The Ramp Power Test: A Power Assessment During a Functional Task for Older Individuals

Joseph F. Signorile,1–3 David Sandler,1,2 Lani Kempner,1 Damian Stanziano,1,2 Fangchao Ma,3 and Bernard A. Roos1–4

1Department of Exercise and Sport Sciences, University of Miami, Coral Gables, Florida.  
2Stein Gerontological Institute, Miami Jewish Home and Hospital, Florida.  
3Geriatric Research, Education, and Clinical Center, Veterans Affairs Medical Center, Miami, Florida.  
4Departments of Medicine and Neurology, University of Miami Miller School of Medicine, Florida.

Background. Power is critical to mobility and activities of daily living and is a key determinant of independence and falls prevention. Therefore, the quantification of power in older persons is critical. The power tests currently available are often expensive, potentially dangerous, and not reflective of everyday activities. We present a modification of an existing field test that uses ambulation up a standard access ramp to quantify functional power in older individuals.

Methods. Three hundred sixty-three women and 157 men, aged 73.1 ± 7.0 years, ambulated up a standard access ramp (1:12 rise/run ratio) as quickly as possible. Each person performed one practice and two timed trials.

Results. Comparisons with accepted power measures and reported patterns of change with aging supported the validity of the ramp power test. The test was found to be reliable across multiple trials and days. Pair-wise comparisons showed that for women the test was sensitive to differences in power output by half-decade, whereas for men it could distinguish between 9 of the 15 comparisons among age groups. Percentile scores are reported by half-decade for power in both genders. In > 1200 trials performed during this study, only one injury (a slightly strained hamstring) occurred.

Conclusions. The ramp power test is valid and reliable and can safely distinguish power by half-decade in women and among the majority of age groups in men. Its safety, low cost, and ease of administration make it a feasible diagnostic tool to assess functional power levels in ambulatory older persons.

From age 65 to 89, lower limb extensor power declines 3.5% annually compared to a 1%–2% drop in strength (1). In men, maximal anaerobic power declines 8.3% per decade from ages 20 to 70 (2). These data reflect a loss of faster contracting motor units (3,4).

Power is associated with independence (5), lower fall risk (6), and a greater likelihood of positive rehabilitation outcomes (7). Additionally, the higher levels of disability in elderly women compared to elderly men appear to be related to lower power production (8) or lower power-to-body weight ratios (9).

The current techniques for measuring power in older persons are often expensive, relatively dangerous, or not task-specific. Methods include vertical jumps (8,10,11), cycle ergometry (2,12,13), isokinetics (6,14,15), and special computerized rigs (5,7,16–18).

The Margaria–Kalame test is a valid and reliable field test of power (19). The test is a modification by Kalame (20) of the original test developed by Margaria and coworkers (21). However, because the test requires bounding up a set of stairs, a task where fall probability and the potential for serious accidents is increased (22–24), its risk-to-benefit ratio may be questionable. Additionally, skill, coordination, leg length, and approach distance all affect power production during the test (25,26), and skill training and practice (25,27) have a greater effect on stair climb than on ramp tests (28). Nevertheless, stair tests are used extensively to measure power (29,30) and functional performance (17,31).

Costa and colleagues (28) compared power outputs during a stair and ramp power test. They found that the ramp test produced significantly greater power than the stair climb. In a separate study, Huskey and coworkers (32) confirmed these results and suggested that the ramp reduced the impact of skill and coordination on power production. They noted that women were especially affected due to their smaller stature and shorter stride lengths.

This article examines a further modification of the ramp test using a standard access ramp commonly found in most facilities (1:12 rise/run) rather than the 30.5° (32) or 35° ramps (28) previously used. The lower incline makes the test accessible to a greater segment of older persons and increases velocity contribution to power production. This is important because this factor makes muscle utilization patterns (14,33–35) and force–velocity relationships (36–38) more specific to walking. The purpose of this article is to examine the validity and reliability of the test and to present age-related norms computed from a sample of independent-living elders.

Methods

Participants

Data from 520 independent, community-dwelling participants were included in this analysis. Independence was defined as living in their own home or apartment without outside help.
Potential participants presented a comprehensive health evaluation from their doctor certifying that they had no acute systemic illnesses or conditions that would disqualify them from participating in an exercise study. These conditions included unstable angina; active pericarditis, myocarditis, or endocarditis; resting electrocardiogram ST segment depression (> 3 mm); myocardial infarction within the past 3 months; uncontrolled symptomatic congestive heart failure, ejection fraction < 30%; moderate to severe aortic stenosis, thrombophlebitis, or recent embolism requiring anticoagulants; resting systolic blood pressure > 164 mmHg or resting diastolic blood pressure > 94 mmHg; or history of hypertension with exercise or resting pulse rate > 90 or < 50. Participants were also free of chronic obstructive pulmonary disease requiring oxygen or steroid medication; renal problems requiring dialysis; anemia (Hgb,10 g/dL); hypothyroidism or adrenal insufficiency; and uncontrolled diabetes mellitus (fasting blood sugar [FBS] > 250 mg/dL). Other exclusion criteria included nutrition problems (30 kg/m² < body mass index [BMI] < 20 kg/m² and/or albumin < 3.2 g/dL), significant unexplained weight change, dementia (Mini-Mental State Examination [MMSE] score < 15), major depression (Geriatric Depression Scale [GDS] score > 5), poor vision (unable to read size 18 font), orthopedic problems that would limit or be aggravated by exercise, or more than six falls in the last 6 months.

Procedures were approved by the Committees for the Protection of Human Subjects of the Miami Jewish Home and Hospital and the University of Miami. Participants signed an informed consent before participation.

**Test Procedure**

Participants were instructed to wear comfortable clothing and walking or running shoes for testing. They were weighed, clothed wearing shoes, on a Detecto double beam scale. The scale had been calibrated against known weights and certified accurate. Trials were performed on a standard concrete access ramp (slope of 1:12) 1.5-m wide. A 3-m platform was at the top of the ramp allowing gradual deceleration; however, participants never reached speeds requiring this distance. Times were assessed to the nearest hundredth of a second using a digital timing system (Lafayette Instruments, Lafayette, IN) interfaced with two portable 43.2 × 70.0 cm pressure-sensitive pads weighing < 1 kg (Tapeswitch Corporation, Farmingdale, NY). Pads were 3.79 m apart, and the vertical distance was 33 cm. A starting line was drawn 4 m from the front edge of the first pad. Participants were instructed to walk or run up the ramp as quickly as possible and touch each pad. They performed one practice and two actual trials. Each participant was allowed a 2-minute recovery between trials. Longer recoveries were allowed but seldom used. Participants successfully completed all the tests.

Power output was computed using the formula (see Figure 1):

\[
\text{Power (W)} = \frac{\text{mass (kg)} \times \text{distance (m)} \times 9.8 \text{ W/kg} \cdot \text{m}^{-1}}{\text{time (s)}}
\]

where mass = body mass in kg, distance = vertical distance between pads in meters, time = time necessary to complete the task in seconds, and 9.8 W/kg-m s⁻¹ = acceleration due to gravity. The better of the two test trials was used for statistical analysis.

**Testers**

Data were collected by students under the supervision of the study coordinator. Each student received individual setup and administration instructions. Students then practiced until procedures and results were consistent among testers. This usually took two to five trials.

**Validity**

Ramp test validity was visualized by plotting power outputs across age groups. Additionally, a subset of the sample (n = 59) volunteered to be tested using an isokinetic dynamometer (Biodex Corporation, Shirley, NY) to evaluate knee extensor and plantar flexor power. These muscle groups are critical to mobility (39). Isokinetic testing occurred within 1 week of the ramp tests. Isokinetics are reliable and valid muscular performance measures commonly used to evaluate strength and power in older persons (6,40–43). Our laboratory has been using isokinetics to assess power in older persons for more than a decade (14,15).

Participants were tested concentrically according to procedures in the Biodex manual. Gravitational corrections were made for limb weight. Before testing, participants were allowed three practice trials, two warm-up repetitions (50% and 75% perceived effort), and a short recovery. Verbal encouragement was provided throughout all sessions. Knee extensor test speeds were 1.05, 3.14, and 5.24 rad · s⁻¹. Plantar flexor speeds were 1.05, 3.14, and 4.72 rad · s⁻¹ (14). Three maximal contractions were performed at each speed with speeds randomized to reduce order effects.

**Reliability**

Another subsample (n = 59), drawn from our larger sample, volunteered for reliability testing. Two trials were performed on each of two test days falling within a 2-week period.

---

Figure 1. Schematic diagram of the ramp power test used in the current study. BW = body weight.
Statistical Analyses

The sample was separated by gender then stratified by age (60–64, 65–69, 70–74, 75–79, 80–84, and 85+ years). Age-group means were used to compute norms. Multiple t tests were used to detect differences among age groups. A Bonferroni adjustment was considered to adjust the probability level for the 15 pairwise t tests; however, because it increases type II error while reducing type I error, we used the false discovery rate (FDR) procedure developed by Benjamini and Hochberg (44).

Bland–Altman plots of the difference versus average scores between the ramp power test and isokinetic knee extension and ankle plantar flexion power data were also used to demonstrate the validity of the test (45).

Test–retest reliability between tests and between testing days was examined using intraclass correlation coefficients (ICC), mean differences, limits of agreement, and Cronbach’s alpha (46–48). Plots of between-test differences against test means were constructed to allow visualization of intertest reliability (45).

RESULTS

Participants

Participant characteristics are presented in Table 1. The subsamples used to assess test validity (42 women, 17 men: 73.2 ± 7.3 years, 72.7 ± 15.3 kg) and reliability (38 women, 21 men: 73.8 ± 6.1 years, 71.9 ± 18.0 kg) had characteristics similar to the sample (363 women, 157 men: 73.5 ± 6.7 years, 73.0 ± 14.9 kg).

Validity

Bar graphs representing means and standard deviations of the power values with best-fit lines and regression equations for women and men across age groups are shown in Figure 2, A and B, respectively. Additionally, the Bland–Altman plots demonstrate that only one or two points fell outside the 95% confidence interval for comparisons between isokinetic plantar flexion testing and the ramp power test and all points fell within the 95% confidence interval for comparisons between isokinetic knee extension testing and the ramp power test (see Figure 3A–F).

Reliability

Table 2 presents reliability analysis results across tests and days. The ICC across all comparisons was 0.966, between Test 1 and Test 2 on each day 0.979 and 0.947, respectively, and between Day 1 and Day 2 for each test 0.921 and 0.925. Cronbach’s alpha among all tests was .966, for Test 1 and Test 2 on each day were 0.979 and 0.947, respectively, and between Day 1 and Day 2 for each test 0.921 and 0.925. Bland–Altman analyses demonstrated good agreement among all tests with small mean differences (−0.08 to −0.5 seconds). Plots of test differences versus test means showed that most points fell within the 95% limits of agreement for between-test (Figure 4A and B) and

Table 1. Participant Characteristics

<table>
<thead>
<tr>
<th>Age Group</th>
<th>60–64 y</th>
<th>65–69 y</th>
<th>70–74 y</th>
<th>75–79 y</th>
<th>80–84 y</th>
<th>85+ y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>40</td>
<td>80</td>
<td>102</td>
<td>83</td>
<td>38</td>
<td>20</td>
</tr>
<tr>
<td>Age</td>
<td>62.03</td>
<td>66.90</td>
<td>71.93</td>
<td>77.10</td>
<td>81.74</td>
<td>88.67</td>
</tr>
<tr>
<td>SD</td>
<td>1.48</td>
<td>1.38</td>
<td>1.47</td>
<td>1.36</td>
<td>1.29</td>
<td>1.57</td>
</tr>
<tr>
<td>Weight</td>
<td>78.54</td>
<td>70.56</td>
<td>69.04</td>
<td>67.70</td>
<td>63.48</td>
<td>61.04</td>
</tr>
<tr>
<td>SD</td>
<td>16.85</td>
<td>13.84</td>
<td>13.05</td>
<td>12.66</td>
<td>9.81</td>
<td>11.77</td>
</tr>
<tr>
<td>Men</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>9</td>
<td>24</td>
<td>48</td>
<td>35</td>
<td>29</td>
<td>12</td>
</tr>
<tr>
<td>Age</td>
<td>62.14</td>
<td>67.52</td>
<td>72.15</td>
<td>76.56</td>
<td>81.52</td>
<td>87.25</td>
</tr>
<tr>
<td>SD</td>
<td>1.57</td>
<td>1.27</td>
<td>1.38</td>
<td>1.31</td>
<td>1.34</td>
<td>1.67</td>
</tr>
<tr>
<td>Weight</td>
<td>96.12</td>
<td>83.16</td>
<td>83.02</td>
<td>81.32</td>
<td>81.64</td>
<td>73.81</td>
</tr>
<tr>
<td>SD</td>
<td>18.21</td>
<td>10.17</td>
<td>11.95</td>
<td>13.13</td>
<td>12.20</td>
<td>11.39</td>
</tr>
</tbody>
</table>

Note: SD = standard deviation.
Figure 3. Scatter plots of difference versus average values for isokinetic ankle plantar flexion tests at 1.05, 3.14, and 4.72 rad \cdot s^{-1} and the ramp power test (A–C) and isokinetic knee extension tests at 1.05, 3.14, and 5.24 rad \cdot s^{-1} and the ramp power test (D–F) with 95% confidence interval (dashed lines).
between-day (Figure 5A and B) comparisons. All figures showed a systematic decline in difference as mean scores increased.

Table 3 presents power means and percentile norms by age for men and women, respectively. Table 4 shows multiple comparisons among age groups for women and men, respectively. For women, values could distinguish between all half-decade groups except for 75- to 79- and 80- to 84-year-olds. For men, significant differences were detected for 9 of the 15 comparisons.

**Table 2. Results for Reliability Between Different Tests for the Same Day and Different Days for the Same Test**

<table>
<thead>
<tr>
<th>Repeatability</th>
<th>ICC Coefficient</th>
<th>95% CI</th>
<th>d (^{-}) (s)</th>
<th>SE of d (^{-}) (s)</th>
<th>95% CI for d (^{-}) (s)</th>
<th>SD (_{diff}) (s)</th>
<th>95% CI (s)</th>
<th>Coefficient (r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day 1</td>
<td>0.979</td>
<td>0.964–0.987</td>
<td>−0.07</td>
<td>0.02</td>
<td>−0.11 to −0.04</td>
<td>0.13</td>
<td>−0.32 to 0.18</td>
<td>.979</td>
</tr>
<tr>
<td>Day 2</td>
<td>0.947</td>
<td>0.911–0.968</td>
<td>−0.05</td>
<td>0.02</td>
<td>−0.09 to −0.01</td>
<td>0.15</td>
<td>−0.35 to 0.25</td>
<td>.947</td>
</tr>
<tr>
<td>Test 1</td>
<td>0.921</td>
<td>0.867–0.953</td>
<td>−0.08</td>
<td>0.03</td>
<td>−0.14 to −0.03</td>
<td>0.22</td>
<td>−0.52 to 0.35</td>
<td>.921</td>
</tr>
<tr>
<td>Test 2</td>
<td>0.925</td>
<td>0.874–0.955</td>
<td>−0.06</td>
<td>0.03</td>
<td>−0.11 to −0.01</td>
<td>0.20</td>
<td>−0.45 to 0.33</td>
<td>.925</td>
</tr>
</tbody>
</table>

*Note: ICC = intraclass correlation coefficients; d \(^{-}\) = mean difference; CI = confidence interval; SE of d \(^{-}\) = standard error of the mean difference; 95% CI for d \(^{-}\) = 95% CI for the mean difference; SD \(_{diff}\) = standard deviation of the differences.*

Figure 4. Scatter plots of between-test difference against the average value for Test 1 and Test 2 on Days 1 (A) and 2 (B) with 95% confidence interval (dashed line).

Figure 5. Scatter plots of between-test difference against the average value on Day 1 and Day 2 for Tests 1 (A) and 2 (B) with 95% confidence interval (dashed line).
Safety

We performed >1200 tests during this study, with only one negative consequence—a slight hamstring pull, which did not require medical attention and resolved within 2 days.

DISCUSSION

Validity

The ramp test has face validity because it measures the time required to move a person’s body mass a specified distance against the force of gravity, which by definition is power. The test has content validity because it uses the same variables and power computation as three documented power tests (20).

Scatter plots of individual and mean values show the expected decline with age reported in muscle cross-sectional area (4), power (1), and functional capacity (1). The $r^2$ values for individual power outputs by women and men show the expected dispersion from the regression line (1,4,49), whereas mean values demonstrate a close fit to age-related patterns of decline. Finally, the concurrent validity of the ramp power test is demonstrated by the fact that only one or two points fell outside the 95% confidence interval for difference versus mean score for isokinetic plantar flexion power and no points fell outside the 95% confidence interval for the confidence interval for knee extension power.

Reliability

The ICC values between tests and testing days support the reliability of the ramp power test (50). The mean differences (range $-0.08–0.05$ seconds) and the standard deviation of the differences $[SD_{diff}]$ (range 0.13–0.22) between tests were low relative to the range of scores (0.51–2.65 seconds); the average $SD$ for the sample (.73 seconds) and the average difference scores between age groups (men: 0.62 seconds; women: 0.24 seconds) indicated good agreement for all comparisons. The plots of differences against means for all comparisons show that nearly all difference values (56 of 59) fell within the limits of agreement. Reliability is further supported by the high Chronbach’s alpha values among all test comparisons. The small negative slopes for the regression lines appear to be the combined results of improvements from Test 1 to Test 2 and Day 1 to Day 2 for the persons with lower power scores and lower performances by participants with higher power scores, whereas difference scores for the participants producing average power outputs clustered near zero. Further study would be necessary to determine the mechanisms that may have been responsible for this pattern; however, possibilities include a motor learning effect, a modification in performance strategy, a regression of both the high performances and low performances toward the mean, or the psychological impact of the initial performance on subsequent performances.

Ability to Discriminate Power Differences Among Age Groups

The discriminatory power among age groups was greater for women than for men. An obvious explanation is the smaller sample size for men. A second explanation is the

<table>
<thead>
<tr>
<th>Age Group</th>
<th>60–64 y</th>
<th>65–69 y</th>
<th>70–74 y</th>
<th>75–79 y</th>
<th>80–84 y</th>
<th>85+ y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60–64 y</td>
<td>—</td>
<td>.017*</td>
<td>&lt; .0001*</td>
<td>.006*</td>
<td>.001*</td>
<td>&lt; .0001*</td>
</tr>
<tr>
<td>65–69 y</td>
<td>—</td>
<td>—</td>
<td>.006*</td>
<td>&lt; .0001*</td>
<td>&lt; .0001*</td>
<td>&lt; .0001*</td>
</tr>
<tr>
<td>70–74 y</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.0002*</td>
<td>&lt; .0001*</td>
<td>&lt; .0001*</td>
</tr>
<tr>
<td>75–79 y</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.05 (NS)</td>
<td>.0003*</td>
</tr>
<tr>
<td>80–84 y</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.005*</td>
</tr>
<tr>
<td>85+ y</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Men</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60–64 y</td>
<td>—</td>
<td>.267 (NS)</td>
<td>.065 (NS)</td>
<td>.011*</td>
<td>.021*</td>
<td>.0003*</td>
</tr>
<tr>
<td>65–69 y</td>
<td>—</td>
<td>—</td>
<td>.051 (NS)</td>
<td>.002*</td>
<td>.004*</td>
<td>&lt; .0001*</td>
</tr>
<tr>
<td>70–74 y</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.101 (NS)</td>
<td>.114 (NS)</td>
<td>&lt; .0001*</td>
</tr>
<tr>
<td>75–79 y</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>.484 (NS)</td>
<td>&lt; .0001*</td>
</tr>
<tr>
<td>80–84 y</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>85+ y</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Notes: *Significance adjusted using method of Benjamini and Hochberg (1995) to control for false discovery rate (FDR). NS = not significant.
slower rate of power loss in men compared to women (1,17). A third explanation is that the most dramatic reductions in power for men occur after age 80 (1). We suggest that further examinations of the test’s discriminatory power be performed using larger sample sizes and/or more homogeneous samples.

**Practical Application**

A number of factors support the ramp power test as a diagnostic tool for older persons. First of all, a functional power test should incorporate a testing methodology similar to a commonly performed activity of daily living (ADL), incorporating muscle groups, joint angles, contractile properties, movement speeds, and ranges of motion similar to the activity being tested (51). For older individuals, it can be argued that mobility is the single most important factor in maintaining independence (52,53). Because the ramp test uses a walk or run up an ascending plane, it replicates a common ADL task and differs considerably from explosive jump (8,10,11,38), cycle ergometer (2,12,13), or isokinetic tests (6,14,15) that are not task-specific. Stair tests are movement-specific; however, recent results (36) indicate that the loads that optimize power for stair climbing, chair stands, and walking speed differ. Additionally, the kinematics and kinetics of stair climbing and walking differ (54–56), arguing that each activity requires its own specific power test.

The second factor to be considered is safety. As noted above, >1200 tests were performed during this study, with only a single negative consequence (which required neither medical care nor prolonged reduction in physical activity). A third factor supporting the ramp power test is the cost. The test requires only a standard access ramp and a timing device. Timing devices can range from a $20 stopwatch to a digital timer interfaced with touch pads or photocells costing about $1,000. Both options are more cost effective than other commonly used devices, costing many thousands of dollars to purchase and maintain. We suggest that a study be conducted comparing the data accuracy and reliability among timing devices for the current test, as well as for other field tests such as stair-climbing and walking speed tests to allow researchers and clinicians the opportunity to choose among these devices.

A fourth factor supporting this test is that it can be used with participants with diverse power outputs. Because the test uses a ramp, we suggest that further research be done examining the feasibility of using this test to measure power in patients with walkers, stroke patients, or other groups whose diminished physical capacity may limit their ability to perform field tests such as stair climb or jump tests.

The final factor is that this test is feasible for use in both community and clinical settings as it is simple to administer, uses only a portion of a standard access ramp, requires minutes to set up, and requires no permanent installation because the touch pads and timing device are portable and weigh < 3 kg.

Given the importance of power to mobility (39,57), independence (1,5,17), and probability of falling (6,34,58) in older persons, the ramp power test could be incorporated into any testing battery used to diagnose performance levels in this population.

**ACKNOWLEDGMENT**

We are grateful for support from the Teaching Nursing Home program, funded by the State of Florida Agency for Health Care Administration, during the data analysis and manuscript preparation portions of this project.

**CORRESPONDENCE**

Address correspondence to Joseph F. Signorile, PhD, Department of Exercise and Sport Sciences, University of Miami, P.O. Box 248965, Coral Gables, FL 33124. E-mail: jsignorile@miami.edu

**REFERENCES**


Received September 5, 2006
Accepted February 8, 2007
Decision Editor: Luigi Ferrucci, MD, PhD