An evaluation of three self-report physical activity instruments for older adults

NANCY D. HARADA, VICKI CHIU, ABBY C. KING, and ANITA L. STEWART

Geriatric Research, Education, and Clinical Center, VA Greater Los Angeles Healthcare System, Los Angeles, CA; UCLA/VA/RAND MEDTEP Center, UCLA School of Medicine, Los Angeles, CA; Stanford Center for Research in Disease Prevention, Stanford University, Palo Alto, CA; and Institute for Health and Aging, University of California San Francisco, San Francisco, CA

ABSTRACT

HARADA, N. D., V. CHIU, A. C. KING, and A. L. STEWART. An evaluation of three self-report physical activity instruments for older adults. *Med. Sci. Sports Exerc.*, Vol. 33, No. 6, 2001, pp. 962–970. **Purpose:** To assess the known-groups and construct validity of measures from the CHAMPS Physical Activity Questionnaire, Physical Activity Survey for the Elderly (PASE), and the Yale Physical Activity Survey (YPAS). **Methods:** The three questionnaires were administered to a convenience sample of older adults ($N = 87$) recruited from community centers and retirement homes. Validation measures included the SF-36 measures of physical functioning, general health, mental health, and pain; body mass index; performance-based tests of lower body functioning and endurance; and Mini-Logger activity monitor data from ankle and waist sensors. Validity was estimated by testing hypotheses about associations between physical activity and validation measures. **Results:** As hypothesized, differences in activity levels on all measures were found between older adults in retirement homes (less active) and community centers (more active) (P-values < 0.0001). Correlations of physical activity measures with performance-based measures ranged from 0.44 to 0.68, conforming to hypotheses; hypotheses regarding associations with the SF-36 measures were also confirmed. Body mass index was not correlated with any of the physical activity measures, contrary to hypotheses. Correlations of physical activity measures with Mini-Logger counts ranged from 0.36 to 0.59 (ankle) and 0.42 to 0.61 (waist) as hypothesized. Correlations among the measures from the three instruments ranged from 0.58 to 0.68. **Conclusions:** The PASE, YPAS, and CHAMPS each demonstrated acceptable validity, as all measures met nearly all hypotheses. Higher validity coefficients were found for subgroups (men, 65-74 yr, retirement home), suggesting that these instruments may perform better for certain segments of the older adult population. **Key Words:** MEASUREMENT, AGING, EXERCISE

Research studies often use surveys to measure the physical activity profile of study participants. These surveys typically ask respondents about the types of physical activity they engage in as well as the intensity, frequency, and duration of these activities over a specified time frame. For older adults, physical activity measurement is complicated because they often engage in lighter activities more frequently than moderate or vigorous activities and they may perform activities on an irregular basis, making it difficult to recall. Older adults may perform physical activity at a substantially lower level of energy expenditure as compared with younger adults. Furthermore, they may have difficulty with memory and cognition, which interferes with their ability to recall past physical activity, especially over long periods of time.

To circumvent these problems, physical activity surveys have been developed specifically for administration to older adults (1,2,20,23,28). The Community Healthy Activities Model Program for Seniors (CHAMPS) Questionnaire, the Physical Activity Scale for the Elderly (PASE), and the Yale Physical Activity Survey (YPAS) were developed for community-dwelling older adults in the United States (2,20,28). These three surveys address measurement issues for older adults by asking about less vigorous activities, such as gardening and walking. To facilitate recall of light-intensity activities, these surveys provide lists of specified activities (YPAS and CHAMPS) or prompt the respondent with examples of specific activities (PASE). In addition, these questionnaires ask about duration of activity performed over 1 wk rather than per session to facilitate recording of irregular activity.

On the surface, these instruments appear similar in their approach to measurement of physical activity. Therefore, we assessed simultaneously the CHAMPS Physical Activity Questionnaire, PASE, and YPAS against the same set of validation criteria to provide researchers and practitioners with a starting point from which to select a tool for inclusion in a research study. In a comparison paper in this issue, Stewart et al. report on the development of the CHAMPS measure and its psychometric characteristics (20). In this study, we evaluate all three physical activity instruments in a sample of older adults recruited from community centers and retirement homes. A secondary study objective was to report on the test-retest reliability of the CHAMPS instrument.

0195-9131/01/3306-0962/$3.00/0
MEDICINE & SCIENCE IN SPORTS & EXERCISE®

Received for publication September 1999.
Accepted for publication September 2000.

TABLE 1. Summary of self-reported physical activity measures.

	PASE	YPAS	CHAMPS
Intended application	To measure age-specific physical activity levels of older adults	To measure age-specific physical activities of older adults	To use as an outcome measure of a physical activity promotion intervention
Method of administration	* Self-administered Telephone-administered Face-to-face interview	* Face-to-face interview Telephone-administered	* Self-administered Telephone-administered Face-to-face interview
Time to complete	5 min	20 min	15 min
Original development and testing sample	Initial version piloted on 36 elderly community-dwelling men and women Reliability studies: 254 men and women Validation studies: 222 men and women ≤65 yr of age living in their own homes	Instrument developed on 222 men and women between 60 and 86 years from senior centers and senior residential communities Reliability studies: subsample of 20 men and 56 women Validation studies: subsample of 14 men and 11 women	Development, reliability and validity testing was conducted on 249 members of a Medicare HMO recruited into a physical activity intervention or an active cohort study
Time frame	Past 7 d	Typical week in past month	Typical week in past month
Derived measures	• Total PASE score = (time spent in each activity (hr/wk) or participation (yes/no) × PASE weight) summed for all activities. (PASE weight is derived from regressing a component score developed from Caltrac, MET total from activity diaries, and global self-reports of activity on responses to PASE) Time spent is assessed in walking outside home, light sport/recreational, moderate sport/recreational, strenuous sport/recreational, muscle strength and endurance exercises. Any participation is assessed in light housework, heavy housework, home repairs, lawn work or yard care, outdoor gardening, caregiving, work for pay or as volunteer.	• Total time summary index (hr/wk) for all activities in checklist • Energy expenditure summary index (kcals/wk) = (time × intensity code) summed for all activities in checklist. Intensity code based on a 60-kg person. • Indices for 5 types of activities as follows: (frequency × duration × weight) (weights based on relative intensity of activity dimension). - vigorous activity: weight = 5 - leisurely walking: weight = 4 - moving on feet: weight = 3 - standing: weight = 2 - sitting: weight = 1 • Seasonal adjustment score • Number of flights of stairs/day Activities included in energy expenditure summary include work, e.g., shopping and home repair; yardwork, e.g., gardening; caretaking, e.g., for older person; exercise, e.g., brisk walking and pool exercises; recreational activities, e.g., needlework and dancing.	• Caloric expenditure per week in moderate intensity or greater and for "all" physical activities • Frequency of activity per week in moderate intensity or greater and in "all" physical activities Moderate activities and greater include walking briskly, jogging, dancing, golfing-walking, singles/doubles tennis, riding a bicycle/stationery cycle, swimming, water exercises, aerobic exercise, heavy housework, and gardening. "All" activity measures include the above plus walking leisurely, golfing with a cart, light housework, yoga, stretching/flexibility, and general conditioning exercises.

* Original testing method of administration.

METHODS

Subjects

Approval for this study was obtained from the Human Subjects Protection Committee at the University of California, Los Angeles. Study procedures were described in detail to potential subjects, and informed consent forms were signed before participation in the study. A convenience sample of 87 older adults was recruited from two community centers and three retirement homes located in Los Angeles through announcements, flyers, and word-of-mouth. These facilities were selected by design to obtain subjects who were known to be more active (those in community centers) and less active (those in retirement homes). One community center and one retirement home served primarily the Japanese-American community, and the other settings served the general population. The size of the total sample was based on a statistical power analysis of the number of cases needed to detect correlations of 0.3 with 80% power. The sample was stratified by age group (65–74 yr and 75 yr and over) and gender. Inclusion criteria were: 1) age 65 yr and over; 2) good cognitive status; and 3) no pacemaker.

Self-Report Physical Activity Surveys

The properties of the three surveys are outlined in Table 1. The PASE was developed to assess physical activity in epidemiologic research and includes items asking about participation in household activities, occupational activities, and leisure time activities over the past 7-d period (27,28). The PASE can be self-administered, telephone-administered, or interviewer-administered. Scoring is based on item weights that are empirically derived through principal components analysis and regression techniques (27,28). The total score is computed in two steps by: a) multiplying an activity frequency value from a conversion of hours per day in six categories of activity (e.g., moderate sports) by the respective weight and summing over these activities; and b) adding a weight to this summated score for each of six other household activities if the activity was reported over the past 7 d. In community-dwelling older adults, PASE scores were moderately correlated with several self-reported health status measures (r ranges between −0.34 and −0.42) (28) as well as grip strength (r = 0.37), static balance (r = 0.33), and leg strength (r ranges between 0.25 and 0.41) (28). In sedentary older adults, low but significant correlations were found between PASE scores and peak oxygen uptake (r = 0.20), systolic blood pressure (r = −0.18), and balance (r = 0.20) (27). The 3- to 7-wk test-retest reliability is 0.75, with the mail-administered version higher than the telephone-administered version (r = 0.84 and 0.68, respectively) (14,28). The internal consistency reliability is 0.69 (28).

The YPAS is an interviewer-administered scale that was developed to assess physical activity in epidemiological

studies (2). The YPAS is divided into two sections: in the first section, there is a comprehensive physical work, exercise, and recreational activities checklist to assess time spent in these types of activities during a typical week in the past month. Recreational activities ranged from low-intensity activities, such as needlework, to high-intensity activities, such as tennis. The second section contains questions to quickly assess an individual's participation in five activity dimensions: vigorous activity, leisurely walking, moving on feet, standing, and sitting. Responses on the YPAS allow eight summary indices to be calculated: total time spent per week in all physical activities, weekly energy expenditure in kcal per week, five individual indices for the activity dimensions, and an activity dimension summary index. Weekly kcal expenditure on the YPAS has been found to be moderately correlated with resting diastolic blood pressure ($r = -0.47$), and poorly correlated with maximum oxygen uptake ($r = 0.20$), percent body fat ($r = -0.07$), and the Caltrac motion sensor ($r = 0.14$) (2). The activity dimensions summary index, however, is moderately correlated with maximum oxygen uptake ($r = 0.58$), percent of body fat ($r = -0.43$), and the motion sensor ($r = 0.37$) but poorly correlated with diastolic blood pressure ($r = -0.21$) (2). The 2-wk test-retest reliability was 0.58 for the kcal measure and 0.65 for the activity summary index, with higher-intensity exercise-related behaviors having higher reliability than low-intensity activities (2).

The CHAMPS survey is the newest self-report measure and was developed to assess outcomes of a physical activity promotion intervention designed to change activity behavior (20,22). Thus, it focuses specifically on the types of activities that the intervention was designed to affect, rather than assessing all physical movements and energy expenditure in a day. The CHAMPS approach is to provide respondents with a comprehensive list of various light, moderate, and vigorous physical activities to which respondents report their weekly frequency of participation and weekly duration in a typical week over the last 4 wk. Rather than asking the duration of participation each time they perform an activity (which assumes a regularity), the duration is asked with respect to the total week. To make the reporting easier, several categories are provided (less than 1 $h \cdot wk^{-1}$, 1–2.5 $h \cdot wk^{-1}$, 3–4.5 $h \cdot wk^{-1}$, 5–6.5 $h \cdot wk^{-1}$, 7–8.5 $h \cdot wk^{-1}$, and 9 or more $h \cdot wk^{-1}$). Scoring of the CHAMPS survey yields four measures: estimated kcal per week and frequency per week spent in moderate-intensity (or greater) activities, and estimated kcal per week and frequency per week spent in "all" physical activities. Moderate-intensity activities are defined as those with an intensity of three METs or greater and include walking briskly, jogging, dancing, golfing without using a cart, singles and doubles tennis, riding a stationery cycle or bicycle, swimming, water exercises, aerobics, heavy household chores, and gardening. The "all activities" measure includes the above plus walking leisurely, golfing with a cart, light housework, yoga, stretching/flexibility exercises, and general conditioning exercises. The CHAMPS measures discriminate well across groups of older adults known to vary in activity levels ($P < 0.001$) (20). Correlations of the CHAMPS kcal per week in moderate-intensity activities measure with physical performance measures of lower body functioning and endurance range from 0.27 to 0.28, and correlations of the "all activities" measures range from 0.22 to 0.27 (20). The 6-month stability of the CHAMPS moderate and "all activities" measures based on intraclass correlation coefficients range from 0.60 to 0.69 (20); this value can be considered as a lower bound estimate of reliability.

Validation Measures

For this study, the overall energy expenditure measures of each physical activity survey were tested against validation measures that could potentially be affected by physical activity. The overall energy expenditure measures of each survey were included in this study because they are the most commonly used in studies of physical activity (13,14,19,21).

Activity monitoring. Each subject wore a Mini-Logger Series 2000 monitor for a seven-day period (Mini-Mitter Company, Sunriver, OR). The Mini-Logger measures activity by counting the number of mercury switch closures, resulting in a "count" of activity, over a predetermined time interval. For this study, the Mini-Logger was configured to record activity level at 1-min intervals through single axis sensors located at the waist and ankle. In addition, heart rate was transmitted to the Mini-Logger by a Polar chest band (Kempele, Finland) with built-in leads.

Manufacturer testing has shown that the Mini-Logger is 85% reproducible, e.g., different sensors held in the same orientation and moved through the same series of motions show a range of activity counts with about 85% agreement. Two Mini-Logger scores were calculated for this study: 1) total counts occurring at the ankle sensor over a 6-d period, and 2) total counts occurring at the waist sensor over a 6-d period. A spreadsheet program designed by researchers at the Stanford Center for Research in Disease Prevention was used to exclude activity counts that did not have corresponding valid heart rate data. In setting up the program, the researchers assumed that the mode heart rate was the resting heart rate, and the program was designed to eliminate activity count data when the recorded heart rate was less than 50% of the mode. Mini-Logger counts at the ankle and waist were then totaled for the 6-d period. In cases where the subject did not wear the Mini-Logger for the entire 6-d period ($N = 3$), total counts were imputed if the Mini-Logger was worn for at least 3 of the 6 d. Total counts over the 6-d period were imputed by summing the total counts for the number of days worn and then adding back the average counts per day for up to 3 missing days.

Short Physical Performance Battery. The Short Physical Performance Battery was developed for the Established Populations for Epidemiologic Studies of the Elderly (EPESE) to measure lower body functioning and consists of three performance-based measures of lower body strength (chair stands), balance (side-by-side, semi-tandem, and tandem), and walking (8-foot walk) (5). Scoring is through a

summary performance scale, which aggregates the quintile scores of all three performance-based tests (5).

Chair stands were conducted by having the subject stand up and down from a chair five times as quickly as possible with the arms placed across the chest. A tester used a stopwatch to time the maneuver, starting when the subject began the first rise from the chair and stopping when the subject stood on the fifth repetition and all body movement had ceased. Standing balance was tested with the feet in three positions: side-by-side, semi-tandem, and tandem. Subjects stood in each position while they were timed with a stopwatch. The timing stopped when the subject moved his or her feet, lost balance, or when 10 s had elapsed. The 8-foot walk was conducted by having the subject walk at their usual speed along an 8-foot walking course with an additional 2 feet at either end marked off. Each participant was timed for two trials. The faster of the two trials was used in the analysis.

6-Minute walk. The 6-min walk was conducted by having the subject walk around a series of traffic cones that were placed to mark off a circular walking area of about 40 feet in diameter (6). Subjects were instructed to walk for 6 min, covering as much ground as possible at a work effort that allowed the person to talk without becoming short of breath. The tester walked alongside the subject, using a stopwatch to time the walk and a counter to record the number of completed laps. At the end of 6 min, a marker was placed on the ground next to the subject. The distance walked during the last lap was measured by a rolling tape measure. The total distance in feet was derived by multiplying the number of laps by the circumference of the walking circle and adding the distance covered on the last lap. One-week test-retest reliability of the 6-min walk was 0.95 in community-dwelling older adults (6).

Body mass index. Body mass index was calculated using the following formula: Body mass index = weight in kg/height in m^2 (4). The subject was weighed using a portable digital scale. A tape measure attached to the wall was used to measure height. This measurement was obtained by alignment of a straight edge from the top of the head.

Self-reported functioning and well-being from the SF-36 Health Survey. The SF-36 health status survey assesses eight domains of health-related quality of life. For this study, four measures were included: physical functioning, general health perceptions, mental health, and pain (24,25). Physical functioning refers to limitations in a variety of activities such as walking, bending, and climbing stairs. General health perceptions are personal beliefs about general health status. Mental health includes positive and negative affective states, such as feeling happy, peaceful, and anxious. Pain refers to subjective feelings of distress or discomfort. The internal consistency reliability of these measures for people over the age of 65 yr is 0.92 for physical functioning, 0.78 for general health perceptions, 0.88 for mental health, and 0.85 for pain (15).

The validity hypotheses were: 1) more active people have higher self-report scores than less active people; 2) the PASE, YPAS, and CHAMPS scores are positively and moderately correlated with activity monitoring, lower body functioning, and endurance; 3) the PASE, YPAS, and CHAMPS scores are negatively and moderately correlated with body mass index; 4) the PASE, YPAS, and CHAMPS are more highly correlated with the physical functioning and general health measures of the SF-36 than the pain and mental health measures; and 5) the three self-report surveys are moderately correlated with each other.

Procedure

Participants were scheduled for two visits, spaced 1 wk apart. All visits were held at either the community center or retirement home from which the subject was recruited. Approximately 1 wk before the first scheduled visit, participants were mailed a packet that contained an informed consent form and a questionnaire asking about demographic and medical information. Participants were instructed to read the informed consent and to bring it and the completed questionnaire to their first visit.

At the first scheduled visit, the study was verbally described to the participant and the informed consent signed. After the informed consent process, each participant was screened for cognitive impairment by using the Folstein Mini Mental State Examination (3). Only one individual was excluded based on a score of less than 24. Individuals receiving a score of 24 or higher went on to receive verbal and written instructions on how to wear the Mini-Logger. Participants were instructed to wear the Mini-Logger during their waking hours for the 1-wk period until their next scheduled visit. All of the subjects were given the Mini-Logger with the built-in waist sensor. However, some of the subjects were not assigned to wear the ankle sensor due to lack of comprehension of procedures ($N = 3$), frailty ($N = 7$), or refusal ($N = 5$). Next, the subject completed the first of three questionnaires asking about their physical activities. The order of administration of the physical activity questionnaires was directed by a slip of paper blindly selected out of a paper bag. The order was set so that the YPAS was always administered first or third because it is interviewer-administered. The participant was instructed to complete the second physical activity questionnaire and the health measures at home and to bring these back to the next visit. At the end of the first visit, participants were scheduled for the second visit 1 wk later and given $10.00 for their participation.

At the second visit, a research assistant reviewed the take-home surveys for completeness, and the participant completed the final physical activity questionnaire. Tests of balance, lower-extremity strength (chair stands), walking (8-foot walk), and endurance (6-min walk) were conducted. Weight and height were recorded. At the end of the visit, the Mini-Logger and Polar band were collected from each subject. Data from the Mini-Logger were downloaded into a laptop computer, and the participant was paid $10.00.

To test the reliability of the CHAMPS measure, the survey was mailed to all participants 2 wk after the second visit.

TABLE 2. Demographic and medical characteristics.

	Retirement Homes (Less Active) ($N = 36$)	Community Centers (More Active) ($N = 51$)	Total ($N = 87$)
Age (yr)[c]			
Mean (SD)	79 (6)	73 (5)	75 (6)
Range	65–89	65–86	65–89
Gender (%)			
Female	61	63	62
Race (%)			
White	34	35	34
African American	8	2	5
Asian American	58	63	61
Education (yr)			
Mean (SD)	13 (3)	13 (3)	13 (3)
Range	5–19	4–20	4–20
Marital status (%)[b]			
Married	6	51	32
Folstein score[b]			
Mean (SD)	27 (2)	29 (1)	28 (2)
Range	24–30	24–30	24–30
Self-reported medical conditions (%)			
Visual deficits	36	22	28
Hearing deficits	31	24	26
Arthritis	39	41	40
Hypertension	49	29	37
Asthma	14	4	8
Diabetes[a]	28	12	18
Falling-last 12 mos	29	24	26
Back problems	33	33	33

[a] $P < 0.05$.
[b] $P < 0.001$.
[c] $P < 0.0001$.

Participants were instructed to mail back the completed questionnaire in a preaddressed, stamped envelope.

Data Analysis

Descriptive statistics, including frequencies and measures of central tendency and dispersion, were conducted to evaluate the distributional characteristics of each measure. To determine known-groups validity, a *t*-test was conducted to determine whether there were significant differences in activity level between community center (more active) and retirement home (less active) groups. To determine construct validity, Pearson's correlation coefficients were calculated between the physical activity scores on each survey with each of the validation measures. Pearson's correlation coefficients were also calculated to assess the relationship of the three surveys to each other. To assess test-retest reliability of the CHAMPS questionnaire, both Pearson's and intraclass correlation coefficients were calculated. All analyses were conducted using the SAS statistical software system (18).

RESULTS

Demographic and medical characteristics of the sample are presented in Table 2. The average age of the total sample was 75 yr; the average age of the retirement home participants was 6 yr older than community-center participants. Sixty-two percent of the total sample was female. The racial breakdown of the participants was 61% Japanese-American, 34% white, and 5% African-American. There were large differences in marital status by site. Fifty-one percent of participants at the community centers were married as compared with only 6% at the retirement homes. The mean Folstein cognitive status score was 28, indicating good cognitive functioning. The most common medical conditions and problems were arthritis (40%), hypertension (37%), back problems (33%), visual deficits (28%), hearing deficits (26%), and falling in the last 12 months (26%).

Known-Groups Validity

Descriptive data for the three activity surveys are displayed by group in Table 3. As hypothesized, community-center participants had significantly higher physical activity scores than those in retirement homes on all measures (*P*-values < 0.0001). The standard deviation for retirement home participants was smaller than in the community center, indicating a wider range of activity levels in the community centers. In the retirement home, the lower boundary for the activity range was zero on all three measures, indicating that some individuals were not physically active. In the community centers, a lower boundary of zero was observed only for the CHAMPS moderate-intensity measure. In contrast, the upper boundary of the activity range for older adults in the community centers was high, indicating that some subjects were extremely physically active.

Construct Validity

Correlations between the activity surveys and construct validity measures are reported in Tables 4a and 4b for the whole sample, and by site, gender, and age group. For all subjects, higher correlations were found for each physical activity questionnaire score and the waist counts of the Mini-Logger than for ankle counts, except the PASE. As hypothesized, correlations between the activity scores and waist counts ranged between 0.42 and 0.61, and with the ankle counts between 0.36 to 0.59. In general, higher correlations were observed for older adults in retirement homes than older adults in community centers; higher correlations were observed for men than women; and higher correlations

TABLE 3. Known-groups validity: scores on CHAMPS, PASE, and YPAS.

	Retirement Homes ($N = 36$)	Community Centers ($N = 51$)	*t*-Test
CHAMPS kcal/wk in all activities			
Mean (SD)	1548 (1767)	3484 (2042)	$t = 4.60$, $P < 0.0001$
Median	805	3243	
Range	0–6477	376–10,243	
CHAMPS kcal/wk in moderate activities			
Mean (SD)	503 (776)	2525 (1847)	$t = 6.99$, $P < 0.0001$
Median	169	1952	
Range	0–2689	0–8067	
PASE total score			
Mean (SD)	50 (44)	158 (65)	$t = 9.26$, $P < 0.0001$
Median	45	150	
Range	0–195	54–372	
YPAS kcal/wk			
Mean (SD)	2313 (2277)	8125 (4125)	$t = 8.41$, $P < 0.0001$
Median	1682	7245	
Range	0–11,310	329–19,680	

TABLE 4a. Validity correlations of CHAMPS with activity monitors, performance-based and self-report measures.

Validation Measures	All Subjects ($N = 87$)	Site		Gender		Age Group	
		C.C. ($N = 51$)	R.H. ($N = 36$)	Male ($N = 33$)	Female ($N = 54$)	65–74 ($N = 43$)	75+ ($N = 44$)
CHAMPS kcal/wk in all physical activities							
Mini-Log ankle[a]	0.36**	0.16	0.16	0.48*	0.27	0.41	0.26
Mini-Log waist[b]	0.42***	0.10	0.21	0.59***	0.31*	0.52***	0.22
EPESE lower-body functioning	0.46**	0.32*	0.38*	0.54**	0.41**	0.55**	0.36*
6-min walk	0.46**	0.20	0.26	0.63**	0.37**	0.57**	0.29*
Body mass index	0.006	0.002	−0.02	−0.002	0.01	0.001	−0.004
SF-36: PF	0.39**	0.25	0.29	0.45**	0.36**	0.58**	0.20
SF-36: GH	0.35**	0.24	0.19	0.34**	0.36**	0.42**	0.26
SF-36: MH	0.25**	0.12	0.15	0.31	0.23	0.31*	0.18
SF-36: Pain	0.26**	0.17	0.19	0.37*	0.20	0.38**	0.11
CHAMPS kcal/wk in moderate activities							
Mini-Log ankle[a]	0.42**	0.20	0.36	0.50*	0.36*	0.40	0.41*
Mini-Log waist[b]	0.48***	0.11	0.30	0.59***	0.41**	0.50***	0.38*
EPESE lower body functioning	0.44**	0.37**	0.32	0.48**	0.42**	0.50**	0.36**
6-min walk	0.54**	0.25	0.30	0.64**	0.49**	0.57**	0.41**
Body mass index	−0.02	−0.06	−0.02	0.06	−0.05	−0.07	−0.0007
SF-36: PF	0.41**	0.26	0.34*	0.48**	0.37**	0.52**	0.22
SF-36: GH	0.42**	0.26	0.36*	0.45**	0.40**	0.42**	0.37
SF-36: MH	0.28**	0.14	0.17	0.29	0.27*	0.27	0.24
SF-36: Pain	0.28**	0.18	0.25	0.39*	0.22	0.34*	0.14

* $P < 0.05$; ** $P < 0.01$, *** $P < 0.001$.
[a] $N = 56$.
[b] $N = 78$.

were observed for the 65- to 74-yr-old age group than the 75 yr and over age group.

Also as hypothesized, for all subjects the correlations of activity scores on all measures with the two performance-based measures ranged from 0.44 to 0.68. Body mass index was not correlated with any of the physical activity measures (r ranging between 0.006 to 0.10), contrary to our hypothesis.

As hypothesized, the physical functioning and general health measures had higher correlations with all physical activity questionnaire measures than the mental health and pain measures among the total sample. The correlations between all physical activity measures and physical functioning and general health measures ranged between 0.26 to 0.42, whereas the correlations with the mental health and pain scores ranged from 0.17 to 0.28. The magnitude of the correlations varied somewhat by gender and age group with men usually having higher correlations than women, and the 65- to 74-yr-old age group having higher correlations than the 75 yr and over age group.

The intercorrelations for the three activity surveys are presented in Table 5. The correlation coefficients among the physical activity measures ranged from 0.58 to 0.68. The two CHAMPS measures were highly correlated ($r = 0.92$) as expected given their overlap.

Reliability of the CHAMPS Survey

The 2-wk test-retest reliability coefficients for the CHAMPS measures are displayed in Table 6. Pearson's and intraclass correlation coefficients are provided for the total sample, and by gender and age group. The reliability for the "all" physical activities measure was 0.62 (Pearson's and intraclass) in the total sample and ranged from 0.59 to 0.69 (Pearson's) and 0.59 to 0.67 (intraclass) in the subgroups. The reliability for the moderate-intensity physical activities measure was 0.76 (Pearson's and intraclass) in the total sample and ranged from 0.72 to 0.81 (Pearson's) and 0.72 to 0.79 (intraclass) in the subgroups. Higher correlations were found for the moderate-intensity measure than for the "all" physical activities measure.

DISCUSSION

In this study, three physical activity surveys were tested for evidence of validity against the same set of criteria. Each of the hypotheses except one was confirmed, thus meeting our criteria for acceptable known-groups and construct validity. This suggests that all three measures exhibited considerable evidence of construct validity in several tests, using three very diverse sources of data: activity monitoring, performance tests, and other self-report measures. There were some variations in the magnitude of the correlations. For example, the PASE demonstrated the highest correlations with the two performance-based measures (EPESE index and 6-min walk).

Higher correlations with validation measures were observed in certain subgroups (e.g. men, the 65- to 74-yr-old group, and people living in retirement homes), suggesting that these subgroups may be better at reporting their physical activities. Men and 65- to 74-yr-old adults may perform more regular and intense physical activities than women and people 75 yr of age and older, and these types of activities are much easier to recall. Other studies have found that the

TABLE 4b. Validity correlations of PASE and YPAS with activity monitors, performance-based, and self-report measures.

Validation Measures	All Subjects ($N = 87$)	Site		Gender		Age Group	
		C.C. ($N = 51$)	R.H. ($N = 36$)	Male ($N = 33$)	Female ($N = 54$)	65–74 ($N = 43$)	75+ ($N = 44$)
PASE total score							
Mini-Log ankle[a]	0.59***	0.38*	0.55*	0.78***	0.49**	0.60**	0.56***
Mini-Log waist[b]	0.52***	0.05	0.24	0.70***	0.42**	0.59***	0.42**
EPESE lower body functioning	0.57**	0.29*	0.57**	0.64**	0.53**	0.68**	0.48**
6-min walk	0.68**	0.27*	0.54**	0.77**	0.63**	0.75**	0.60**
Body mass index	−0.07	−0.18	−0.02	−0.14	−0.04	−0.18	0.007
SF-36: PF	0.30**	−0.03	0.16	0.42*	0.23	0.62**	0.03
SF-36: GH	0.26*	−0.07	−0.01	0.32	0.22	0.40**	0.12
SF-36: MH	0.23*	0.03	0.01	0.45**	0.12	0.27	0.17
SF-36: Pain	0.17	−0.08	0.11	0.36*	0.06	0.23	0.07
YPAS kcal/wk							
Mini-Log ankle[a]	0.46***	0.17	0.55**	0.38	0.55***	0.63***	0.33
Mini-Log waist[b]	0.61***	0.28*	0.20	0.64***	0.60***	0.60***	0.58***
EPESE lower body functioning	0.49**	0.20	0.51**	0.50**	0.48**	0.55**	0.42**
6-min walk	0.58**	0.27*	0.16	0.61**	0.57**	0.62**	0.49**
Body mass index	−0.10	−0.19	−0.10	−0.12	−0.10	−0.19	−0.05
SF-36: PF	0.31**	0.02	0.23	0.37*	0.27*	0.48**	0.12
SF-36: GH	0.31**	−0.04	0.23	0.26	0.34*	0.36**	0.22
SF-36: MH	0.24*	−0.11	0.31	0.25	0.24	0.25	0.19
SF-36: Pain	0.23*	0.02	0.24	0.25	0.20	0.27	0.13

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.
[a] $N = 56$.
[b] $N = 78$.

reliability of vigorous activities is higher than moderate-and low-intensity physical activities (7).

The one hypothesis that was not confirmed was the relationship of physical activity scores to body mass index. Although our hypothesis was based on several prior studies documenting a positive correlation between energy expenditure and body mass index (12,17), our findings are nevertheless consistent with other studies that have found no association between body mass index and activity scores (26–28). Taken together, these findings suggest that the empirical evidence concerning the relationship between physical activity and body mass index is still inconclusive.

Correlations between the CHAMPS measures and validation measures in our study are consistently higher than those found by Stewart et al. (20). We do not believe these differences are due to measurement technique, because our methods were standardized to those used in the CHAMPS study. These differences may be due to differences in the sources of the samples or to statistical anomalies in the bivariate distributions that we have not detected. Clearly, more studies are needed to continue to acquire evidence of the construct validity of these measures in a variety of populations of older adults.

Our findings do not strongly support the use of one measure over the others. Indeed, evidence of validity was confirmed for all measures and patterns were fairly similar. Thus, in attempting to choose among these surveys for a particular study, investigators will need to compile information from this and other studies. One strategy would be to pretest each measure on the target population to determine whether any one instrument presents particular problems in that group.

However, other characteristics not tested in this study are also important to consider when selecting a measurement tool. Sensitivity to change is an important characteristic when selecting an instrument to measure outcomes of a physical activity intervention (11). The CHAMPS measures have demonstrated sensitivity to change after a 6-month (21) and a 12-month (22) physical activity promotion program for seniors, and a 12-month moderate-intensity endurance and strengthening physical activity program for community-dwelling older adults (10). They were also found to be sensitive to change at both 6 and 12 months in response to a moderate-intensity walking program for older women

TABLE 5. Intercorrelations among the PASE, YPAS, and CHAMPS.

	PASE	YPAS	CHAMPS "Moderate" Activities
PASE	—	—	—
YPAS	0.61***	—	—
CHAMPS "moderate" activities	0.64***	0.68***	—
CHAMPS "all" activities	0.58***	0.64***	0.92***

*** $P \leq 0.0001$.

TABLE 6. Two-week test-retest reliability of CHAMPS measures.

		Kcals/wk-"All" Activities		Kcals/wk-"Moderate" Activities	
	N	Pearson's	ICC*	Pearson's	ICC*
Total	80	0.62	0.62	0.76	0.76
By gender					
Male	30	0.69	0.67	0.81	0.79
Female	50	0.62	0.59	0.75	0.74
By age group					
65–74 yrs	42	0.59	0.59	0.72	0.72
75 and over	38	0.62	0.61	0.78	0.78

* Intraclass correlation coefficient.

serving as family caregivers for an older infirm relative (8,9). The PASE, in contrast, was not found to be sensitive to change in either the 12-month trial aimed at moderate-intensity endurance and strengthening (10) or the 12-month trial aimed at increasing moderate-intensity walking in older women caregivers (8). The PASE has demonstrated sensitivity to change after a pilot study of a 6-wk physician counseling intervention (13). The YPAS has demonstrated sensitivity to change at 3 months after a moderate-intensity aerobic exercise intervention (29).

The 2-wk test-retest reliability of the CHAMPS moderate-intensity measure was 0.76 (Pearson's and intraclass). Given that correlations above 0.70 are sufficient in early stages of development for group studies (16), this indicates a good reliability value. The reliability was 0.62 (Pearson's and intraclass) for the "all" activities measure, indicating that this could be improved. Our values are similar to the 6-month stability estimates in the CHAMPS study for the moderate-intensity measure of 0.67 (intraclass correlation coefficient, which tends to be lower) and for the "all" activities measure of 0.66 (intraclass correlation coefficient) (20). The reliability of the moderate-intensity measure is roughly comparable to the test-retest reliability coefficient for the PASE ($r = 0.75$) (28), and the all activities measure is roughly comparable to the 3- to 7-wk test-retest reliability of the YPAS kcal index ($r = 0.58$) (2).

The subjects in this study represent a convenience sample of older adults drawn by design from community centers and retirement homes. Further, our subjects volunteered to have their physical activity assessed; thus, they may be among the more physically active even in those settings. Because all of the Japanese-American participants were English speaking, it is improbable that ethnicity would limit the generalizability of this study. However, further studies are needed to assess the validity of physical activity questionnaires in other minority groups. In this respect, the methodology employed in this study can be used as a model for future work of this type.

Our findings that validity coefficients sometimes differed by site, gender, and age suggest that further studies are warranted with larger subgroup sample sizes. These studies should report validity coefficients not only for the total sample but also for separate subgroups of older adults. In addition, because we did not include all of the indices that can be derived from each of the instruments, important aspects of validity from these instruments are yet to be explored. This study, however, is a first step toward understanding the different approaches to measurement of physical activity in older adults.

This study was supported by the National Institute on Aging, grant number 1735SC, a supplement to AG09931. The authors would like to acknowledge Leslie A. Pruitt, Ph.D., and Tadashi Nakamura, M.D., for their work in developing the software program for analyzing activity monitor data. We also thank William L. Haskell, Ph.D., for his expertise throughout this project. The results of this study do not constitute endorsement of the Mini-Logger equipment by the authors or ACSM.

Address for correspondence: Nancy Harada, Ph.D., GRECC (11G), VA Greater Los Angeles Healthcare System, 11301 Wilshire Blvd., Los Angeles, CA 90073; E-mail: nharada@ucla.edu.

REFERENCES

1. Caspersen, C., B. Bloemberg, W. Saris, R. Merritt, and D. Kromhout. The prevalence of selected physical activities and their relation with coronary heart disease risk factors in elderly men: the Zutphen study, 1985. *Am. J. Epidemiol.* 133:1078–1092, 1991.
2. Dipietro, L., C. Caspersen, A. Ostfeld, and E. Nadel. A survey for assessing physical activity among older adults. *Med. Sci. Sports Exerc.* 25:628–642, 1993.
3. Folstein, M., S. Folstein, and P. Mchugh. "Mini-Mental state": a practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* 12:189–198, 1975.
4. Garrow, J., and J. Webster. Quetelet's Index (W/H^2) as a measure of fatness. *Int. J. Obes.* 9:147–153, 1985.
5. Guralnik, J., E. Simonsick, L. Ferrucci, 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. Med. Sci.* 49:M85–M94, 1994.
6. Harada, N. D., V. Chiu, and A. L. Stewart. Mobility-related function in older adults: assessment with a six-minute walk test. *Arch. Phys. Med. Rehabil.* 80:837–841, 1999.
7. Jacobs, D., B. Ainsworth, T. Hartman, and A. Leon. A simultaneous evaluation of 10 commonly used physical activity questionnaires. *Med. Sci. Sports Exerc.* 25:81–91, 1993.
8. King, A. C., K. Baumann, P. O'Sullivan, C. Castro, and S. Wilcox. Optimizing health in older family caregivers: the Teaching Healthy Lifestyles to Caregivers Trial. *Ann. Behav. Med.* 22:42, 2000.
9. King, A. C., C. Castro, P. O'Sullivan, K. Baumann, and S. Wilcox. Promoting exercise adoption in at-risk older adults: a clinical trial aimed at family caregivers. *Ann. Behav. Med.* 21:S65, 1999.
10. King, A. C., L. A. Pruitt, W. Phillips, R. Oka, A. Rodenburg, and W. L. Haskell. Comparative effects of two physical activity programs on measured and perceived physical functioning and other health-related quality of life outcomes in older adults. *J. Gerontol. Med. Sci.* 55A:M74–M83, 2000.
11. King, A. C., J. Rejeski, and D. Buchner. Physical activity interventions targeting older adults: a critical review and recommendations. *Am. J. Prev. Med.* 15:316–332, 1998.
12. Kromhout, D., W. Saris, and C. Horst. Energy intake, energy expenditure, and smoking in relation to body fatness: the Zutphen Study. *Am. J. Clin. Nutr.* 47:668–674, 1988.
13. Marcus, B., M. Goldstein, A. Jette, et al. Training physicians to conduct physical activity counseling. *Prev. Med.* 26:382–388, 1997.
14. Martin, K., W. Rejeski, M. Miller, M. James, W. Ettinger, and S. Messier. Validation of the PASE in older adults with knee pain and physical disability. *Med. Sci. Sports Exerc.* 31:627–633, 1999.
15. Mchorney, C. A., J. E. Ware, J. Rachel Lu, and C. Sherbourne. The Mos 36-Item Short-Form Health Survey (SF-36). III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med. Care* 32:40–66, 1994.
16. Nunnally, J. *Psychometric Theory*. New York: McGraw-Hill; 1978, p. 245.
17. Paeratakul, S., B. Popkin, G. Keyou, L. Adair, and J. Stevens. Changes in diet and physical activity affect the body mass index of Chinese adults. *Int. J. Obes. Relat. Metab. Disord.* 22:424–431, 1998.
18. SAS Institute Inc. *SAS/Stat User's Guide*. Cary, NC: SAS Institute Inc., 1993.
19. Starling, R., D. Matthews, P. Ades, and E. Poehlman. Assessment of physical activity in older individuals: a doubly labeled water study. *J. Appl. Physiol.* 86:2090–2096, 1999.

20. STEWART, A. L., K. M. MILLS, A. C. KING, W. L. HASKELL, D. GILLIS, and P. RITTER. CHAMPS physical activity questionnaire for older adults: outcomes for interventions. *Med. Sci. Sports Exerc.* (in press).
21. STEWART, A. L., K. M. MILLS, P. SEPSIS, et al. Evaluation of CHAMPS, a physical activity promotion program for older adults. *Ann. Behav. Med.* 19:353–361, 1997.
22. STEWART, A. L., C. J. VERBONCOEUR, B. Y. MCLELLAN, et al. Physical activity outcomes of CHAMPS II: a physical activity promotion program. *J. Gerontol. Med. Sci.* (in press).
23. VOORRIPS, L., A. RAVELLI, P. DONGELMANS, P. DEURENBERG, and W. VAN STAVEREN. A physical activity questionnaire for the elderly. *Med. Sci. Sports Exerc.* 23:974–979, 1991.
24. WARE, J. E. *SF-36 Health Survey: Manual and Interpretation Guide*. Boston: The Health Institute, New England Medical Center, 1993.
25. WARE, J. E., and C. D. SHERBOURNE. The MOS 36-Item Short -Form Health Survey (SF-36). I. Conceptual framework and item selection. *Med. Care* 30:473–483, 1992.
26. WASHBURN, R. A., and J. FICKER. Physical Activity Scale for the Elderly (PASE): the relationship with activity measured by a portable accelerometer. *J. Sports Med. Phys. Fitness* 39:336–340, 1999.
27. WASHBURN, R.A., E. MCAULEY, J. KATULA, S. MIHALKO, and R. BOILEAU. The Physical Activity Scale for the Elderly (PASE): evidence for validity. *J. Clin. Epidemiol.* 52:643–651, 1999.
28. WASHBURN, R. A., K. SMITH, A. M. JETTE, and C. JANNEY. The Physical Activity Scale for the Elderly (PASE): development and evaluation. *J. Clin. Epidemiol.* 46:153–162, 1993.
29. YOUNG, D. R., L. APPEL, S. JEE, and E. MILLER III. The effects of aerobic exercise and T'ai Chi on blood pressure in older people: results of a randomized trial. *J. Am. Geriatr. Soc.* 47:277–284, 1999.