Reliability and Validity of Responses to Submaximal All-Extremity Semirecumbent Exercise in Older Adults

Marissa E. Mendelsohn, Denise M. Connelly, Tom J. Overend, and Robert J. Petrella

Although popular in clinical settings, little is known about the utility of all-extremity semirecumbent exercise machines for research. Twenty-one community-dwelling older adults performed two exercise trials (three 4-min stages at increasing workloads) to evaluate the reliability and validity of exercise responses to submaximal all-extremity semirecumbent exercise (BioStep). Exercise responses were measured directly (Cosmed K4b²) and indirectly through software on the BioStep. Test–retest reliability (ICC²,1) was moderate to high across all three stages for directly measured METs (.92, .87, and .88) and HR (.91, .83, and .86). Concurrent criterion validity between the K4b² and BioStep MET values was moderate to very good across the three stages on both Day 1 (r = .86, .71, and .83) and Day 2 (r = .73, .87, and .72). All-extremity semirecumbent submaximal exercise elicited reliable and valid responses in our sample of older adults and thus can be considered a viable exercise mode.

Key Words: psychometric properties, community dwellers, oxygen consumption

Aerobic capacity reflects the overall function and health of the cardiovascular and respiratory systems. Measurement of maximal oxygen consumption (VO₂max) using inspired- and expired-gas analysis during work is considered the criterion standard for determining aerobic fitness. Maximal exercise tests require the use of sophisticated and expensive equipment, are very stressful on the individual being tested, require physician supervision, and are associated with a higher risk of adverse events (Jones, Vandervoort, & Overend, 2005). Measuring an accurate VO₂max might not be appropriate in older adults because of chronic medical conditions and physical limitations that make it difficult to meet the criteria for attaining VO₂max or might even prevent older adults from receiving medical clearance to be tested.

A comprehensive recent review by Huggett, Connelly, and Overend (2005) suggested the need to develop new testing protocols to adequately measure aerobic...
capacity in older adults. Aerobic fitness is often measured using traditional exercise machines such as stationary cycles, treadmills, and upper body ergometers. New exercise modalities, however, such as all-extremity semirecumbent trainers might be better suited to individuals who are non–weight bearing or have balance, coordination, or gait impairments or lower extremity weakness or pathology (Huggett et al.).

When maximal exercise testing is deemed inappropriate, submaximal exercise testing should be considered (Loudon, Cagle, Figoni, Nau, & Klein, 1998). Submaximal tests are easier to administer to a broader range of people, have lower risks of adverse events, require less time and expense, and are conducted at lower exercise intensities (typically around 70–85% of age-predicted maximum heart rate; American College of Sports Medicine [ACSM], 2000). Submaximal exercise tests such as the YMCA Protocol (Golding, Myers, & Sinning, 1989) and Astrand–Ryhming cycle-ergometer test (Astrand & Ryhming, 1954) can be administered on either arm or leg cycle ergometers.

Older adults, however, have indicated that they do not like the uncomfortable seating and the unfamiliar and nonfunctional movements required during leg cycle-ergometry tests (Looney & Rimmer, 2003). Furthermore, because only the lower extremity muscles are being used, localized leg-muscle fatigue often has a greater influence on test outcome than true aerobic capacity (Looney & Rimmer). Leg cycle ergometers have recently been redesigned so that testing can be conducted in a semireclined position and incorporate the upper body, such as the NuStep TRS 4000 recumbent cross-trainer (NuStep, Inc., Ann Arbor, MI) or the BioStep Clinical Pro semirecumbent elliptical cross-trainer (Biodex, Shirley, NY). Simultaneous use of the arms and legs during an aerobic-fitness test decreases the localized fatigue of arm or leg muscles in arm cranking or leg cycling exercise, respectively; decreases the likelihood of premature fatigue resulting in cessation of exercise; and eliminates the requirement for weight bearing in the legs (Looney & Rimmer).

Using all four extremities, patients might be more capable of maintaining exercise at any absolute work rate with lower ratings of perceived exertion. In addition, for rehabilitation, blood pressure can be more easily monitored because one arm can be released from the arm lever for measurement while the remaining three limbs continue the pedaling motion. Approximately 10% of nondisabled community-dwelling older adults over 75 years of age develop a need for assistance with basic activities of daily living (Katz et al., 1983; Manton, Corder, & Stallard, 1993). Seated submaximal exercise on a NuStep or BioStep machine might be useful as part of a “prehabilitation program” aimed at preventing disability and maintaining function in community-dwelling older adults.

These new exercise machines are marketed primarily to hospital rehabilitation centers, health clubs, and seniors’ fitness programs. Little is known about the utility of such all-extremity exercise machines for research purposes. Research involving all-extremity exercise might provide rehabilitation clinicians with methods of improving aerobic fitness in older adults who are unable to weight bear, have impaired balance or osteoarthritis, or are recovering from a hip fracture or stroke.

The purpose of the present study was to investigate, in older adults, the test–retest reliability and concurrent criterion validity of heart rate and submaximal-exercise oxygen consumption (VO$_2$) measured directly by a breath-by-breath
portable metabolic gas-analysis system and indirectly through values provided by software from an all-extremity semirecumbent exercise trainer.

Methods

Participants
Based on a sample-size calculation (Hulley et al., 2001), 21 healthy older adults (14 women age 63–89 years and 7 men age 65–86 years) were recruited from an exercise program (40-min aqua fitness program three times per week at approximately 60% of age-predicted maximal heart rate) located at a retirement home. Participants were included in the study if they were at least 60 years of age, lived in their own homes, scored more than 24 on the Mini-Mental State Exam (Folstein, Folstein, & McHugh, 1975; Tombaugh & McIntyre, 1992), and had physician approval to participate in the study. The ACSM (2000) guidelines for absolute and relative contraindications to exercise testing were used as exclusion criteria. These included a range of cardiac, neurological, and musculoskeletal symptoms and conditions. Each participant’s physician completed a checklist to confirm the absence of all exclusion criteria. The study was approved by the Health Sciences Research Ethics Board at the University of Western Ontario, and informed written consent was obtained from all participants before the study began.

Instrumentation
The BioStep is a hybrid between a recumbent bicycle and an elliptical exercise machine (Figure 1). Each arm handle is coupled with the contralateral foot pedal, allowing the user to exercise in a reciprocal pattern at a 1:1 ratio with the upper and lower extremities simultaneously. The upper extremities move forward and backward while the lower extremities follow an elliptical pattern. The BioStep provides 30 levels of resistance. Heart-rate monitoring via a telemetry system (Polar Vantage NV) provides a visual cue for maintaining an appropriate training intensity. Metabolic equivalent (MET) values provided by the BioStep software are based on the unit’s power output (in watts). The power output from the alternator is converted to a corresponding watt value that is then converted to an MET value as per the ACSM equations (ACSM, 2000).

\( \text{VO}_2 \), carbon-dioxide production, respiratory-exchange ratio, and heart rate (HR) were measured using a portable breath-by-breath metabolic unit (Cosmed K4b², Rome, Italy). The K4b² is a lightweight (925-g) system that is worn on the participant’s torso and includes a face mask and turbine flowmeter allowing real-time data collection of \( \text{VO}_2 \). A validation study comparing the Cosmed K4b² with the Douglas-bag method over a wide range of cycling exercise intensities reported no significant differences in \( \text{VO}_2 \) between the K4b² and the Douglas-bag method at rest or at 250 W, thus suggesting that this portable unit is acceptable for measuring \( \text{VO}_2 \) (McLaughlin, King, Howley, Bassett, & Ainsworth, 2001).

Participants were seated on the BioStep in a standardized position with a 5° knee bend at maximal leg extension and a 5° bend in the elbow at maximal arm extension. A metronome provided audible cues for a rate of 60 rpm during exercise; participants also used the cadence value from the BioStep display panel to monitor and adjust their push–pull rate with arms and legs.
Test–Retest Reliability
Participants completed two exercise test sessions, one week apart, using the same submaximal multistage exercise protocol on each day.

Concurrent Criterion Validity
For each participant, breath-by-breath VO\textsubscript{2} was measured at rest and during exercise on the BioStep. The VO\textsubscript{2} values obtained from the K4b\textsuperscript{2} during the exercise protocol were converted to MET values and then compared with the MET values obtained from the BioStep software to assess concurrent criterion validity.

Protocol for the Submaximal-Exercise Test
The exercise test was modeled after the YMCA Protocol (Golding et al., 1989). The current study lengthened each of three constant-work-rate exercise stages to 4 min, to allow more time for the older adults to reach a steady-state HR and VO\textsubscript{2} response at each of the exercise stages. The exercise protocol thus consisted of a 3-min warm-up, 4 min of exercise at each of three incremental workload levels, and a 5-min cool-down. Resting HR and blood-pressure values were obtained before the exercise test.

Practice Session. During the practice session, each participant completed one trial of the Berg Balance Scale (Berg, Wood-Dauphinee, Williams, & Gayton, 1989) and one practice and two trials of the timed up-and-go (TUG; Podsiadlo &
Richardson, 1991) to assess functional mobility. The participant then exercised on the BioStep while wearing the K4b2 and HR monitor to become familiar with the equipment, exercise protocol, and exercise motion. An age-predicted maximal heart rate (HR_{max}) for each participant was calculated using the equation 208 – (0.7 \times \text{age}; Tanaka, Monahan, & Seals, 2001). Approximate starting workload levels for the exercise test were determined for each participant from his or her HR response during the practice session.

**Exercise Testing.** During the 3-min warm-up (Resistance Level 1), the participants were instructed to exercise on the BioStep while keeping time with the metronome (60 rpm). When the warm-up period was complete, the work rate was set for the first exercise stage and the participant exercised for 4 min. Participants then exercised through the second and third 4-min exercise stages at incremental workload levels. HR was recorded for each minute throughout the exercise. Participants were asked to give a rating of perceived exertion, using the Borg 0–10 category-ratio scale, CR10 (ACSM, 2000; Borg, 1982), at the end of the second minute in each 4-min stage of exercise. Participant performance (HR response and rating of perceived exertion) during the exercise stages informed the investigators about the subsequent workload levels required to obtain an HR response equal to 75% of the participant’s predicted HR_{max} at the end of the third stage. The criteria for terminating an exercise test included participants’ inability to maintain HR at their 75% HR_{max} target, inability to maintain a cadence of 60 rpm, volitional stopping, or any of the ACSM guidelines for stopping an exercise test (ACSM, 2000).

In the cool-down period, the workload was reduced to the warm-up workload level and the participant was instructed to exercise slowly for 5 min. The recovery stage was extended if HR had not returned to within five beats of resting HR within 5 min. Blood pressure was measured on completion of the cool-down.

**Data Reduction**

Breath-by-breath VO_{2} data (ml O_{2} · kg^{-1} · min^{-1}) from the K4b2 were averaged over the last minute of each 4-min exercise stage. MET values were calculated from the last minute of averaged VO_{2} values (METs = VO_{2} ÷ 3.5). For the concurrent validity analysis, MET values were also recorded from the BioStep panel display.

**Statistical Analyses**

**Descriptive Data for Participants.** All data were analyzed using SPSS-PC (Version 11.0; SPSS, Chicago). Descriptive statistics were compiled for age, body-mass index, health conditions, daily medications, residence, gait aids, and the Berg Balance Scale. Paired t tests were performed on the TUG scores to determine whether a significant difference occurred between Trials 1 and 2.

Separate one-way repeated-measures ANOVAs were conducted to determine differences in resting values for HR or blood pressure between Day 1 and Day 2. Two-way repeated-measures ANOVAs were carried out separately for Day 1 and Day 2 to determine whether MET values from the K4b2 and the BioStep were different across the three stages.
Test–Retest Reliability. Two-way repeated-measures ANOVAs were completed for each of METs, HR, and watts to determine whether these variables increased with the progressively higher workloads across the three 4-min stages of exercise on Day 1 and Day 2. The level of significance was set at \( p < .05 \) for all comparisons. All data are reported as \( M \pm SD \).

Test–retest reliability of the BioStep MET data was determined using intraclass correlation coefficients (ICC\(_{2,1}\)) calculated from results of the two-way repeated-measures ANOVAs. ICC values below .80 are questionable for physiological data, values of .80 to .89 indicate moderate reliability, and values greater than .90 are considered high (Vincent, 2005).

Concurrent Criterion Validity. Concurrent criterion validity of the MET values from the K4b\(^2\) and the BioStep for each of the three exercise stages was determined separately for Day 1 and Day 2 using Pearson correlation coefficients \( (r) \). The strength of the association between the MET values from the BioStep and the K4b\(^2\) was characterized according to Colton (1974): 0–.25, no relationship; .25–.50, fair; .50–.75, moderate to good; >.75, very good to excellent.

Results

Descriptive Data for Participants

Data were collected for 21 participants (14 women age 63–89 years and 7 men age 65–86 years). Chronic conditions reported by the participants included high blood pressure \( (n = 6) \), high cholesterol \( (n = 4) \), arthritis \( (n = 4) \), osteoporosis \( (n = 3) \), osteoarthritis \( (n = 2) \), diabetes \( (n = 2) \), and hyperthyroidism \( (n = 1) \). Participants took a daily average of 2.0 ± 2.1 prescription medications (range 0–9). Their Berg Balance Scale scores were within the range of values (29–56) found in previous work with community-dwelling older adults (Berg, Wood-Dauphinee, Williams, & Maki, 1992), and their mean TUG scores (<10 s) indicated that they were independent with daily activities and going outside alone (Podsiadlo & Richardson, 1991). One participant used a cane. Eleven of the participants lived at home with family, 5 lived at home alone, and 3 lived alone and 2 with a spouse in a retirement-home setting. The average TUG score was reported because there was no significant difference between the two TUG trials (Table 1).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>( M \pm SD )</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>75.1 ± 7.7</td>
<td>63–89</td>
</tr>
<tr>
<td>Body-mass index (kg/m(^2))</td>
<td>26.2 ± 2.9</td>
<td>21.1–34.8</td>
</tr>
<tr>
<td>Berg Balance Scale score (maximal score 56)</td>
<td>52.4 ± 4.5</td>
<td>36–56</td>
</tr>
<tr>
<td>Timed up-and-go score (s)</td>
<td>8.9 ± 1.4</td>
<td>6.3–11.8</td>
</tr>
</tbody>
</table>
Exercise Intensities

There were no significant differences between resting values for HR, blood pressure, or respiratory-exchange ratio recorded on Day 1 and Day 2 (Table 2). As expected, HR, MET values, and watts increased significantly ($p < .05$) with the increase in workloads across the three 4-min stages of exercise in both test sessions. A stepwise increase in HR and rating of perceived exertion was achieved across the stages, with participants achieving 75% of their age-predicted HR_max at the end of the third stage (Table 3). Extrapolation of the HR data suggested that the participants reached 60–65% of their predicted VO_{2max} at the end of the third stage of exercise. Figure 2 illustrates the steady-state exercise MET values (K4b²) achieved on Day 1 during each of the three 4-min stages of exercise.

Test–Retest Reliability

The average BioStep workload levels for the participant group across the three 4-min exercise stages were $2.5 \pm 1.12$ (range 2–6), $3.5 \pm 1.12$ (range 3–7), and $4.6 \pm 1.9$ (range 4–9), while the average watt levels were $36.8 \pm 13.5$, $44.0 \pm 15.4$, and $53.4 \pm 18.8$ on Day 1 and $39.6 \pm 15.8$, $46.2 \pm 18.3$, and $56.4 \pm 21.7$ on Day 2. There were no significant differences between watt levels on Day 1 versus Day 2 for the three stages of exercise.

Test–retest reliability of the MET and HR values obtained from the BioStep was reported as the reliability of two single measures on two separate test occasions (ICC_{2,1}). Reliability of directly measured (K4b²) METs data was moderate to high across the three stages (ICC_{2,1} = .92, .87, and .88). The ICC_{2,1} values for HR were also moderate to high across the three stages (.91, .83, and .86).

Table 2  Baseline Values ($M \pm SD$) for Resting Heart Rate and Blood Pressure, $N = 21$

<table>
<thead>
<tr>
<th>Variable</th>
<th>Day 1</th>
<th>Day 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart rate (beats/min)</td>
<td>71.8 ± 9.3</td>
<td>70.5 ± 6.7</td>
</tr>
<tr>
<td>Systolic blood pressure (mmHg)</td>
<td>142.5 ± 15.1</td>
<td>135.7 ± 14.3</td>
</tr>
<tr>
<td>Diastolic blood pressure (mmHg)</td>
<td>76.9 ± 9.1</td>
<td>74.9 ± 8.2</td>
</tr>
</tbody>
</table>

Table 3  Age-Predicted Maximal Heart Rate (HR_max) and Ratings of Perceived Exertion at Three Levels of Submaximal Exercise on Two Test Occasions ($M \pm SD$), $N = 21$

<table>
<thead>
<tr>
<th>Stage</th>
<th>Age-predicted HR_max</th>
<th>Ratings of perceived exertion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Day 1</td>
<td>Day 2</td>
</tr>
<tr>
<td>1</td>
<td>64.6% ± 2.6%</td>
<td>2.5 ± .81</td>
</tr>
<tr>
<td>2</td>
<td>69.9% ± 2.0%</td>
<td>3.5 ± .92</td>
</tr>
<tr>
<td>3</td>
<td>75.9% ± 2.4%</td>
<td>4.3 ± 1.3</td>
</tr>
<tr>
<td></td>
<td>Day 1</td>
<td>Day 2</td>
</tr>
<tr>
<td>1</td>
<td>65.5% ± 2.3%</td>
<td>2.4 ± .67</td>
</tr>
<tr>
<td>2</td>
<td>70.4% ± 2.0%</td>
<td>3.3 ± 1.0</td>
</tr>
<tr>
<td>3</td>
<td>75.9% ± 2.3%</td>
<td>4.6 ± 1.1</td>
</tr>
</tbody>
</table>
Concurrent Criterion Validity

There were no significant differences between Day 1 and Day 2 BioStep and K4b² MET values for any of the three exercise stages (Table 4). Correlations between the BioStep MET values and K4b² MET values across the three stages on both Day 1 \((r = .86, .71, \text{ and } .83)\) and on Day 2 \((r = .73, .87, \text{ and } .72)\) ranged between moderate to good and very good to excellent, respectively. The BioStep slightly and systematically underestimated MET values compared with the K4b², but the differences were not statistically significant (Table 4).

![Graph showing heart rate and MET values](image)

**Figure 2** — Group heart-rate and MET (metabolic equivalent) values \((M \pm SD)\) collected by the K4b² \((N = 21)\) at each minute of submaximal exercise during the three 4-min stages on Day 1. Stage 1 included Minutes 1–4, Stage 2 included Minutes 5–8, and Stage 3 included Minutes 9–12.

**Table 4** Comparison and Error Between the BioStep and K4b² MET Values at Three Levels of Submaximal Exercise on Two Test Occasions \((M \pm SD), N = 21\)

<table>
<thead>
<tr>
<th>Stage</th>
<th>BioStep METs</th>
<th>K4b² METs</th>
<th>Error</th>
<th>BioStep METs</th>
<th>K4b² METs</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.87 ± 0.54</td>
<td>2.99 ± 0.57</td>
<td>4.0%</td>
<td>2.98 ± 0.74</td>
<td>3.14 ± 0.70</td>
<td>5.1%</td>
</tr>
<tr>
<td>2</td>
<td>3.41 ± 0.63</td>
<td>3.57 ± 0.63</td>
<td>4.5%</td>
<td>3.40 ± 0.77</td>
<td>3.52 ± 0.78</td>
<td>3.4%</td>
</tr>
<tr>
<td>3</td>
<td>3.69 ± 0.80</td>
<td>3.86 ± 0.85</td>
<td>4.4%</td>
<td>3.83 ± 0.87</td>
<td>4.08 ± 0.92</td>
<td>6.1%</td>
</tr>
</tbody>
</table>

*Note. MET = metabolic equivalent.*
Discussion

This is the first study to our knowledge that has examined the test–retest reliability and concurrent criterion validity of \( \text{VO}_2 \) on the BioStep in community-dwelling older adults. Our results provide preliminary evidence of reliable and valid estimates of oxygen consumption by the BioStep when compared with the criterion standard of direct \( \text{VO}_2 \) measurement in our sample.

An all-extremity machine such as the BioStep might thus provide a useful alternative exercise modality for rehabilitation, aerobic training, and testing in older adults. Loudon et al. (1998) developed a linear-regression equation to determine the relationship between \( \text{VO}_{2\text{peak}} \) values measured during a maximal treadmill test and a maximal all-extremity ergometer test (Pro II Power Trainer) in healthy women age 30–60 years. Treadmill \( \text{VO}_{2\text{peak}} \) values correlated highly with the all-extremity \( \text{VO}_{2\text{peak}} \) values \((r = .918)\) and were not significantly different, suggesting that an all-extremity test might be a reliable and valid estimator of \( \text{VO}_{2\text{peak}} \) (Loudon et al.). Therefore, all-extremity exercise might be a reliable and valid mode of testing aerobic fitness and a viable alternative to using a treadmill. Loudon et al. also suggested the need for the development of submaximal tests to determine \( \text{VO}_{2\text{peak}} \) for clinical populations.

Using a combined upper and lower limb ergometer (Power Trainer), Hill, Ethans, MacLeod, Harrison, and Matheson (2005) assessed functional cardiopulmonary-exercise tolerance in patients with stroke. The results suggested that the all-extremity Power Trainer is a safe exercise machine on which to assess cardiopulmonary fitness in patients with neurological impairments. Hill et al. also suggested that proper exercise prescription using the all-extremity trainer might be beneficial to high-risk populations.

This study found moderate to very good concurrent criterion validity between MET values reported by the BioStep and those measured by the K4b2 on both Day 1 and Day 2. The results indicated that the BioStep slightly underestimated MET values when compared with the criterion standard direct measurement (K4b2). Although the percentage difference, which ranged from 3.4% to 6.1% (0.42–0.88 ml · kg\(^{-1}\) · min\(^{-1}\)), is not likely to make a practical difference in exercise performance or prescription, the MET values provided by the BioStep can be used as a guide for exercise but not for stringent assessment purposes.

Huggett et al. (2005) indicated that there are limited data on \( \text{VO}_{2\text{max}} \) testing in older adults over the age of 65 years and suggested the need to refine existing criteria for \( \text{VO}_{2\text{max}} \) testing in order to make these criteria more applicable to older adults, as well as the need to develop new protocols for testing aerobic capacity. An all-extremity exercise machine such as the BioStep might provide an alternative modality for aerobic testing in older adults. Although maximal-exercise testing is considered the criterion standard for measuring \( \text{VO}_{2\text{max}} \), such testing might be inappropriate or not possible for older adults because of surgery, pain, fatigue, frailty, and chronic disease. Many submaximal tests have been developed on traditional exercise equipment such as the treadmill and the cycle ergometer, but such testing protocols have until now not been used with newer equipment such as an all-extremity exercise machine. For these reasons, the BioStep might also provide an alternative mode of exercise training in a “prehabilitation” setting for older adults.

A limitation of the study was that the sample was a relatively small and homogeneous group of community-dwelling older adults, so our results are not
representative of all older adults. A second limitation was that the workloads during Exercise Stage 3 ranged only from 4 to 9 out of the 30 available workload levels on the BioStep. In the tested range of workloads, the BioStep provided valid MET values when compared with a direct measurement of VO₂. Future studies, however, should determine the reliability and validity of the BioStep at higher workload levels in order to assess the full range of available workloads. Examining the reliability of this device on a younger population exercising at higher intensities might also determine whether there is an increase in variance as the workload increases.

Conclusions

In summary, our sample of healthy older adults provided preliminary evidence to support the use of the BioStep for testing aerobic fitness, albeit in a narrow workload range. Although the BioStep appears to be an alternative exercise modality to the treadmill or cycle ergometer, different all-extremity exercise machines might not provide the same results. Future research should assess the reliability and validity of the BioStep in specific populations of older adults including people with higher fitness levels, frail people, and those with cardiac, musculoskeletal, or neurological impairment.

All-extremity semirecumbent exercise elicited reliable and valid HR and VO₂ responses at submaximal exercise intensities in our sample of healthy older adults. It can thus be considered a viable alternative mode of exercise for this population, who commonly cannot use traditional exercise ergometers.

Acknowledgments

We thank the study volunteers and staff at Chelsey Park Health Club and Canada Foundation for Innovation (CFI), as well as Biodex Medical Systems for the use of the BioStep.

References


