Journal of Aging and Physical Activity, 1998, **6**, 363-375

The Reliability and Validity of a 6-Minute Walk Test as a Measure of Physical Endurance in Older Adults

Roberta E. Rikli and C. Jessie Jones

The purpose of this study was to assess the reliability and validity of a 6-min walk test as a measure of physical endurance in older adults. Seventy-seven subjects, ages 60–87, performed three separate 6-min walk tests and a treadmill test and completed questionnaire items assessing physical activity level and functional status. The 6-min walk had good test–retest reliability ($.88 < R < .94$), particularly when a practice trial preceded the test trial. Convergent validity of the 6-min walk was demonstrated by its moderate correlation ($.71 < r < .82$) with treadmill performance. Construct validity was assessed by determining the ability of the test to detect differences between different age and activity level groups. As expected, walking scores decreased significantly across decades and were significantly lower for low-active subjects compared to high-active subjects. There was a moderate relationship between 6-min walk scores and self-reported functional ability. It was concluded that the 6-min walk can be used to obtain reasonably reliable and valid measures of physical endurance in older adults and that it moderately reflects overall physical functional performance.

Key Words: functional performance, aging, aerobic fitness

Maintaining adequate physical endurance in later years is important to prevent frailty and loss of independence. Considering the typical decline in aerobic power ($\dot{V}O_2$peak) of approximately 5 $ml \cdot kg^{-1} \cdot min^{-1}$ per decade, many inactive older adults will reach or surpass the threshold required for independent living by their late 70s or early 80s unless steps are taken to reduce the rate of decline (Jackson et al., 1995, 1996; Shephard, 1997). Evaluating and managing physical decline during aging are limited by a lack of suitable measurement tools, particularly measures that can be administered in a nonlaboratory setting to older adults with a wide range of ability levels.

Physical endurance, defined as the ability to maintain submaximal aerobic exercise for an extended time (American College of Sports Medicine [ACSM], 1997), generally is considered to be more relevant to older adult functioning than are common maximal measures such as $\dot{V}O_2$max or $\dot{V}O_2$peak. In fact, several studies have revealed that oxygen consumption measures, whether based on

The authors are with the Division of Kinesiology and Health Promotion, California State University, Fullerton, Fullerton, CA 92834.

maximal or submaximal performance, have questionable relevance to the demands of everyday activities for older adults and do not correlate well to measures of everyday functioning (Bittner et al., 1993; Guyatt et al., 1985b; Steele, 1996). Evidence also suggests that, with training, submaximal physical endurance can be improved even though no changes are observed in $\dot{V}O_2$max or $\dot{V}O_2$peak (ACSM, 1997). Certainly, submaximal endurance is important in many everyday activities such as shopping or other errands or participating in recreational activities.

Existing submaximal measures of physical endurance, such as 1-mile or 1/2-mile walk tests, can assess performance in young people and in high functioning older adults (Bravo et al., 1994; Fenstermaker, Plowman, & Looney, 1992; Kline et al., 1987; Osness et al., 1996) but are too strenuous for many older people who are incapable of walking the required distances for these tests. Statistics indicate that 40% of community-residing adults over 75 have difficulty walking even 1/4 mile (Select Committee on Aging, 1992).

A time-constant walking test (e.g., a 6-min walk), as compared to a distance-constant test such as the 1/2- or 1-mile test, can be used to measure the performance of all older adults who are ambulatory, regardless of their ability level. With time-constant tests, scores can be obtained for individuals who can walk only a few feet as well as for those who can cover several hundred yards within the time period. To date, however, validation studies on timed walking tests for older adults have been conducted primarily on populations with specific medical conditions such as pulmonary disease (Butland, Pang, Gross, Woodcock, & Geddes, 1982; Knox, Morrison, & Muers, 1988), heart disease (Bittner et al., 1993; Guyatt et al., 1985a), and osteoarthritis (Peloquin, Gauthier, Bravo, Lacombe, & Billiard, 1998). The one study we found that evaluated timed walking performance of nondiseased individuals involved only a small sample ($N = 16$) of very elderly subjects (mean age 86.8 years) (Stillwell, Forman, McElwain, Simpson, & Garber, 1996). Although data from the Stillwell et al. study revealed a moderately high correlation between 6-min walking scores and maximal aerobic capacity ($r = .76$), further research is needed before the 6-min walk can be considered a viable measure of physical performance in the general population of older adults.

The purpose of this study was to test the effectiveness of a 6-min walk as a measure of physical endurance in a sizable sample of generally healthy, community-residing adults over age 60. Specifically, the aims of the study were (a) to determine the test–retest reliability of the 6-min walk, and (b) to provide support for the test's validity by conducting several different types of comparative analyses (no single criterion measure exists for evaluating submaximal physical endurance in older adults). Specifically, as an indication of convergent validity (Safrit & Wood, 1995), scores on the 6-min walk were compared to treadmill walking performance using a modified Balke protocol (ACSM, 1991). To assess the construct validity of the test, 6-min walk scores of subjects from various age categories (60s, 70s, and 80s) and from different activity levels (high active vs. low active) were compared. The ability of a test to discriminate among various groups with known or presumed differences in the construct of interest indicates the construct validity of a test (Morrow, Jackson, Disch, & Mood, 1995; Rikli & Jones, 1997). Walking scores were also compared to self-reported overall functional ability to indicate the degree to which walking endurance reflects general physical functioning.

Methods

SUBJECTS

Participants were 77 volunteers (48 females and 29 males, mean age 73.1 years, *SD* 7.2) who were solicited from a nearby senior housing complex and from a university-sponsored exercise program. All subjects signed an informed consent and completed a written questionnaire concerning their health, functional status, and physical activity level. To be included in the study, participants had to be over age 60, community residing, functionally independent, and ambulatory (without the use of assistive devices); they could have no medical conditions that would contraindicate submaximal testing according to ACSM guidelines (ACSM, 1995); and they could not have been advised by their physician to refrain from exercise for health reasons. Medical clearance was obtained only for the subset of subjects who volunteered to participate in the treadmill testing.

PROCEDURES

During the first 2 weeks of the study, all participants completed the written questionnaire and performed three separate 6-min walking tests, with each test scheduled 2 to 5 days apart. A subsample of 37 volunteers (20 females and 17 males) completed submaximal treadmill tests during Weeks 3 and 4 of the study. Subjects were asked not to engage in strenuous exercise 24 hr prior to testing and to avoid smoking, heavy eating, or consuming caffeine within 2 hr prior to the study. Descriptions of the various tests and testing protocols are as follows:

6-Min Walk. Prior to each of the three walking tests, participants performed an 8-min warm-up and static stretch routine emphasizing the trunk and lower body. Walks were conducted outdoors (temperature range 68–75 °F) on a flat 50-yard rectangular course, marked off in 5-yard segments. Bright-colored cones were used to mark the four corners of the course. Participants were tested in groups (3 to 6 at a time), using 10-s staggered starting and stopping times to promote individual pacing and to discourage walking together in groups or pairs. Subjects were instructed to walk as fast as they comfortably could, trying to cover the maximum distance possible without overexerting or pushing themselves beyond what they thought was safe. Encouragement phrases such as "you are doing well" and "keep up the good work" were called out at 30-s intervals by the test administrator. The importance of standardized encouragement protocols relative to scoring accuracy in walking tests has been well-documented (Guyatt et al., 1985a; Steele, 1996). To assist with pacing, elapsed time was announced when the subjects were approximately halfway through the test, when they had 2 min left, and when they had 1 min left.

At the signal stop, participants were instructed to step aside and wait for the test administrator to note the yardage and record the score to the nearest 5-yard mark. The score was the total yards walked, determined by having an observer record a hash mark on the score card for each completed lap, then multiplying the total number of completed laps by 50 and adding the additional yards covered. If necessary, participants were allowed to stop and rest (sitting in chairs that had been positioned along the course) and then resume walking within the 6-min time period.

Rate of perceived exertion (RPE) scores utilizing a 20-point scale (Borg, 1970) were recorded within the last 2 min of the test and immediately following the test. Each participant's heart rate was monitored using Polar heart monitors to provide additional information about exercise intensity.

Treadmill Testing. Treadmill walking performance was assessed using a modified Balke graded exercise test recommended for older adults (ACSM, 1991; Spirduso, 1995). Prior to testing, all participants completed a 5-min general warm-up consisting of light calisthenics and stretching, plus a 3-min warm-up on the treadmill at 0.0% grade and 1.5 mph. The test began with the grade at 0.0% and the speed at 2.0 mph. At the end of each 3-min interval, the grade was increased by 3.5%. All participants had prior treadmill experience and were able to walk comfortably without using the handrails for support. Heart rates (from Polar heart monitors) and RPE values were recorded at the end of each 3-min period. The test was stopped when 85% of predicted maximal heart rate (PMHR = 220 — age) was reached or when the subject asked to stop or showed signs of overexertion. The score was the total time on the treadmill prior to reaching the 85% PMHR endpoint. A preliminary pilot study involving 46 older adults (mean age 69 years, *SD* 5.7) indicated that the intraclass test–retest reliability for treadmill walking time (time to 85% PMHR) was .87, using the formula for single-trail estimation (Baumgartner & Jackson, 1995).

Functional Ability. Physical functional ability was assessed through self-report using a composite physical function (CPF) scale developed for this study. The CPF scale (Table 1), which contains a wide range of functional items, is an adaptation and extension of three shorter, previously published scales: a 4-item scale by Rosow and Breslau (1966) and the 5-item and 6-item scales developed by Siu, Reuben, and Hays (1990). Also, three of the CPF items were taken from the National Health Interview Survey (National Center for Health Statistics, 1991). Notations in Table 1 indicate which of the CPF items have been adapted from each source. The newly developed CPF 12-item scale, similar to the Rosow-Breslau (1966) and Siu et al. (1990) scales, was designed to assess and discriminate across wide ranges of functional abilities, that is, abilities ranging from those associated with basic activities of daily living (ADLs; e.g., dressing and bathing oneself), to instrumental or intermediate ADLs (IADLs; housework, shopping, etc.), to advanced activities such as strenuous sports/exercise activities. The somewhat longer CPF scale differs from the Rosow-Breslau and Siu et al. scales in that the CPF includes items that reflect upper body function (i.e., Items i and j: ability to lift 10 lb and 25 lb, respectively; see Table 1), items that reflect greater variations in walking ability (i.e., Items d, g, and h), and two items, instead of one, concerning the ability to perform household tasks (Items e and k). The expanded CPF scale, with its increased number of items, should improve the discrimination power of the test as well as its ability to provide meaningful feedback to evaluators and to clients.

The scoring protocol for the CPF required that participants check one of three responses associated with each of the 12 items—*can do*, *can do with difficulty or with assistance*, or *cannot do*. A score of 2, 1, or 0, respectively, was assigned to the responses. Scores were summed over the 12 items, resulting in a potential range of scores from 24, indicating full function on all items, to 0 (unable to perform any of the items). The single-trial intraclass test–retest reliability estimate for this scale

Table 1 A Composite Physical Function (CPF) Scale

Instructions: Please indicate your ability to do each of the following. (Your response should indicate whether you are able to do these activities, not if you actually do the activities):

	Can do	Can do with difficulty or with help	Cannot do
a. Take care of own personal needs—like dressing yourself[a]	2	1	0
b. Bathe yourself, using tub or shower[a]	2	1	0
c. Climb up and down a flight of stairs[b] (like to a second story in a house)	2	1	0
d. Walk outside (one or two blocks)	2	1	0
e. Do light household chores—like cooking, dusting, washing dishes, sweeping a walkway	2	1	0
f. Shop for groceries or clothes[a,b]	2	1	0
g. Walk 1/2 mile (6-7 blocks)[a,b]	2	1	0
h. Walk 1 mile (12-14 blocks)[c]	2	1	0
i. Lift and carry 10 lb (full bag of groceries)[c]	2	1	0
j. Lift and carry 25 lb (medium to large suitcase)[c]	2	1	0
k. Do heavy household activities—like scrubbing floors, vacuuming, raking leaves[a,b]	2	1	0
l. Do strenuous activities—like hiking, digging in garden, moving heavy objects, bicycling, aerobic dance activities, strenuous calisthenics, etc.[a]	2	1	0

[a]Adapted from the 5- and/or 6-item scales published by Siu et al. (1990). [b]Adapted from the Rosow-Breslau Scale (1966). [c]Adapted from the National Health Interview Survey (National Center for Health Statistics, 1991).

Table 2 Correlations (r) and 95% Confidence Intervals for Comparing CPF Scores to Other Measures of Functional Performance

	CPF scores	
	r	95% CI
Rosow-Breslau	.96	.94–.98
5-item Siu et al.	.92	.88–.95
6-item Siu et al.	.95	.92–.97
Treadmill performance	.69	.48–.81

was .94, computed on scores obtained 2 to 4 weeks apart from 66 older adults (42 females, 24 males, mean age 71.6 years, *SD* 5.7).

The validity of the CPF was addressed in three ways: (a) by determining its correlation with previously published scales (a measure of convergent validity), (b) by comparing CPF scores with treadmill performance, a common criterion measure of functional capacity (ACSM, 1995), and (c) by looking at the sensitivity of the test with respect to identifying hierarchical levels of functional ability (an indication of discriminant validity).

In support of convergent validity, scores on the expanded CPF scale were highly correlated ($.92 < r < .96$, see Table 2) with previously validated scales, that is, the Rosow-Breslau Scale (Reuben, Siu, & Kimpau, 1992; Rosow & Breslau, 1966; Siu et al., 1990) and the 5-item and 6-item scales developed by Siu et al. (1990). The moderate correlation ($r = .69$) between CPF scores and treadmill performance, as described in the preceding section of this paper, supported the criterion validity of the test. As an indication of discriminant validity of the test, *t*-test analysis showed that the 26 subjects who were classified as having advanced functional ability (those receiving the maximum CFP score of 24) had significantly higher treadmill scores ($t = 2.26$, $p < .05$) than the 16 subjects classified as having intermediate functional ability (those who scored between 18 and 23). Specifically, treadmill scores (time on treadmill to 85% PMHR) for the advanced functional subjects were 14.3 min ($SD = 12.9$) versus 11.3 min ($SD = 17.7$) for those identified as having an intermediate level of functional ability. In this study, most subjects classified below the intermediate level of functional ability were either unwilling or unable to obtain medical clearance for the treadmill test, thus preventing us from extending the treadmill comparisons to the lower functioning group.

Physical Activity Level. Physical activity level was assessed through self-report, using a simple questionnaire item that asked about the frequency of moderate to strenuous exercise habits. Specifically, the question in this study, which was taken from Paffenbarger's College Alumnus Questionnaire (Paffenbarger, Blair, Lee, & Hyde, 1993), asks the subject, "Do you currently participate in regular physical exercise that is strenuous enough to cause a noticeable increase in breathing, heart rate, and perspiration?" If the answer was yes, the participant was asked to indicate the number of days per week that he or she engaged in such activity. Although a "global" question such as this cannot provide detailed information about one's activity level or energy expenditure, numerous studies have shown that it is a reliable and valid way of classifying subjects into "high" versus "low" physical activity and fitness categories (Ainsworth, Montoye, & Leon, 1994; Ainsworth, Richardson, Jacobs, & Leon, 1992; Godin, Jobin, & Bouillon, 1986; Paffenbarger et al., 1993; Schectman, Barzilai, Rost, & Fisher, 1991; Siconolfi, Lasater, Snow, & Carleton, 1985). In fact, some have found that a single-item question such as this is a more accurate indication of general physical activity/fitness level than are the more detailed physical activity indexes (Ainsworth et al., 1994; Siconolfi et al., 1985).

Based on their questionnaire responses, subjects were categorized as either "high active" or "low active." High active subjects were those who reported participating in strenuous activity three or more times per week. Low active subjects either did not participate in strenuous activity (as defined in the question) or were active on an irregular basis ($\leq$2 times a week). To pilot test the reliability of the

scoring procedures, the questionnaire was administered to 63 subjects (41 females, 22 males, mean age 71.4 years, *SD* 5.1) on two different occasions, 2 to 4 weeks apart, and was evaluated by two different scorers. The degree of agreement from the first questionnaire to the second was high, with 89% of the subjects receiving the same physical activity classification on both tests.

DATA ANALYSIS

Test–retest reliability for the walking test was established by determining the intraclass correlation coefficient (R) for a single trial using one-way ANOVA procedures (Baumgartner & Jackson, 1995). As an indication of convergent validity, Pearson correlation analysis was used to compare 6-min walk scores with treadmill walking performance. Construct validity was addressed by determining (through ANOVA analyses) whether the 6-min walk test could discriminate among groups with presumed differences in ability level. Specifically, ANOVA procedures were used to compare walking scores of people in various age groups (i.e., those in their 60s, 70s, and 80s) and of subjects with different activity levels (i.e., high-active vs. low-active subjects). Correlation analysis also was used to compare 6-min walk scores with self-reported overall functional ability.

Results

On the 6-min walk test, high intraclass correlations for total subjects and for men and women separately were found between Trials 1 and 2 ($.88 < R < .94$) and between Trials 2 and 3 ($.91 < R < .97$), indicating that scores had good relative reliability across trials (i.e., scores within the group tended to be in the same relative position to each other from trial to trial). However, the ANOVA analysis revealed a significant overall improvement across trials for total subjects, $F(2, 130) = 15.6$, $p < .0001$, as well as for males only, $F(2, 44) = 7.6$, $p < .001$, and for females only, $F(2, 84) = 8.4$, $p < .001$, indicating that there was a lack of absolute consistency (i.e., a lack of stability of group means) over the three trials. Follow-up analyses (Tukey post hoc comparisons of group means) indicated that in all cases there were significant increases in mean walking distances between Trials 1 and 2 ($p < .01$) but not between Trials 2 and 3, suggesting that two trials (with the first trial considered a practice trial) are needed to achieve stable scores across trials—scores that do not continue to change as a result of practice. Trial mean scores, standard deviations, and test–retest correlation values are presented in Table 3.

The correlation between 6-min walk scores and treadmill performance (time to 85% PMRH) was moderate to moderately high for men and women ($r = .71$ and .82; see Table 4), supporting the convergent validity of the test. (The 6-min walk scores used throughout the validity analysis were those achieved during Trial 2 of the testing, the first of the three trials administered that produced both relative and absolute scoring stability.) The correlation between walking scores and self-reported functional ability was somewhat lower for both males and females ($r = .71$ and .61) but still highly significant ($p < .0001$). All r values and corresponding 95% confidence intervals are presented in Table 4.

Table 3 Mean Yards Walked, Standard Deviations, and Test–Retest Intraclass Correlation Estimates (*R*) for the 6-Min Walk

	Test 1		Test 2*		Test 3**		*R* (Tests 1 & 2)	*R* (Tests 2 & 3)
	M	*SD*	*M*	*SD*	*M*	*SD*		
Total ($n = 66$)	595.1	109.3	616.5	98.9	619.6	99.2	.91	.94
Men ($n = 23$)	619.3	128.2	641.4	119.5	650.2	115.2	.94	.97
Women (n = 43)	582.2	96.9	603.2	84.4	604.4	86.7	.88	.91

*Test 2 scores were significantly greater than Test 1 scores for all three comparisons: total subjects, males only, and females only ($p < .0001$). **There was no significant change from Test 2 to Test 3 for any of the three comparison groups: total subjects, males only, or females only.

Table 4 Correlations (*r*) and 95% Confidence Intervals (in Parentheses) for Comparing 6-Min Walk Scores With Treadmill Performance and Self-Reported Functional Ability

	Total *r*	Men *r*	Women *r*
Treadmill scores	$n = 37$ .78 (.62–.88)	$n = 17$ .82 (.52–.93)	$n = 20$ .71 (.40–.88)
Functional ability	$n = 77$ .63 (.48–.75)	$n = 29$ .71 (.47–.86)	$n = 48$ .61 (.38–.75)

Table 5 contains means, standard deviations, and *F* values for comparing 6-min walk scores of subjects of different age and activity level groups. Significant ANOVA *F* values (and Tukey post hoc comparisons) indicate that walking scores, as would be expected, declined significantly across all age groups—from the 60s to the 70s to the 80s ($p < .0001$). Also, as expected, walking scores were higher for subjects with self-reported high physical activity levels than for those with low levels of physical activity ($p < .0001$). To avoid cells with small *n* values, male and female scores were combined for the age group and activity level analyses. However, any gender differences in walking performance should not have affected the overall interpretation of results since the proportion of men to women was similar (approximately 35–39%) across the various age and activity group comparisons.

Table 5 Mean Scores (Yards Walked), Standard Deviations, and ANOVA *F* Values for the 6-Min Walk Test for Age Group and Activity Level Categories

	n	*M*	*SD*	*df*	*F*	*p*
Age groups				2, 74	10.7	<.0001
60–69 years	20	677.8	95.0			
70–79 years	37	621.0	82.4			
80–89 years	20	550.1	86.7			
Activity groups				1, 74	36.5	<.0001
High active	59	647.6	81.5			
Low active	17	513.2	77.9			

Other results indicated that RPE measures taken during the 6-min walk did not differ significantly from those taken during the treadmill test. The mean peak RPE (20-point scale) reported during the 6-min walk was 13.6 (*SD* 1.7) versus 13.9 (*SD* 1.9) during the treadmill test. However, heart rates recorded during the two tests did differ significantly ($p < .01$), with a mean peak heart rate of 133.8 (*SD* 18.5) observed during the walk test compared to 128.7 (*SD* 7.8) during the treadmill test. No signs of overexertion were observed for any subject during either of the two tests.

Summary and Discussion

Previously published walking tests, such as the 1/2- or 1-mile tests, are relatively good indicators of physical endurance in high functioning older adults. However, these protocols are not feasible for measuring across the full range of ability levels typically found in the community-dwelling population of older adults, many of whom cannot walk even 1/4 mile. Shorter walking tests, such as a 5-or 6-min walk, have been used to measure physical endurance in patients with various medical conditions but have not been validated on generally healthy subjects. Therefore, the purpose of this study was to evaluate the reliability and validity of the 6-min walk test on a sample of generally healthy, community-residing older adults.

Results of the reliability analysis indicated that 6-min walk scores were highly correlated (had relative consistency) across the three trials administered, with *R* values ranging from .88 to .97. However, the ANOVA analysis indicated that to achieve absolute stability of 6-min walk scores (i. e., where no significant day-to-day changes occur due to testing effects), the first trial needs to be treated as a practice trial. For all comparison groups in this study (i.e., for total subjects and for men and women separately), there were significant increases in number of yards walked from Trial 1 to Trial 2 but not between Trial 2 and Trial 3, thus indicating a plateau in walking performance by Trial 2. Achieving scoring stability, of course, is especially important in any type of intervention study. If stable baseline scores are not obtained, it is difficult to differentiate between treatment effects and testing effects during the course of the study. However, it should be noted that the plateau

in walking performance observed by Trial 2 in this study occurred earlier than in most studies involving patients with medical conditions. Generally, it has been found that at least three trials are required for individuals with special conditions, such as heart and/or respiratory disease, before 6-min walk scores stabilize (Butland et al., 1982; Guyatt et al., 1985a; Knox et al., 1988).

To address the validity of the 6-min walk, scores were compared to performance on a submaximal treadmill test, using a modified Balke protocol. The .78 correlation found between 6-min walk scores and treadmill performance in this study was similar to the .76 correlation found between 6-min walk scores and $\dot{V}O_2$max scores by Stillwell et al. (1996), the only other study we found that tested the 6-min walk in a general (nondiseased) population. In studies involving subjects with cardiac and respiratory disease, the correlations between 6-min walk performance and various measures of aerobic endurance tend to be somewhat lower. Guyatt et al. (1985b) reported a .58 correlation between 6-min walk performance and a maximal cycle ergometer test for 43 patients with either heart or lung disease. More recently, a correlation of .68 was found between 6-min walk scores and scores on a shuttle walking test in a group of patients with chronic obstructive pulmonary disease (Singh, Morgan, Scott, Walters, & Hardman, 1992).

Other findings in the study indicated that there also was a moderate correlation (r = .61–.71) between 6-min walk scores and self-reported functional ability, indicating that walking performance may be a positive indicator of overall physical functioning as well as of cardiovascular functioning. However, the common variance (r^2) associated with an r value as low as .61 is only .37 or 37%, indicating that while the association is positive and highly significant ($p < .001$), it is not a strong relationship.

The expected declines found in walking performance among the various age groups in this study (60-, 70-, and 80-year-olds), as well as the significant differences in the high-active versus the low-active groups, support the construct validity of the test, that is, the responsiveness or sensitivity of the 6-min test relative to its ability to detect differences where differences normally exist. Other evidence of the sensitivity of the 6-min walk with respect to its ability to detect change was noted in a recent 3-month exercise training study. In this study, the 6-min walk test detected physical endurance differences between a group of exercising subjects compared to nonexercising subjects (Fitts & Guthrie, 1995).

Finally, it is interesting to note that although subjects' recorded heart rates were significantly higher during the 6-min walk compared to the treadmill test, RPE scores were very similar during both tests. It is possible that the higher heart rates during the walk are a result of greater overall exertion caused by increased involvement of the upper body (more vigorous swinging of the arms) during the 6-min walk, a phenomenon that was clearly observed by test administrators. However, because of the familiarity of regular walking (compared to treadmill walking), it is also possible that subjects do not perceive that they are doing more strenuous activity. Whether subjects are, in fact, underestimating their exertion rate during regular walking or perhaps overestimating it during treadmill walking is not clear from these data. The predicted relationship between RPE and heart rate scores has not been as well defined for older adults as for younger people. Although no complications were observed during the current study, we recommend that partici-

pants be monitored closely for signs of overexertion during the administration of walking tests.

Conclusion

The results of this study suggest that the 6-min walk test can be used to assess physical endurance in generally healthy older adults with varying levels of physical capacity. The 6-min walk has good test–retest reliability, particularly when a practice trial is given before the test trial, is positively associated with both treadmill and self-report measures of functional performance, and is sensitive to change as indicated by its ability to detect differences across age and activity groups.

References

American College of Sports Medicine (ACSM). (1991). *Guidelines for exercise testing and prescription* (4th ed.). Philadelphia: Lea & Febiger.

American College of Sports Medicine (ACSM). (1995). *ACSM's guidelines for exercise testing and prescription* (5th ed.). Baltimore: Williams & Williams.

American College of Sports Medicine (ACSM). (1997). *Exercise management for persons with chronic diseases and disabilities.* Champaign, IL: Human Kinetics.

Ainsworth, B.E., Montoye, H.J., & Leon, A.S. (1994). Methods of assessing physical activity during leisure and work. In C. Bouchard, R.J. Shephard, & T. Stephens (Eds.), *Physical activity, fitness, and health* (pp. 146-159). Champaign, IL: Human Kinetics.

Ainsworth, B.E., Richardson, M.T., Jacobs, D.R., & Leon, A.S. (1992). Prediction of cardiorespiratory fitness using physical activity questionnaire data. *Research Quarterly for Exercise and Sport*, **63**(Suppl.), A-20.

Baumgartner, T.A., & Jackson, A.S. (1995). *Measurement for evaluation in physical education and exercise science*. Dubuque, IA: Brown & Benchmark.

Bittner, V., Weiner, D.H., Yusuf, S., Rogers, W.J., McIntyre, K.M., Bangdiwala, S.I., Kronenberg, M.W., Kostis, J.B., Kohn, R.M., Guillotte, M., Greenberg, B., Woods, P.A., & Bourassa, M.G. (1993). Prediction of mortality and morbidity with a 6-minute walk test in patients with left ventricular dysfunction. *Journal of the American Medical Association*, **270**, 1702-1707.

Borg, G.A.V. (1970). Perceived exertion as an indicator of somatic stress. *Scandinavian Journal of Rehabilitative Medicine*, **2**, 92-98.

Bravo, G., Gauthier, P., Roy, P., Tessier, D., Gaulin, P., Dubois, M., & Peloquin, L. (1994). The functional fitness assessment battery: Reliability and validity data for elderly women. *Journal of Aging and Physical Activity*, **2**, 67-79.

Butland, R.J.A., Pang, J., Gross, E.R., Woodcock, A.A., & Geddes, D.M. (1982). Two-, 6-, and 12-minute walking tests in respiratory disease patients. *British Medical Journal*, **284**, 1607-1608.

Fenstermaker, K.L., Plowman, S.A., & Looney, M.A. (1992). Validation of the Rockport Fitness Walking Test in females 65 years and older. *Research Quarterly for Exercise and Sport*, **63**, 322-327.

Fitts, S.S., & Guthrie, M.R. (1995). Six-minute walk by people with chronic renal failure. *American Journal of Physical Medicine and Rehabilitation*, **74**, 54-58.

Godin, G., Jobin, J., & Bouillon, J. (1986). Assessment of leisure time exercise behavior by self-report: A concurrent validity study. *Canadian Journal of Public Health*, **77**, 359-361.

Guyatt, G.H., Sullivan, M.J., Thompson, P.J., Fallen, E.I., Pugsley, S.O., Taylor, D.W., & Berman, L.B. (1985a). The 6-minute walk: A new measure of exercise capacity in patients with chronic heart failure. *Canadian Medical Association Journal*, **132**, 919-923.

Guyatt, G.H., Thompson, P.J., Berman, L.B., Sullivan, M.J., Townsend, M., Jones, N.L., & Pugsley, S.O. (1985b). How should we measure function in patients with chronic heart and lung disease. *Journal of Chronic Disabilities*, **38**, 517-524.

Jackson, A.S., Beard, E.F., Wier, L.T., Ross, R.M., Stuteville, J.E., & Blair, S.N. (1995). Changes in aerobic power of men, ages 25-70 years. *Medicine and Science in Sports and Exercise*, **27**, 113-120.

Jackson, A.S., Wier, L.T., Ayers, G.W., Beard, E.F., Stuteville, J.E., & Blair, S.N. (1996). Changes in aerobic power of women, ages 20-64 yr. *Medicine and Science in Sports and Exercise*, **28**, 884-891.

Kline, G.M., Porcari, J.P., Hintermeister, R., Freedson, P.S., Ward, A., McCarron, R.F., Ross, J., & Rippe, J.M. (1987). Estimation of $\dot{V}O_2$max from a one-mile track walk, gender, age, and body weight. *Medicine and Science in Sports and Exercise*, **19**, 253-259.

Knox, A.J., Morrison, J.F., & Muers, M.F. (1988). Reproducibility of walking test results in chronic obstructive airways disease. *Thorax*, **43**, 388-392.

Morrow, J.R., Jackson, A.W., Disch, J.G., & Mood, D.P. (1995). *Measurement and evaluation in human performance*. Champaign, IL: Human Kinetics.

National Center for Health Statistics. (1991). *National Health Interview Survey 1991 (10): Vital Health Statistics*. Washington, DC: National Center for Health Statistics

Osness, W.H., Adrian, M., Clark, B., Hoeger, W., Rabb, D., & Wiswell, R. (1996). *Functional fitness assessment for adults over 60 years*. Dubuque, IA: Kendall/Hunt.

Paffenbarger, R.S.J., Blair, S.N., Lee, I-M., & Hyde, R.T. (1993). Measurement of physical activity to assess health effects in free-living populations. *Medicine and Science in Sports and Exercise*, **25**, 60-70.

Peloquin, L., Gauthier, P., Bravo, G., Lacombe, G., & Billiard, J-S. (1998). Reliability and validity of the 5-minute walking field test for estimating $\dot{V}O_2$peak in elderly subjects with knee osteoarthritis. *Journal of Aging and Physical Activity*, **6**, 36-44.

Reuben, D.D., Siu, A.L., & Kimpau, S. (1992). The predictive validity of self-report and performance-based measures of function and health. *Journal of Gerontology: Medical Sciences*, **47**, 106-110.

Rikli, R.E., & Jones, C.J. (1997). Assessing physical performance in independent older adults: Issues and guidelines. *Journal of Aging and Physical Activity*, **5**, 244-261.

Rosow, I., & Breslau, N. (1966). A Guttman Health Scale for the aged. *Journal of Gerontology*, **21**, 556-559.

Safrit, M.J., & Wood, T.M. (1995). *Introduction to measurement in physical education and exercise science*. St. Louis: Mosby-Year Book.

Schectman, K.B., Barzilai, B., Rost, K., & Fisher, E.B., Jr. (1991). Measuring physical activity with a single question. *American Journal of Public Health*, **81**, 771-773.

Select Committee on Aging. (1992). *Aging research: Benefits outweigh the costs* (Publ. No. 102-871). Washington, DC: U.S. Government Printing Office.

Shephard, R.J. (1997). *Aging, physical activity, and health*. Champaign, IL: Human Kinetics.

Siconolfi, S.F., Lasater, T.M., Snow, R.C., & Carleton, R.A. (1985). Self-reported physical activity compared with maximal oxygen uptake. *American Journal of Epidemiology*, **122**, 101-105.

Singh, S.J., Morgan, M.J., Scott, S., Walters, D., & Hardman, A. (1992). Development of a shuttle-walking test of disability in patients with chronic airways obstruction. *Thorax*, **47**, 1019-1024.

Siu, A.L., Reuben, D.B., & Hays, R.D. (1990). Hierarchical measures of physical function in ambulatory geriatrics. *Journal of the American Geriatrics Society*, **38**, 1113-1119.

Spirduso, W.W. (1995). *Physical dimensions of aging*. Champaign, IL: Human Kinetics.

Steele, B. (1996). Timed walking tests of exercise capacity in chronic cardiopulmonary illness. *Journal of Cardiopulmonary Rehabilitation*, **16**, 25-33.

Stillwell, K.M., Forman, D.E., McElwain, D., Simpson, C., & Garber, C.E. (1996). The 6 minute walk test for evaluation of functional capacity in elderly adults. *Medicine and Science in Sports and Exercise*, **28**, S152.

Acknowledgment

This research was supported by PacifiCare Health Systems.