

Continuous-Scale Physical Functional Performance Test: Validity, Reliability, and Sensitivity of Data for the Short Version

Background and Purpose. The Continuous-Scale Physical Functional Performance Test (CS-PFP) can be used to obtain valid, reliable, and sensitive measurements of physical functional capacity. This test requires a fixed laboratory space and approximately 1 hour to administer. This study was carried out in 4 steps, or substudies, to develop and validate a short, community-based version (PFP-10) that requires less space and equipment than the CS-PFP. **Subjects and Methods.** Retrospective data (n=228) and prospective data (n=91) on men and women performing the CS-PFP or the PFP-10 are reported. A 12-week exercise program was used to examine sensitivity to change. Data analyses were done using paired *t*-test, Pearson correlation, intraclass correlation coefficient (ICC), and delta index (DI) procedures. **Results.** The PFP-10 total score and 4 of the 5 domain scores were statistically similar (within 3%) to those of the CS-PFP. The PFP-10 upper-body strength domain score was 17% lower, but was highly correlated (ICC=.97). Community and established laboratory PFP-10 scores were similar (ICC=.85-.97). The PFP-10 also is sensitive to change (DI=.21-.54). **Discussion and Conclusion.** The PFP-10 yields valid, reliable, and sensitive measurements and can be confidently substituted for the CS-PFP. [Cress ME, Petrella JK, Moore TL, Schenkman ML. Continuous-Scale Physical Functional Performance Test: validity, reliability, and sensitivity of data for the short version. *Phys Ther.* 2005;85:323-335.]

Key Words: *Older adults, Physical capacity, Physical function.*

M Elaine Cress, John K Petrella, Trudy L Moore, Margaret L Schenkman

Use of physical performance measures provides valuable insight into the ability of older people to perform specific tasks that are important for daily living. Measures of task performance, such as the Short Physical Performance Battery¹ and the Physical Performance Test,² have been used in several epidemiological studies (eg, Epidemiologic Studies of Established Populations in the Elderly,¹ Women's Health and Aging Study,³ Health ABC Study⁴), establishing evidence for mobility and balance performance as risk factors for institutionalization and mortality. A measure, such as the Physical Performance Test, typically requires very little equipment, and tasks can be administered in most locations.² However, these measures may not adequately challenge individuals with higher-level functional ability and, therefore, often are unable to quantify function in people with higher physical ability.⁴ For example, the ability to pick up an item from the floor is not difficult for many older individuals, whereas carrying a heavy bag of groceries as fast as possible over a considerable distance is demanding. Furthermore, many performance-based measures (eg, Functional Reach Test,⁵ 6-minute walk test⁶) assess only a single task. These single-item tests have the advantage of a low burden of testing and administration. These tests, however, do not measure the physical endurance that is required when tasks are performed serially, which more closely mimics how function normally takes place within the home. Tests are needed that involve serial performance of tasks that quantify performance of individuals in an ecological context, that are applicable to a wide range of environments, and that are adequately demand-

ing for individuals with a broad range of functional abilities.

One such test, the Continuous-Scale Physical Functional Performance Test (CS-PFP), offers many attractive features for measuring physical function across a wide range of functional abilities.⁷ The CS-PFP consists of 16 household tasks that are performed serially, in a manner of usual function (eg, in a person's preferred manner rather than in a constrained fashion). Task performance reflects the person's ability because each task is performed at maximal effort within the person's judgment of comfort and safety. The CS-PFP yields measurements that are reliable,⁷ valid,⁷ and sensitive to change.⁸ The CS-PFP uses time, distance, and weight to evaluate functional ability, based on the performance of 16 sequentially performed common household tasks.

The CS-PFP was validated on ambulatory older adults with a broad range of physical abilities but without a focus on specific underlying disorders. The CS-PFP also has been used to assess function in several groups of people with specific disorders, including cardiovascular disease,⁹ post-burn injury,¹⁰ Parkinson disease,¹¹ and low back pain.^{12,13} This measure has good sensitivity to detect effect sizes in the range of 0.5 to 0.7^{8,9,14} with sample sizes of less than 15 people per group.¹⁴ The 16 tasks of the CS-PFP were selected, in part, to include a spectrum of physiological demands. The 16 tasks of the CS-PFP are reported as a total score and 5 domain scores that uniquely reflect the physical demands of the separate domains. The domains are upper-body strength,

ME Cress, PhD, FACSM, is Associate Professor, Department of Exercise Science, and The Gerontology Center, University of Georgia, 300 River Rd, Athens, GA 30602-6554 (USA) (mccress@coe.uga.edu). Address all correspondence to Dr Cress.

JK Petrella, PhD, is Postdoctoral Research Fellow, Geriatric Research, Education Clinical Center, Department of Physiology and Biophysics, University of Alabama-Birmingham, Birmingham, Ala.

TL Moore, MS, is a graduate assistant and doctoral student in the Department of Exercise Science, University of Georgia.

ML Schenkman, PT, PhD, is Professor of Physical Therapy, Department of Rehabilitation Medicine, University of Colorado Health Sciences Center, Denver, Colo.

Dr Cress provided concept/idea/research design, writing, data collection, project management, fund procurement, subjects, facilities/equipment, institutional liaisons, and consultation (including review of manuscript before submission). Dr Petrella and Ms Moore provided data collection and analysis and writing. Dr Schenkman provided writing and conceptual assistance.

This study was approved by the human subjects institutional review committees of the University of Washington and Northwest Hospital, University of Georgia, Medical College of Georgia, Georgia State Department of Human Resources, and Athens Community Council on Aging.

The results of this study were presented, in part, at the Annual Conference of the American College of Sports Medicine; May 28-31, 2003; San Francisco, Calif.

Funding for this research was provided by grants from National Institutes of Health grants 1 R29 AG10267, 1 R01 AG10853, and R010943; a Medical College of Georgia-University of Georgia Biomedical Research Institution Grant; and a University Georgia Gerontology Consortium Student Seed Grant.

This article was received May 31, 2004, and was accepted September 21, 2004.

lower-body strength, upper-body flexibility, balance and coordination, and endurance. The “strength” domains are reflective of force production but will be called “strength” to remain consistent with the domain names used in previous publications of the test.^{7-10,14-16} The scoring was developed through measurement of older adults with a wide range of physical abilities (from independent community dwellers to people in assisted-living facilities). From these data, ranges were established on the 16 unique tasks, which were then used in the development of the 1-to-100 scale.^{8,15} As a consequence of the sample tested and the scoring, the total score and domain scores are without known ceiling or floor effects in both physically robust and frail adults.

The CS-PFP is one of the few measures that has been validated against physiological capacity measures (maximal oxygen consumption, peak torque, reaction time, and range of motion).⁷ Therefore, the domain scores (upper-body strength, lower-body strength, flexibility, balance and coordination, and endurance) can be used to gain insight into the impairments underlying those aspects of function.⁷ In addition, thresholds of function have been identified that define a physical reserve required to live independently.¹⁷ People with a higher physical reserve complete the test more quickly and carry more weight.⁷ The ability to quantify physical function in adults with higher function is important because this is a rapidly growing segment of the population.¹⁸

Applicability of the CS-PFP is constrained for both research and clinical purposes because the CS-PFP uses a fixed laboratory space and takes approximately 1 hour to administer. To enhance its applicability, a shorter version of the CS-PFP was developed that requires less equipment and space and can be used in the community setting. This shorter version consists of 10 of the original 16 tasks and is called the Physical Functional Performance 10 Test (PFP-10). The PFP-10 was designed to retain the ability to evaluate interventions using small sample sizes while reducing administration time from 60 to 30 minutes. As a result of these changes, the PFP-10 test is more feasible for use in clinical practice settings and in clinical trials. The purposes of this article are: (1) to describe the strategy used to identify tasks for inclusion in a short form of the CS-PFP, (2) to establish convergent validity of the PFP-10 scores by comparing the performance on the CS-PFP and PFP-10 for people with lower and higher levels of function, (3) to establish parallel-forms reliability of data obtained in the community setting with short form of the CS-PFP (PFP-10) through comparison with a laboratory-based version of the test, and (4) to establish the ability of the PFP-10 test to detect change in function with adequate sensitivity.

Methods

Validation of the PFP-10 took place in several steps (or substudies), with each substudy requiring a different sample and a different method. The overall methods are described, followed by the purpose, specific procedures, and results of each substudy.

Subjects

In substudy 1, a pre-existing data set (n=228) was used consisting of data from previous studies.^{7,8} In substudies 2, 3, and 4, new participants (n=91) were recruited separately for each study. Participants of all substudies met the following criteria. Male and female participants who were over the age of 60 years were included. To ensure that participants could safely participate in aerobic conditioning programs in the dataset of 228 participants, each participant’s personal physician provided a medical clearance using the following exclusion criteria: poorly controlled or unstable cardiovascular disease, diabetes, thyroid disease, obesity (body mass index ≥ 30), recent unhealed bone fracture, severe hypertension ($>160/90$ mm Hg), anatomic deformity (amputation), excessive alcohol intake (>3 drinks per day), severe spinal osteopenia accompanied by severe kyphosis or classic anterior compression fractures, a life expectancy of <1 year, chronic vestibular disease, disorders with a highly variable course (eg, multiple sclerosis), malnutrition, the inability to speak English, or the inability to follow directions.

Participants were differentiated into categories using previously described criteria.¹⁷ Specifically, those who lived independently in the community, referred to as the “community dweller” (CD) group, were differentiated from those who lived in congregate housing. *Congregate housing* (CH) is defined as a multilevel retirement community with light housekeeping, at least one meal per day, and the option for more support services. In addition, participants living in congregate housing were further differentiated according to their self-reported difficulty with functional tasks. Those who scored less than 65 on the Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36) physical function scale (SF-36PF) were classified as dependent (CH-D group), whereas those who scored 65 or higher were classified as independent (CH-I group).¹⁷ Prior to entering the study, participants signed an informed consent form approved by the appropriate institutional review boards.

Measures

CS-PFP. The CS-PFP is reported to yield valid and reliable measurements⁷ and is used to quantify physical function based on performance of 16 different tasks that reflect ability in 5 physical domains (upper-body strength, lower-body strength, upper-body flexibility, bal-

Table 1.

Tasks of the Continuous-Scale Physical Functional Performance Test (CS-PFP) and Those Included in the Physical Functional Performance 10 Test (PFP-10)

CS-PFP Performance Tasks (Short Name)	Tasks Included in PFP-10
Low difficulty	
Carry a weighted pan a distance of 1 m (pan carry)	Yes
Pouring water from a jug into a cup (pouring)	No
Donning and removing a jacket (jacket)	Yes
Place a sponge on and remove it from an adjustable shelf (shelf reach)	Yes
Moderate difficulty	
Floor sweeping with broom and dustpan (floor sweep)	Yes
Transfer clothes from washer to dryer (laundry 1); transfer clothes from dryer to basket (laundry 2)	Yes
Open and pass through a fire door (door pull)	No
Making a bed (bed)	No
Vacuuming (vacuum)	No
Place a strap over a shoe (shoe strap)	No
Pick up 4 scarves from the floor (scarves)	Yes
High difficulty	
Carry weighted bag up and down simulated bus stop (bus)	No
Carry groceries 70 m (grocery)	Yes
Sit and stand up from the floor (floor sit/rise)	Yes
Climb stairs (stairs)	Yes
6-min walk (walk)	Yes
Total of timed tasks (total timed tasks)	Yes

ance and coordination, and endurance).⁷ Physical functional performance was measured as weight, time, or distance. Tasks from the CS-PFP and those selected for the PFP-10 are shown in Table 1. Scores are scaled from 0 to 100 utilizing the following formula based on lower and upper extremes of performance from previously tested older adults^{7,8}:

$$\text{CS-PFP score} = \frac{(\text{observed score} - \text{lower limit})}{(\text{upper limit} - \text{lower limit})} \times 100$$

The CS-PFP total score is the average corrected score of all tasks, and the total score for each domain is the average of the scores for the tasks in that domain. The same ranges and procedure were used to calculate the CS-PFP and the PFP-10 scores. Exertion during the test was assessed using the Borg 6–20 Rate of Perceived Exertion Scale (RPE).¹⁹ If an individual did not attempt a CS-PFP or PFP-10 task, it was scored as 0. Scores range from 0 to 100, with higher scores reflecting higher fitness and higher function.⁷ Depending on the ability of

the older adults being tested, the CS-PFP requires approximately 1 hour to complete. Details of the equipment, dimensions for setting up the established laboratory, and data reduction are available at the CS-PFP Web site.²⁰ The CS-PFP test was performed in a laboratory registered at the Aging and Physical Performance Laboratory at the University of Georgia. All testing followed the published protocol for the CS-PFP.^{7,8} Summary scores for the CS-PFP were generated from the CS-PFP Web-based program.²⁰ Summary scores for the PFP-10 were generated using a syntax algorithm for the SPSS version 11.1 statistical package.* This program is available upon request from the first author.

SF-36PF. The SF-36PF was used to assess the participants' self-rated function. The SF-36 is reported to yield valid and reliable measurements of health status.²¹ The SF-36 consists of 8 subscales that measure different health concepts or domains: vitality, role–physical, role–emotional, physical function, mental health, general health, bodily pain, and social functioning.²¹ Only one SF-36 domain (physical function) is reported in this study. The SF-36PF is scored from 0 to 100, with higher scores indicating better function. A score of less than 65 is associated with a transition to disability.^{17,22} Individuals scoring ≥ 65 on the SF-36PF were categorized as independent (CH-I group). Individuals scoring less than 65 on the SF-36PF were classified as dependent (CH-D group).

Data Analysis

All analyses were conducted using the SPSS version 10.0 statistical package.* The following analytic techniques were applied for the different substudies. In substudy 1, to establish the ability of the tasks to discriminate among groups, a one-way analysis of variance (ANOVA) was used, followed by a Bonferroni or Games-Howell adjusted *post hoc* analysis for contrasts.²³ The Levene test of equality of error variances²³ was used to determine the appropriate adjustment for *post hoc* testing (ie, a Levene test result with a value of $P < .1$ indicated the need for a Games-Howell adjustment). Tasks of the domains of the CS-PFP have internal consistency of .74 to .97. The Cronbach alpha (α) statistic was used to determine that internal consistency of the domains was retained in the PFP-10.²⁴ To determine whether comparable information was provided between the CS-PFP and PFP-10 or the community and the established laboratory site, the results of the total scores and domain scores were compared using paired *t* tests, Pearson correlations, and intraclass correlations (goodness-of-fit model) in substudies 1, 2, and 3. Finally, in substudy 4, to establish sensitivity of the PFP-10, paired *t* tests were used to detect change after the 12-week exercise program with a Bon-

* SPSS Inc, 233 S Wacker Dr, Chicago, IL 60606.

ferroni correction (probability if .05/5 planned comparisons) to establish the *P* value of <.01 for significance. The delta index (DI) was calculated for the pretest and posttest values ($[\text{posttest-pretest}]/\text{baseline standard deviation}$) and assessed using values of <.2 as indicating a small change, values of .3 to .7 as indicating a moderate change, and values of .8 or higher as indicating a large change.²³ Results are reported as means and standard deviations.

Substudy 1: Identification of Tasks for Inclusion in the PFP-10

Purpose

The first substudy identified tasks for inclusion in a short form of the CS-PFP. Using a pre-existing data set of the CS-PFP (16 tasks), a subset of 10 tasks was identified to comprise a short form (PFP-10). Using the same data set, scores were then calculated for the PFP-10 and compared with the CS-PFP scores.

Method and Subjects

A pre-existing data set was used¹⁷ that consisted of data from 228 individuals, including 134 participants (58.8%) who were community dwelling and independent (CD group), 49 participants (21.5%) in a congregate care facility who were independent (CH-I group), and 45 participants (19.7%) in a congregate care facility who were dependent (CH-D group). Mean perceived function as measured with the SF-36PF was lower for the CH-D group ($\bar{X}=47.3$, $SD=13$) than for the CD group ($\bar{X}=90.3$, $SD=8$) or the CH-I group ($\bar{X}=86.8$, $SD=9$).

Criteria used for constructing the short form of the CS-PFP. The following criteria were used to determine which items from the original CS-PFP to include in the shorter version of the test:

1. For domains with fewer than 3 items (eg, flexibility), all items were retained to preserve the integrity of that domain, even if they did not discriminate across the 3 groups.
2. Tasks that did not discriminate among the 3 groups were removed in order to eliminate ceiling or floor effects on individual tasks (eg, pouring water from a jug into a cup).
3. When 2 items contained similar components, the item that provided the most information was retained. For example, carrying groceries and carrying luggage both require carrying weight up a platform. Carrying luggage was deleted, but carrying groceries was kept because the task of carrying groceries involves walking a greater distance and balancing the groceries while opening a door. For other items, the item that had the

least burden of setup and administration was retained. For example, floor sweeping and vacuuming are similar, but vacuuming requires a vacuum cleaner with specific features. Therefore, the vacuuming task was deleted, and the floor-sweeping task was retained.

Results and Discussion

The average age of the sample of 228 older men and women was 76.0 years ($SD=7.1$, range=60–97). The sample was 78.5% female and 59% community dwellers, and the average SF-36PF score was 78.9 ($SD=19.3$). Tasks that comprise the CS-PFP and those tasks retained in the PFP-10 are identified in Table 1. Results of the ANOVA and *post hoc* analyses are shown in Table 2. Eleven variables discriminated among all 3 groups. The variables are reported in raw scores for each task. Some tasks (eg, pouring water) include both a time variable and a weight variable. Using this information as well as the previously described criteria, the following tasks were eliminated: pouring water (did not discriminate), placing a strap over a shoe (did not discriminate), making a bed (burden of setup and administration), vacuuming (components similar to those of floor sweeping, greater burden of setup), opening and passing through a fire door (burden of setup), and carrying a weighted suitcase up a 3-step bus platform (components similar to carrying groceries) (Tab. 1). The resulting shorter version of the test comprised 10 tasks and 13 unique variables.

Domain scores and total scores as calculated for the PFP-10 ($n=228$) are reported in Table 3. With the exception of the upper-body strength domain, the mean difference between the domain scores of the CS-PFP and the PFP-10 was less than 1, and the difference in the total score was 2.1 units. Pearson correlation coefficients between the domain scores of the PFP-10 and the CS-PFP were high ($r=.97-1.00$, Tab. 3).

Before proceeding to substudy 2, it was necessary to determine which items of the CS-PFP would provide the greatest information and be retained for the PFP-10. In this sample of 228 subjects, the paired *t* test revealed differences between the CS-PFP and the PFP-10 on many of the domain scores (Tab. 3). These differences were significant because of the large sample size ($n=228$), increasing the likelihood of statistical significance even though the mean differences were small. These small differences between total scores and lower-body strength, endurance, and balance and coordination domain scores of the CS-PFP versus the PFP-10 were not considered to be clinically meaningful. The difference between the 2 tests for the upper-body strength domain was 9.08 units.

Several factors may have contributed to the upper-body strength domain score being lower for the PFP-10 than

Table 2.Comparison of Task Performance for 3 Groups of People With Different Living Environments and Perceived Function (Analysis of Variance)^a

Task	CD	CH-I	CH-D	F	CD vs CH-I	CD vs CH-D	CH-I vs CH-D
	$\bar{X} \pm SD$	$\bar{X} \pm SD$	$\bar{X} \pm SD$		P	P	P
N	134	49	45				
Pan carry (s)	4.64±1.2	5.42±1.8	6.12±1.8	18.74	.016	.000	.140
Pan carry (kg)	11.27±6.7	10.27±6.5	6.96±3.2	8.41	.627	.000	.006
Pouring (s)	8.98±2.0	10.59±2.6	12.02±3.4	28.21	.001	.000	.065
Pouring (kg)	4.31±0.05	4.31±0.1	4.23±0.3	4.83	.638	.219	.369
Jacket (s)	17.37±6.5	18.86±6.99	32.7±17.9	41.30	.406	.000	.000
Shelf-reach (cm)	217.77±10.4	213.82±12.6	202.74±20.6	21.08	.238	.000	.000
Shelf-reach height/standing height	1.32±0.03	1.30±0.04	1.28±0.11	8.20	.016	.057	.493
Floor sweep (s)	34.67±11.3	40.74±14.6	50.85±14.3	27.79	.028	.000	.003
Laundry 1 (s)	26.34±6.3	33.00±10.0	45.70±10.9	93.44	.000	.000	.000
Laundry 2 (s)	20.77±4.6	25.12±7.8	32.77±8.9	59.59	.002	.000	.000
Opening a door							
Door pull time (s)	3.59±0.7	4.14±0.7	4.45±0.7	6.81	.022	.017	1.000
Fire door weight (kg)	15.52±14.0	13.83±11.7	15.10±7.2	19.91	.022	.000	.003
Bed (s)	83.02±24.0	116.20±54.9	153.98±66.3	48.48	.000	.000	.100
Vacuum (s)	42.13±14.0	48.29±19.0	61.74±19.6	24.29	.104	.000	.003
Shoe strap ^b	7.01±1.8	8.34±3.0	8.68±2.7	3.56	.200	.363	.966
Scarves ^b	4.79±1.4	5.64±2.3	6.61±1.5	4.21	.206	.043	.676
Bus (s)	19.30±7.6	22.3±6.7	39.6±19.3	60.57	.028	.000	.000
Bus (kg)	12.42±5.4	10.3±5.5	5.18±1.9	35.78	.066	.000	.000
Grocery (s)	54.29±14.2	63.06±18.8	89.3±29.8	56.22	.011	.000	.000
Grocery (kg)	14.79±5.1	11.77±4.9	6.65±2.8	50.49	.001	.000	.000
Getting down and up							
Bath time	7.33±7.0	10.28±9.8	36.66±38.2	36.18	.01 (n=81)	.000 (n=30)	.002 (n=28)
Floor sit/rise	9.23±3.7	9.61±3.2	23.62±13.6	22.94	.912 (n=43)	.106 (n=18)	.114 (n=6)
Stairs (s)	5.82±0.8	6.20±0.9	8.72±3.7	43.63	.023	.000	.000
Walk (m)	560.89±75.6	506.21±79.5	351.80±107.7	104.16	.000	.000	.000
Total of timed tasks	374.65±35.3	202.75±51.1	299.02±89.8	75.09	.000	.000	.000

^a Refer to Table 1 for full names of tasks. CD=community dweller, CH-I=congregate housing-independent, CH-D=congregate housing-dependent. Italics indicate tasks that were used in the earliest version of the CS-PFP. These tasks were modified for the current version. Specifically, the bath time⁷ initially was obtained by using an actual bathtub (CD, n=216; CH-I, n=48; CH-D, n=34), but was changed so the participant sits down and then stands up from the floor between 2 chairs⁸ (CD, n=53; CH-I, n=16; CH-D, n=6). The fire door was assessed using a spring scale attached to the door handle, and the person pulled as if opening a heavy door⁷ (CD, n=216; CH-I, n=48; CH-D, n=34), and this task was replaced with timing how long it took for a person to open and pass through a fire door⁸ (CD, n=53; CH-I, n=16; CH-D, n=6). Italicized tasks were discontinued and another task was substituted ("fire door weight" was changed to "door pull time," and "bath time" [sitting down and rising from a bathtub] was changed to "floor sit/rise").

^b Published in Cress et al.⁸

for the CS-PFP. In the CS-PFP, 6 variables (weight carried in a pan, weight of jug from which water is poured, opening a fire door, transferring laundry, weight carried in luggage, and weight of groceries carried), contribute to the domain, whereas in the PFP-10, only 3 of the variables (weight carried in a pan, weight of groceries carried, transferring laundry) are used to determine the score. The most obvious difference comes from the elimination of the task where individuals carry a jug of water weighing 4.5 kg and pour a cup of water. Ninety-six percent of the subjects (n=220) carried the

maximum weight the jug could hold. Elimination of this task contributed to lowering the overall upper-body strength domain score. This may have accounted, in part, for the statistically significant differences. Weight carried on to and off of the bus was eliminated from the upper-body strength domain. That weight carried may be more reflective of the confidence and ability in lower-body force production rather than upper-body force production alone. The ability to open a heavy fire door also was eliminated from the upper-body strength domain. The elimination of these tasks from the upper-

Table 3.

Comparison of the Continuous-Scale Physical Functional Performance Test (CS-PFP) and the Physical Functional Performance 10 Test (PFP-10) Using Retrospective Data (n=228)

Variable	CS-PFP	PFP-10	t	P	Pearson r
	$\bar{X} \pm SD$	$\bar{X} \pm SD$			
Total score	47.96±18.5	45.86±19.0	12.28	.00	.99
Upper-body strength	48.42±16.5	39.34±19.3	25.02	.00	.967
Upper-body flexibility ^a	45.05±23.8	45.05±23.8	0.00	N/A ^b	1.00
Lower-body strength	46.50±19.9	46.80±20.1	-1.13	.26	.98
Balance and coordination	51.09±22.0	50.13±21.6	2.60	.01	.97
Endurance	48.72±19.9	47.98±19.6	3.25	.00	.98

^aThis domain is the same for the CS-PFP and the PFP-10 because it has only 2 tasks.

^bN/A=not applicable.

body strength domain did not affect the total PFP-10 score or other domain scores.

In summary, the PFP-10 was largely comparable to the CS-PFP, with only the upper-body strength domain scores differing substantially. Based on these results, the PFP-10 was subjected to further investigation.

Substudy 2: Comparison of Performance on the CS-PFP With Performance on the PFP-10

Purpose

The second substudy was conducted to determine convergent validity of the PFP-10 scores by comparing the scores for the PFP-10 and the CS-PFP administered on different occasions in a new sample of participants.

Method and Subjects

Flyers, recruitment presentations, and referrals were used to recruit men and women aged 60 years and older from public housing and the Athens, Ga, community at large. All participants received a medical clearance to establish inclusion and exclusion criteria. Potential participants were excluded if they had poorly controlled or unstable cardiovascular disease, heart failure, uncontrolled arrhythmias, severe and symptomatic aortic stenosis, or uncontrolled casual blood glucose >200 mg/dL. *Casual blood glucose* is defined as blood glucose taken at any time of day without regard to the time of the last meal.²⁵ They also were excluded due to inability to follow directions or keep appointments, uncontrolled hypertension, leg or arm amputation, excessive alcohol intake (>3 drinks per day), or life expectancy of less than 1 year.

All testing was performed at the Aging and Physical Performance Laboratory, University of Georgia. Participants performed the 16-task CS-PFP and the 10-task PFP-10 on separate days approximately 1 week apart. The order of testing was randomized.

Results and Discussion

Sixty people expressed interest in participation. Eight people were not cleared by their physician. Twelve people, after receiving physician clearance, decided not to participate due to scheduling problems, personal reasons, or lack of transportation. The 40 participants had a mean age of 77.2 years (SD=6.4, range=65–94), were 65% female and 50% community dwellers, and had an average SF-36PF score of 83.38 (SD=11.9). Correlations between the PFP-10 scores and the CS-PFP scores ranged from $r = .86-.95$, with upper-body flexibility having the lowest domain correlation (Tab. 4). Based on RPE, the participants gave a similar effort on both the CS-PFP and the PFP-10, with mean RPE scores of 12.0 (SD=1.4) for the CS-PFP and 11.9 (SD=1.8) for the PFP-10.

The mean upper-body strength domain score was lower for the PFP-10 than for the CS-PFP. All other domain scores were within 2 units of each other. The Cronbach α was used to establish internal consistency among the variables of each domain. The Cronbach α was high for all domains (upper-body strength, $\alpha = .83$; lower-body strength, $\alpha = .87$; balance and coordination, $\alpha = .90$; and endurance, $\alpha = .91$), except for the upper-body flexibility domain, in which the Cronbach α was moderate ($\alpha = .69$). Based on these results, the PFP-10 was subjected to further investigation.

Substudy 3: Community-Based Testing

Purpose

The next step to establishing the broader utility of the PFP-10 was to determine whether equipment set up at a community site would yield results that were comparable to those in the established laboratory. We hypothesized that older adults with a broad range of physical abilities would have similar physical function scores in the community setting using the portable CS-PFP compared with the established laboratory.

Table 4.

Comparison of the Continuous-Scale Physical Functional Performance Test (CS-PFP) and the Physical Functional Performance 10 Test (PFP-10) Using Prospective Data (n=40)

Variable	CS-PFP	PFP-10	t	P	Pearson r^a (Intraclass Correlation Coefficient)
	$\bar{X}\pm SD$	$\bar{X}\pm SD$			
Total score	48.4±15.6	49.0±16.4	0.46	.646	.95* (.91)
Upper-body strength	49.2±15.2	60.3±18.2	5.80	.000	.95* (.90)
Upper-body flexibility ^b	70.3±13.3	68.6±15.1	0.91	.369	.86* (.74)
Lower-body strength	48.9±16.7	45.9±19.4	2.24	.570	.94* (.88)
Balance and coordination	46.9±16.0	43.9±16.1	1.06	.321	.95* (.90)
Endurance	46.5±15.9	45.8±16.4	0.78	.440	.93* (.91)

^a Asterisks after Pearson r correlation coefficients indicate statistical significance at $P<.003$.

^b This domain is the same for the CS-PFP and the PFP-10 because it has only 2 tasks.

Method and Subjects

The strategies for recruitment and inclusion and exclusion criteria were the same as described for substudy 2. All equipment and props for testing were taken to the site to replicate the procedure in the established laboratory. In order to make the PFP-10 feasible for use in the community, some modifications to the equipment and the test procedures were necessary. The equipment was transported in a van to the public housing site, Denney Tower, where it was set up in a room (approximately 6 × 7.3 m [20 × 24 ft]) to the exact dimensions for each task of the PFP-10 as described in the published CS-PFP protocol^{7,8} and detailed at the University of Georgia College of Education Web site.²⁶ A long hallway (approximately 105 m) with minimal traffic was used for the 6-minute walk test.

For the floor sit/rise task, a 5-cm-thick mat was used when the floor did not have a carpet. Rubber stoppers were put on the chair legs used for support in standing to prevent them from sliding on a tile floor. Tile (121.9 × 91.4 cm) was mounted on a wooden back for use for test purposes at sites that did not have a tile floor. If a smooth-surface floor is available, the dimensions can be marked using tape applied directly on the floor. The color of the tile was chosen to contrast with the kitty litter used during the floor sweep task.

Several items were constructed of lightweight materials to allow for easy use at a variety of community sites. Included were the washer and dryer and the platform. These lightweight items could be transported by one person and set up in the remote sites in 1½ to 2 hours.

The washer with a door on top and a dryer with a door on the front were constructed out of plastic tubs (depth=54.6 cm, diameter=49.5 cm). The tub was mounted on a stand with a final height of 88.9 cm for the washer and 83.8 cm for the dryer. These lightweight items simulated the dimensions of a commercial washer and dryer.

The adjustable shelf for the reaching task was mounted on a stand and was free-standing (maximum height=243.8 cm). Kitchen counters were constructed to the final dimensions (91.4 cm high) and mounted with wheels to enhance portability. The distance between counters could be adjusted to meet the final required dimension of 59 cm.

Because many senior centers and facilities built to accommodate people with disabilities do not have a staircase, the bus platform was modified to have 4 stairs instead of 3 stairs as in the CS-PFP protocol. The stair platform was made in 2 pieces to enhance portability. Separate handrails were made from polyvinyl chloride (PVC) pipe, which is lightweight and sturdy. For the stair-climbing task, the participant was asked to ascend and descend the 4 stairs. The calculation of the score is based on time per stair. Therefore, for the 8 stairs of the community-based laboratory, the calculated score was time to complete the task divided by 8. This procedural modification was used as a substitution for the original CS-PFP task of climbing 9 to 11 stairs where time per stair was used in the calculations. For the grocery task, which also requires the bus platform, the procedure was the same as for the CS-PFP except that the person climbed 4 stairs rather than 3 as in the established laboratory. The community-based laboratory equipment (mock washer and dryer, 2 counters, shelf, tile flooring, 4-step platform, measuring wheel, 29.5-kg weight, groceries, broom and dustpan) can be transported in a van in one trip. The equipment was designed for carrying by one person. Initial setup requires approximately 1½ hours due to the measurements required; however, subsequent testing periods entail less time. Participants were tested at the established laboratory and at a public housing site (community-based laboratory) within 2 weeks. The order was randomized to control for an order effect.

Results and Discussion

The 20 participants had a mean age of 69.3 years (SD=7.1, range=60–86); were 75% female, 65% Caucasian, 35% African American, and 100% community

Table 5.

Comparison of the Physical Functional Performance 10 Test (PFP-10) Scores for the Established Laboratory and Community-Based Setting (n=20)

Summary and Domain Scores	Established Laboratory	Community Setting	P	Pearson <i>r</i> ^a (Intraclass Correlation Coefficient)
	$\bar{X}\pm SD$	$\bar{X}\pm SD$		
Total score	44.9±17.5	42.7±19.0	.115	.95 (.973)
Domain scores				
Upper-body strength	45.4±20.5	46.9±22.4	.331	.95* (.97)
Upper-body flexibility	55.3±14.1	54.2±14.7	.646	.74* (.85)
Lower-body strength	39.0±18.4	35.6±19.7	.066	.92* (.96)
Balance and coordination	45.8±18.4	43.2±20.1	.102	.94* (.97)
Endurance	46.1±18.3	43.3±20.2	.073	.95* (.97)
Task ^b scores				
Pan carry (s)	4.2±1.1	4.3±1.4	.557	.99*
Pan carry (kg)	14.4±7.5	14.5±7.4	.683	.97*
Jacket (s)	15.1±7.0	16.0±6.6	.461	.92*
Shelf reach (cm)	211.3±16.0	212.8±16.5	.338	.89*
Floor sweep (s)	33.2±10.2	28.6±11.0	.007	.71*
Laundry 1 (s)	34.8±13.8	33.3±12.8	.303	.91*
Laundry 2 (s)	20.8±9.9	19.7±6.7	.474	.82*
Scarves (s)	8.1±4.7	7.2±3.2	.161	.82*
Grocery (s)	71.2±26.7	68.1±24.4	.080	1.00*
Grocery (kg)	13.0±6.8	11.7±6.3	.086	.88*
Floor sit/rise (s)	12.9±10.5	15.6±11.6	.003	.76*
Stair climb (s)	9.4±7.8	10.4±5.8	.253	.90*
Walk (m)	416.8±159.0	389.8±163.1	.113	.92*
Total of time tasks (s)	199.0±60.2	192.0±60.8	.861	.96*

^a Asterisks after Pearson *r* correlation coefficients indicate statistical significance at $P<.003$ for the difference between the mean scores for the established laboratory and community-based setting.

^b Refer to Table 1 for full names of tasks.

dwelling (57% living in a house and 43% living in high-rise, low-income housing); and their average SF-36PF score was 76.3 (SD=24.9).

Whether administered in the established laboratory or the community setting, the PFP-10 total scores were similar (established laboratory: $\bar{X}=44.9$, $SD=17.5$; community-based: $\bar{X}=42.7$, $SD=19.0$), as were the 5 domain scores (Tab. 5). Correlations were $r>.9$ for all domains, except for upper-body flexibility ($r=.74$). Scores for the individual tasks also were moderately to highly correlated (Tab. 5).

Substudy 4: Sensitivity to Change

Purpose

The final substudy was designed to examine whether the community-based version of the PFP-10 was sensitive to change. We hypothesized that a 12-week evidence-based aerobic and muscle force production training program would improve physical function.

Methods and Subjects

Inclusion and exclusion criteria were the same as those described in substudy 2. In addition to recruiting subjects from the public housing facility, participants were recruited from 2 senior centers, located in different

counties, that offered congregate meal programs. The PFP-10 procedure as described in substudy 3 was used as an outcome measure. The PFP-10 testing was conducted at a county senior center (Newton County Senior Center crafts room) and a public housing facility (Denney Tower, described in substudy 3). All participants were tested at the same site before and after the intervention.

Prior to the administration of the CS-PFP or the PFP-10, testers were required to establish competency in administration by comparison with a trained administrator. Competency on test administration was defined as an interrater reliability value of .95 or better on timing of tasks as established using double timing with a proficient administrator of the test. Test-retest reliability of the community-based version of the PFP-10 was established at baseline for substudy 4 with data from 3 participants who also participated in substudy 3. The reliability criterion was that all domains be greater than $r=.75$.

The exercise intervention was based on the Lifetime Fitness Program, an evidence-based exercise program designed by researchers and specialists at the University of Washington and Group Health Cooperative.²⁷ The exercise program focused on endurance, muscle force production, posture, balance, and flexibility. Classes were led by a certified exercise leader 3 times a week for

Table 6.

Sensitivity to Change: Physical Functional Performance 10 Test (PFP-10) Scores at Baseline and 12 Weeks (n=31)

Summary and Domains	Baseline	12 Weeks	Delta Index	P
	$\bar{X} \pm SD$	$\bar{X} \pm SD$		
Total score	39.2 ± 13.5	45.7 ± 15.1	0.48	.000
Domain scores ^a				
Upper-body strength	39.3 ± 15.0	42.5 ± 16.8	0.21	.042
Upper-body flexibility	56.9 ± 14.9	61.5 ± 14.4	0.30	.003
Lower-body strength	31.5 ± 11.6	37.1 ± 14.1	0.48	.000
Balance and coordination	39.7 ± 14.9	47.8 ± 16.3	0.54	.000
Endurance	40.5 ± 15.2	47.9 ± 16.5	0.49	.000

^aDomain scores are the average of scores for tasks within a domain (range=0–100).

12 weeks. Prior to each session, blood pressure was measured and attendance was recorded. Each class consisted of 7 to 10 minutes of warm-up exercises and 20 minutes of aerobic conditioning, followed by 3 to 5 minutes of cool-down and then 20 minutes of progressive resistance training and ending with 8 to 10 minutes of stretching. Exercise intensity was monitored using RPE values.¹⁹ Participants were instructed to work at an aerobic exercise intensity that elicited an RPE of 12 to 13, reflecting a physical exertion of “somewhat hard” on a self-rated perceived exertion scale.¹⁹ The muscle force production training was set at an RPE of 15 to 17 (“hard”) using 8 muscle groups. Balance and flexibility exercises were included in the warm-up and cool-down segments. Participants were encouraged to attend all exercise classes (36 sessions); participants who missed any sessions were required to attend classes for 2 consecutive weeks before exit testing. The exercise classes at one senior center were held in a large banquet room or gymnasium. The exercise classes at the other senior center were held in a small meeting room, limiting the number of participants who were able to join the program.

Results and Discussion

Seventy individuals volunteered for substudy 4. Twenty-nine volunteers were either ineligible by the health screening or unable to obtain physician clearance to participate. Ten individuals (9 women and 1 man [24.4%]) dropped out of the study. Reasons for dropping out were transportation issues, caregiving responsibilities, a car accident, and orthopedic problems that did not arise from the class.

The final sample (n=31) comprised 27 women (87.1%) and 4 men (12.9%). Of these participants, 21 (67.7%) were Caucasian and 10 (32.3%) were African American. The subjects’ mean age was 73.5 years (SD=6.7, range=60–83), 100% were community dwellers (80% living in a house, 20% living in a high-rise apartment), and their average SF-36PF score at baseline was 74.8 (SD=24.8). Data for the 2 sites (Newton County, n=21;

Athens-Clarke County, n=10) were combined for analysis. The test-retest reliability of data for the procedure in the community was $r=.93-.98$ for the total score and all domain scores. The PFP-10 total scores and all domain scores are shown in Table 6. Four of the 5 summary scores improved (Bonferroni correction, $P<.01$). The total score increased by approximately 16.5%, with a change in domain scores from 20.4% (balance and coordination domain) to 8.1% (upper-body flexibility domain). Only the upper-body strength domain score did not improve. Participants’ adherence to attendance (74%) was not related ($r=.21$) to the extent of change in the PFP-10 total score. The SF-36PF scores (preintervention: $\bar{X}=74.8$, SD=24.8; postintervention: $\bar{X}=77.1$, SD=21.8) did not show a change in function as a result of this program.

Discussion and Conclusions

One of the critical challenges for physical therapist practice is to determine the efficacy of physical interventions. Of particular importance is determining whether functional change has occurred. Functional improvement should be the central goal of many physical interventions. Despite the importance of determining overall functional outcomes, few measures are available for this purpose. This is particularly problematic when measuring improvement for individuals who are already functioning independently and at a relatively high level. One such measure, the CS-PFP, is a measure of functional ability and yields data that are valid, reliable, and sensitive to change.^{7,8} This tool is comprised of tasks performed serially which is important because it can uncover limitations in capacity that are not apparently when tasks are performed individually. The PFP-10 retains this important attribute of the CS-PFP. The CS-PFP has been used as an outcome measure in several clinical trials.^{7-11,14,16} The application of this measure for clinical trials and in clinical practice settings is limited by the need for an established laboratory and by the length of time to administer the test (approximately 1 hour). These limitations were the impetus for creating a shorter version of this test, called the PFP-10.

The PFP-10 uses 10 of the CS-PFP's 16 tasks, takes approximately 30 minutes to administer, and requires less equipment for setup. The PFP-10 is being used in several randomized, controlled exercise interventions and other studies (personal communications; Petrella and associates, Wood and associates, White and associates, Kressig and associates). In addition, the shortened PFP-10 could be used in certain clinical settings (eg, to enhance evaluation of patients in rehabilitation settings to determine their readiness for independent living, to examine the impact of orthopedic conditions such as medial meniscal tears or low back pain).

The results of this investigation demonstrate that the psychometric properties of the PFP-10 meet requirements for clinical and research application. The PFP-10 has convergent validity with the CS-PFP and high test-retest reliability. Its sensitivity to change appears to be similar to that of the CS-PFP, although there was no control group in this investigation. These data suggest that the PFP-10 can be substituted for the CS-PFP without losing important information.

One of the primary advantages of the CS-PFP is its excellent sensitivity to change. This feature of the CS-PFP is retained in the PFP-10, allowing for relatively small sample sizes when testing the efficacy of interventions to change physical function. The PFP-10 was sensitive to change with a sample of only 31 participants who participated in the Lifetime Fitness Program, an evidence-based exercise program. The DI of the change in the PFP-10 total score and domain scores showed moderate levels of improvement. The PFP-10 detected change, whereas the SF-36PF did not. These data are similar to those previously reported from randomized controlled trials where the CS-PFP detected change in function and the SF-36PF did not.^{8,9}

The third substudy examined the use of the PFP-10 in a community-based setting. The ability to use the PFP-10 within the community greatly enhances its applicability, both for research studies that must be carried out at specific sites (eg, senior centers, community sites, congregate housing facilities) and for clinical application in facilities that cannot maintain a permanent testing laboratory.

There were no systematic differences between the scores in the established laboratory and those at the remote sites. The correlations were similar for each task and for the PFP-10 domain scores (Tab. 5). The main difference between the established laboratory and remote sites is the need for a portable washer and dryer, shelf, and stair platform that can be transported in a van. This substudy provided evidence that it is feasible to set up the PFP-10 in community settings.

The value of the PFP-10 is in its validity, reliability, and apparent sensitivity to change. The PFP-10 domain scales provide unique information compared with other performance-based measures. The time required to administer the PFP-10 was approximately 30 minutes, enhancing the potential for its use in clinical settings. Because it is expected that more Americans will live healthy and long lives, it is increasingly important to have the PFP-10 for clinical and research applications.¹⁸

One of the primary concerns was whether the PFP-10 scores would demonstrate convergent validity with the CS-PFP scores. This convergent validity was established in the first 2 substudies. The 2 versions were correlated when using the original data set ($n=228$, Tab. 3) and when tested with a new sample ($n=40$, Tab. 4). The differences in the PFP-10 total score and 4 of the 5 domain scores were within 5% of the original method, and when tested in a new sample the domain scores of the CS-PFP and the PFP-10 were within 1% to 3% of each other, except for the upper-body strength domain score, which was 17% lower in the PFP-10.

Differences in the upper-body strength domain score (Tabs. 3 and 4) may have resulted, in part, from the elimination of one task that had a ceiling effect. The upper-body strength domain for the CS-PFP is derived from 6 variables (ie, weight of jug in the water-pouring task, weight in the pan-carrying task, time to pass through a fire door, time to transfer laundry [1 and 2], weight carried in bus stop task, and weight carried in grocery task). The score for the water-pouring task, which had a demonstrable ceiling effect, was eliminated from the PFP-10 upper-body strength domain score calculation. Specifically, the average corrected score in substudy 1 ($n=228$) for the water-pouring task was 98 out of the possible 100. Because several tasks make up the domain, this ceiling effect did not translate to a ceiling effect for the entire test.

Other variables that were eliminated from the upper-body strength domain when calculating the PFP-10 scores included opening a heavy fire door and the weight carried up and down a 3-stair bus platform. Despite mean differences between the upper-body strength domain on the CS-PFP and the PFP-10, correlations were generally high, and removing tasks in general did not alter the other domains or the total score. Scores for the upper-body strength domain of the CS-PFP were correlated ($r=.55$) with measurements of biceps muscle peak torque.⁷ Scores for the upper-body strength domain of the PFP-10 and the CS-PFP were highly correlated ($r=.94$); therefore, a higher score on the upper-body strength domain of the PFP-10 reflects a higher biceps muscle peak torque. The upper-body strength domain of the PFP-10 may be a better repre-

sentation of upper-body function than the upper-body strength domain of the CS-PFP. The CS-PFP and the PFP-10 are perhaps the only performance-based tests that measure upper-body force production. Continuing research is needed to provide a better understanding of upper-body function and appropriate interventions to improve upper-body function. The advantages of shorter administration time and portability of the PFP-10 far outweigh the slight difference in upper-body strength domain scores.

Further supporting the validity of the PFP-10 scores, the CS-PFP and PFP-10 total scores were not different and were highly correlated with each other. This finding suggests that the PFP-10 also retains a relationship to the underlying physiological capacity measures (peak oxygen consumption and peak torque of the isokinetic knee extensor)^{7,17} against which the CS-PFP was validated. Because a score of 57 on the PFP-10 is essentially the same as a score of 57 on the CS-PFP, it would appear that the underlying thresholds for peak oxygen consumption ($20 \text{ mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$) and knee extensor peak torque adjusted for body height and weight ($2.5 \text{ N}\cdot\text{m}/\text{kg}\cdot\text{m}^{-1}$) remain comparable.¹⁷ For final validation, further study is needed.

Several limitations should be considered. This research is a compilation of several studies that took place in different locations and over a period of 2 years. Although this is a limitation to the study, it also demonstrates the strength of the CS-PFP and the shorter version, the PFP-10. No known problems with the data resulted from these differences in time and distance of administration. Procedural requirements of the PFP-10 include use of a standard dialogue and replication of the testing environment, which may make important contributions to reducing nonspecific variability and enhancing sensitivity to change.

The fourth substudy, which examined sensitivity to change, did not include a control group. Although the CS-PFP has demonstrated sensitivity to change in several randomized controlled trials,^{8,9,14,16} additional research is needed to confirm the PFP-10's sensitivity to change in a controlled sample in which the test administrator is blinded to group assignments.

In summary, the PFP-10 was developed to improve the feasibility of implementing this test for clinical outcomes studies and for clinical practice. The test was constructed to retain as much functional performance information as possible while reducing the burden of setup and administration. The serial testing of function, a unique feature of the CS-PFP test was retained in the PFP-10 measure provides a measure of capacity that can identify limitations in higher functioning individuals including

those at the threshold of disability. The CS-PFP and the PFP-10 tests are among the very few tools applicable to this growing segment of the population.

The protocol for the PFP-10 was tested in 2 different environments, showing that it can be useful in the administration of studies across multiple sites such as clinical trials and in testing older adults in the community. The results of this investigation should encourage investigators and clinicians to consider this measure to examine outcomes in clinical practice settings and for clinical trials that examine the overall functional impact of physical therapy interventions.

References

- 1 Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J Gerontol.* 1994;49:M85-M94.
- 2 Reuben DB, Siu AL. An objective measure of physical function of elderly outpatients: the Physical Performance Test. *J Am Geriatr Soc.* 1990;38:1105-1112.
- 3 Fried LP, Bandeen-Roche K, Kasper JD, Guralnik JM. Association of comorbidity with disability in older women: The Women's Health and Aging Study. *J Clin Epidemiol.* 1999;52:27-37.
- 4 Simonsick EM, Newman AB, Nevitt MC, et al. Measuring higher level physical function in well-functioning older adults: expanding familiar approaches in the Health ABC study. *J Gerontol A Biol Sci Med Sci.* 2001;56:M644-M649.
- 5 Duncan PW, Weiner DK, Chandler J, Studenski S. Functional reach: a new clinical measure of balance. *J Gerontol.* 1990;45:M192-M197.
- 6 Butland RJA, Pang J, Gross ER, et al. Two-, six-, and 12-minute walking tests in respiratory disease. *Br Med J (Clin Res Ed).* 1982; 284(6329):1607-1608.
- 7 Cress ME, Buchner DM, Questad KA, et al. Continuous-scale physical functional performance in healthy older adults: a validation study. *Arch Phys Med Rehabil.* 1996;77:1243-1250.
- 8 Cress ME, Buchner DM, Questad KA, et al. Exercise: effects on physical functional performance in independent older adults. *J Gerontol A Biol Sci Med Sci.* 1999;54:M242-M248.
- 9 Brochu M, Savage P, Lee M, et al. Resistance training in older women with coronary heart disease: a randomized controlled trial. *J Appl Physiol.* 2002;92:672-678.
- 10 Nakamura DY, Moore ML, Esselman PC, et al. Use of the CS-PFP (Continuous-Scale Physical Function Performance) Test as a functional outcome measure. *J Burn Care Rehabil.* 2000;21:S192.
- 11 Schenkman ML, Cutson TM, Kuchibhatla M, et al. A new measure of physical functional performance applied to those with PD. *Neurology Report.* 2002;26:130-138.
- 12 Dishman RK, Sallis JF. Determinants and interventions for physical activity and exercise. In: Bouchard C, Shephard RJ, Stevens T, eds. *Physical Activity, Fitness, and Health.* Champaign, Ill: Human Kinetics Inc; 1994:214-238.
- 13 Oka R, King A. Sources of social support as predictors of exercise adherence in women and men age 50 to 65 years. *Women's Health.* 1995;1:161-175.

- 14 Miszko TA, Cress ME, Slade JM, et al. Effect of strength and power training on physical function in community-dwelling older adults. *J Gerontol A Biol Sci Med Sci*. 2003;58:171–175.
- 15 Cress ME. Quantifying physical functional performance in older adults. *Muscle Nerve*. 1997;Suppl 5:S17–S20.
- 16 Ades PA, Savage P, Cress ME, et al. Resistance training improves performance of specific measured daily activities in disabled women with coronary heart disease. *Med Sci Sports Exerc*. 2003;35:1265–1270.
- 17 Cress ME, Meyers M. Maximal voluntary and functional performance levels needed for independence in adults aged 65 to 97 years. *Phys Ther*. 2003;83:37–48.
- 18 Centers for Disease Control and Prevention Web site. The state of aging and health in America, 2004. Available at: <http://www.cdc.gov/aging/publications.htm#sah>. Accessed December 23, 2004.
- 19 Borg G. Perceived exertion as an indicator of somatic stress. *Scand J Rehabil Med*. 1970;2:92–98.
- 20 CS-PFP Web site. Available at: <http://www.coe.uga.edu/cs-pfp>. Accessed December 23, 2004.
- 21 Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36), I: conceptual framework of item selection. *Med Care*. 1992;30:473–483.
- 22 Pinsky JL, Jette AM, Branch LG, et al. The Framingham Disability Study: relationship of various coronary heart disease manifestations to disability in older persons living in the community. *Am J Public Health*. 1990;80:1363–1367.
- 23 Cohn J. *Statistical Power Analysis for Behavioral Sciences*. 2nd ed. San Diego, Calif: Academic Press; 1988.
- 24 Bernstein IH, Garbin CP, Teng GK. *Applied Multivariate Analysis*. New York, NY: Springer-Verlag; 1988.
- 25 American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care*. 2004;27(suppl 1):S5–S10.
- 26 University of Georgia College of Education Web site. Available at: <http://www.coe.uga.edu>. Accessed December 23, 2004.
- 27 Wallace J, Buchner D, Grothaus L, et al. Implementation and effectiveness of a community-based health promotion program for older adults. *J Gerontol A Biol Sci Med Sci*. 1998;53: M301–M306.