted to slow down, stop or rest as necessary, but you should resume stepping as soon as you are able.**”

“As you are stepping up and down, you will be informed of the time when there is 1 minute remaining, and when there are 15 seconds remaining. You will be encouraged to do your best. Please do not talk during the test unless you have chest pain or dizziness. **When the 2 minutes are up, I will ask you to stop. Please stop immediately.**”

*Do you have any questions?*

“**Remember that the object is to step up and down as many times as you can in 2 minutes, so remember to pace yourself.**

*Start stepping now!*”

**Data analysis:** A power analysis was performed to assure that the sample size was sufficient to guarantee the statistical data obtained. A construct validity of the 2MST was determined by a Pearson correlation analysis between number of steps attained during the 2MST and distance covered during the 6MWT collected from a chart review of subjects enrolled in inpatient cardiopulmonary re-
habilitation, the Day One trial of outpatients, and the trial of healthy volunteers. Physiologic response differences to the tests were evaluated using t-test analysis on HR, O₂ sat, BPs, BPd, and RPE.

Reproducibility was evaluated in the outpatient and healthy subgroups by comparing the two 2MSTs and two 6MWTs performed within the same week using paired t-test analysis. A power analysis was performed to justify the sample size of each group.

The sensitivity of the 2MST (i.e., its ability to distinguish between levels of fitness) was evaluated in two ways. We compared the number of steps attained respectively by inpatients, by outpatients, and by healthy subjects. In addition, we assessed the 2MST’s ability to measure the functional results of cardiopulmonary rehabilitation within the outpatient group by comparing the pre- and post-rehabilitation data from the 2MST to the pre- and post-rehabilitation data for the 6MWT. Two-group sensitivity was determined by t-test analysis. Multigroup sensitivity was assessed using analysis of variance (ANOVA) followed by Tuckey’s Comparison test.

3. Results

A total of 159 subjects (Table 1) were tested: 105 were inpatient cardiopulmonary rehabilitation patients, 32 were enrolled in outpatient cardiopulmonary rehabilitation, and 22 were healthy volunteers.

Validity was determined by comparing the basic data generated by the 2MST and 6MWT performed by inpatients, by the outpatients on Day One, and by healthy subjects. The number of steps attained in the 2MST and the distance covered in the 6MWT were highly correlated (r = 0.93, P < 0.001) (Figure 1).

During the 2MST, heart rate increased more than it did during the 6MWT by a small but significant amount (Table 2). There were no significant differences between the 2MST and the 6MWT in perceived exertion, oxygen saturation, or blood pressure.

We assessed the reproducibility of each test by comparing the results of the first and second 2MST and the first and second 6MWT in the combined groups of 32 outpatients and 22 healthy subjects (Figure 2). A small but significant increase was observed in both the number of steps accomplished on the repeat 2MST (trial 1 = 48 ± 2.2 steps, trial 2 = 52 ± 2.5 steps, mean ± SE, P < 0.001)

<p>| Table 1. Population demographics (mean ± SD). |
|-----------------|-----------------|-----------------|-----------------|</p>
<table>
<thead>
<tr>
<th>N</th>
<th>Age (years)</th>
<th>Height (cm)</th>
<th>Weight (kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inpatients</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>48</td>
<td>77 ± 8</td>
<td>156 ± 67</td>
</tr>
<tr>
<td>Men</td>
<td>57</td>
<td>76 ± 9</td>
<td>170 ± 7</td>
</tr>
<tr>
<td>Outpatients</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>13</td>
<td>64 ± 14</td>
<td>165 ± 10</td>
</tr>
<tr>
<td>Men</td>
<td>19</td>
<td>69 ± 11</td>
<td>179 ± 8</td>
</tr>
<tr>
<td>Healthy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>12</td>
<td>26 ± 12</td>
<td>161 ± 7</td>
</tr>
<tr>
<td>Men</td>
<td>9</td>
<td>34 ± 19</td>
<td>175 ± 7.6</td>
</tr>
</tbody>
</table>
Figure 1. Scattergram showing the relationship between 2MST and 6MWT.

Figure 2. Test-retest data demonstrate minimal differences in both the 2MST and the 6MWT (mean ± SD).

Table 2. Maximum heart rate (HR), perceived exertion (RPE), oxygen saturation (O₂ sat), Systolic blood pressure (BPsys) and diastolic blood pressure (BPdias) attained during the 2MST and the 6MWT (mean ± SD).

<table>
<thead>
<tr>
<th></th>
<th>2MST</th>
<th>6MWT</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart Rate</td>
<td>100.6 ± 27.9</td>
<td>96.8 ± 24.0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>RPE</td>
<td>14.0 ± 2.3</td>
<td>13.5 ± 7.1</td>
<td>NS</td>
</tr>
<tr>
<td>O₂ Sat</td>
<td>94.9 ± 4.1</td>
<td>95.0 ± 7.9</td>
<td>NS</td>
</tr>
<tr>
<td>BPsys</td>
<td>130.7 ± 25.6</td>
<td>130.9 ± 23.7</td>
<td>NS</td>
</tr>
<tr>
<td>BPdias</td>
<td>63.1 ± 12.0</td>
<td>64.2 ± 11.4</td>
<td>NS</td>
</tr>
</tbody>
</table>

and the distance achieved during the repeat 6MWT (trial 1 = 455 ± 17.7 m, trial 2 = 472 ± 15.5 m, P < 0.005). The two methods increased by a similar degree (8% in the 2MST, 5% in the 6MWT, P > 0.1). There were no significant differences in HR, perceived exertion, or BP between the two trials in either the 2MST or 6MWT (Table 3).

Sensitivity of the 2MST was evaluated from two perspectives. One concerned the ability to distinguish between different levels of function, and thus compared the results from low-level subjects (the 105 inpatients), mid-level subjects (the 32 outpatients), and healthy subjects (the 22 volunteers). Figure 3 illustrates the ability of the 2MST to distinguish between these functional levels. Inpatients averaged 23.6 ± 11.1 steps (mean ± SD), compared to 48.4 ± 13.7 steps for the outpatient group and 64.4 ± 18.1 steps in the healthy population. ANOVA indi-
Figure 3. The 2MST distinguishes between respective fitness levels of low-level inpatients, mid-level outpatients, and high-level healthy subjects (mean ± 95% CI, P < 0.001 between each group as assessed by Tuckey’s post-test).

Table 3. Maximum heart rate (HR), perceived exertion (RPE), Systolic blood pressure (BPsys) and diastolic blood pressure (BP dias) attained during repeated 2MST and 6MWT (mean ± SE).

<table>
<thead>
<tr>
<th></th>
<th>2MST</th>
<th>6MWT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial 1</td>
<td>Trial 2</td>
<td>Trial 1</td>
</tr>
<tr>
<td>Heart Rate</td>
<td>111 ± 3.9</td>
<td>115 ± 4.3</td>
</tr>
<tr>
<td>RPE</td>
<td>13.2 ± 0.3</td>
<td>13.1 ± 0.4</td>
</tr>
<tr>
<td>BPsys</td>
<td>140 ± 4.0</td>
<td>140 ± 3.9</td>
</tr>
<tr>
<td>BP dias</td>
<td>65 ± 1.4</td>
<td>65 ± 1.3</td>
</tr>
</tbody>
</table>

cated a significant difference among the three groups (F = 116.8, P < 0.001). Tuckey’s post-test confirmed significant differences between inpatient and outpatient, inpatient and healthy, and outpatient and healthy (P < 0.001 between each group). The other perspective concerned the ability to accurately identify change in function within an individual. This was carried out in the outpatient group, comparing the results from the 2MST and the 6MWT performed at the beginning and then at completion of their outpatient pulmonary rehabilitation program. Each post-test showed significant improvement, with no significant difference between the 2MST and 6MWT in the degree of improvement identified (Table 4). The 2MST performance improved by 20 ± 4% compared to 18.1 ± 4% on the 6MWT (t-test, P = NS), which indicates comparable sensitivity between the two tests in assessing improvement.

4. Discussion

The initial population that we tested comprised patients discharged to inpatient cardiopulmonary rehabilitation following acute hospitalization. They were all low-level patients with significant functional limitations. As most current step tests are 6 minutes in duration [22] [23], which our experience suggests is too strenuous for this patient population to complete, we chose a 2-minute step test
Table 4. Improvement in 2MST and 6MWT following outpatient pulmonary rehabilitation in 32 patients (mean ± SE).

<table>
<thead>
<tr>
<th>Test</th>
<th>Pre-rehab.</th>
<th>Post-rehab.</th>
<th>% change</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>2MWT (steps)</td>
<td>35.7 ± 1.9</td>
<td>41.6 ± 2.0</td>
<td>20 ± 4%</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>6MWT (meters)</td>
<td>338.5 ± 16.3</td>
<td>385.7 ± 16.6</td>
<td>18.1 ± 4%</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

For assessment. Our goal was a feasible and highly flexible test for functional assessment with results comparable to a known standard. As shown in Table 2, peak HR, O₂ saturation, and blood pressure response on the 2MST were similar to those obtained in this low-level patient group with the 6MWT. This similarity of physiologic variables was corroborated by the similarity in perceived exertion score. Then, we found that mid-level patients (i.e., those able to perform CPETs) reached similar HR levels during both the 2MST and 6MWT (Table 3). These values are similar to those observed in previous studies using the 6MWT [3] [24]. These consistent similarities suggest that, despite the absence of published values for HR during self-paced step tests, the similarity in HR attained during the 2MST and the 6MWT in both patient populations indicate that the 2MST is equally stressful for assessment purposes and is, therefore, a valid substitute for the 6MWT.

In our use of the 2MST, we attempted to control for external factors that have been identified as sources of variability in the 6MWT that could potentially influence results (Inherent factors, i.e., patient sex, age, weight, height, and muscle function [18], cannot be controlled for). Performance encouragement and familiarity with the test have been shown to increase the distance walked [10] [24] [25]. We accounted for both of these factors in our study design. In addition to having subjects repeat each test twice, we included standardized encouragement given by the test administrator modeled on ATS/ACCP recommendations for the 6MWT [11]. This involved time cues given every 30 seconds along with such comments as “you are doing well!” or “good work.”

Because learning effects with repeated testing have been shown to improve outcomes of the 6MWT [26] as well as step tests [22] we evaluated the effects of familiarity by having subjects perform each test twice. The 5% improvement noted in the 6MWT is consistent with previous reports [5] [10], and the 7% improvement in the 2MST reflects a learning effect that is not significantly different. Although we did not establish the number of repetitions of the 2MST needed to establish a true learning plateau, the relatively small difference that we observed between the first two repetitions indicates that using the 2MST as an end point in an interventional investigation would require performing the test a minimum of two times prior to assessing the intervention.

Several limitations in our study underline the need for further study to explore the sensitivity of this new field test in greater detail. Because our control group was too small to distinguish between men and women, accurate normative data could not be derived. A second limitation concerns the age difference between the patient groups and the healthy control group. We chose a young
healthy population to demonstrate that this new field test accurately assesses capability over a very broad spectrum of fitness. The study thus must be repeated with an older healthy aging population. In addition, we used very broad diagnostic categories—i.e., cardiac, pulmonary—and future studies should be more specific in characterizing pathologies, e.g., COPD vs. restrictive pulmonary disease.

In addition to confirming our findings, future studies are also needed to establish the minimal clinically significant performance difference in the 2MST. This would, in part, require the evaluation of floor and ceiling effects. Independent evaluation of predictive performance threshold values is required before these values can be used clinically with confidence. Finally, our single study assessing this new field test must be followed by more to gauge the interpretability of the 2MST and determine the degree to which change in test performance reflects clinical and functional changes.

5. Conclusion

In conclusion, the measurement of functional status and exercise capacity is an integral part of evaluating patient fitness level and the impact of interventions such as cardiopulmonary rehabilitation. The present study shows the 2MST to be a low-cost, well-tolerated, reproducible, and accurate field test that can assess functional status and exercise tolerance in patients across a spectrum of fitness levels and in the limited physical space typically found in medical offices, acute care facilities, and patient homes.

References


https://doi.org/10.1378/chest.118.1.98

https://doi.org/10.1136/thx.39.11.818

https://doi.org/10.1136/thx.43.5.388

---

**Submit or recommend next manuscript to SCIRP and we will provide best service for you:**

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.  
A wide selection of journals (inclusive of 9 subjects, more than 200 journals)  
Providing 24-hour high-quality service  
User-friendly online submission system  
Fair and swift peer-review system  
Efficient typesetting and proofreading procedure  
Display of the result of downloads and visits, as well as the number of cited articles  
Maximum dissemination of your research work  
Submit your manuscript at: http://papersubmission.scirp.org/  
Or contact ojtr@scirp.org