The Effects of a Community-Based Exercise Program on Function and Health in Older Adults: The EnhanceFitness Program

Basia Belza
Anne Shumway-Cook
Elizabeth A. Phelan
Barbara Williams
University of Washington, Seattle

Susan J. Snyder
Senior Services of Seattle/King County

James P. LoGerfo
University of Washington, Seattle

This study examined the effectiveness of participation in EnhanceFitness (EF) (formerly the Lifetime Fitness Program), an established community-based group exercise program for older adults. EF incorporated performance and health status measure testing in year 2000. Initial performance was compared to age- and gender-based norms to classify participants as within or at or above normal limits (WNL) or below (BNL). In 2,889 participants who participated in outcomes testing, improvements were observed at 4 and 8 months on performance tests for both subgroups. Participants' self-rating of health improved at 8 months. All participants improved on performance tests. Implementation of performance-based measures in community studies is possible. Challenges included selecting measures, staff training, collecting performance measures, and deciding on time points for data collection. Older adults can maintain and/or improve physical function through participation in EnhanceFitness.

Keywords: exercise; aging; older adults; physical performance; physical function

---

AUTHORS' NOTE: This research was funded by the Centers for Disease Control and Prevention, Prevention Research Center Program, with a grant to the University of Washington Health Promotion Research Center (U48/CCU009654). We sincerely thank the instructors and participants of EnhanceFitness and the staff of Senior Services Project Enhance including Meghan Thompson, Brenda Barkey, Chris Grekoff, and Kari Washburn.

The Journal of Applied Gerontology, Vol. 25 No. 4, August 2006 291-306
DOI: 10.1177/0733464806290934

Physical activity has been shown to benefit a number of chronic conditions prevalent among older adults (Batty & Lee, 2004; Chavannes, Vollenberg, van Schayck, & Wouters, 2002; Singh, 2004). In addition, physical activity and exercise improve balance and reduce risk for falls (Chang et al., 2004; Clemson et al., 2004; Gardner, Robertson, & Campbell, 2000; Gillespie et al., 2003) and is associated with lower risk for physical disability in later life (Fried & Guralnik, 1997). Among older adults, physical inactivity is a strong predictor of physical disability (Cress et al., 1999; Hirvensalo, Rantanen, & Heikkinen, 2000; Van Den Brink et al., 2005).

The surgeon general's report on physical activity and health recommends that all adults accumulate 30 minutes of moderate activity on all or most days of the week (U.S. Department of Health and Human Service [USDHHS], 1996). In addition to endurance exercises, the report also stresses the importance of strength training exercises twice a week. Despite these recommendations, only 21.4% of older adults (age 65 years and older) are meeting these recommendations (Centers for Disease Control and Prevention [CDC], 2002).

Structured exercise programs are known to improve aerobic capacity (Kostka, Drai, Berthouze, Lacour, & Bonnefoy, 2000), muscle strength (Fiatarone, 2002), and endurance (Ades et al., 1996). Physical performance measures are independently related to independence in instrumental activities of daily living (IADLs) in frail older adults (Coleman, Buchner, Cress, Chan, & de Lateur, 1996). Lower-extremity functional limitation, and low level of physical activity in general, has been associated with functional status decline in community-living elderly (Stuck et al., 1999). However, there is limited information about using physical performance and health status as outcome measures in community-based, group exercise programs in nonresearch settings. And there are also few studies that examine differences and similarities in outcomes across ethnic groups who are participating in exercise programs.

The current study was undertaken to examine the effectiveness of EnhanceFitness (EF) (formerly the Lifetime Fitness Program), a community-based group exercise program on improving physical performance in older adults. In 2000, the EF changed procedures to formally include periodic measurement of physical performance and health status in the ongoing program as part of an effort to monitor quality, provide evaluation data for administrators and funders, and to generate feedback for participants. The objectives of the current study were to use the routinely gathered data to (a) examine the effectiveness of 4 and 8 months' participation in a community-based group exercise program on physical performance, health status, and falls in White and ethnically diverse older adults and (b) summarize lessons learned about

implementing performance-based measurement in a widely disseminated, community-based program.

Methods

Participants

Participants included 2,889 older adults (mean age 75.5 years, *SD* ± 6.6) who enrolled in and participated in EF from April 2002 to September 2005.

Intervention

EF is a community-based group exercise program involving supervised classes that meet 3 times per week for 1 hour. Classes emphasize moderate-intensity aerobic conditioning, strength training, flexibility, and balance exercises. EF was initially studied in a randomized, controlled trial at a senior center in Bothell, Washington (Wallace et al., 1998). Participants for the current study were recruited from a total of 116 classes in nine states. The four states with the majority of facilities and thus the majority of participants were: 81 facilities in Washington (1,941 participants), 8 facilities in Texas (143 participants), 7 facilities in Maine (319 participants), and 7 facilities in Michigan (188 participants). The remaining five states in which participants resided were California; Washington, D.C.; Georgia; New York; and South Carolina.

Classes meet in various community locations including churches, senior centers, public housing, hospitals, YMCAs, ethnic community centers, and fitness centers. Class sizes vary but on average include about 14 participants per class. Instructors are certified as fitness trainers from an external professional organization (such as American College of Sports Medicine [ACSM] or American Council on Exercise [ACE]). Instructors receive an additional 12 hours of training related to special issues in older adults and EF procedures and protocols. All EF classes follow a standardized format: 5 minutes of warm-up, 20 to 25 minutes of moderate intensity aerobics (12 to 13 on the Borg [6 to 20] scale; Borg, 1998), 20 minutes of resistance strength training (with cuff weights), and 10 minutes of flexibility and balance training. The aerobic program content varies across sites and includes adaptations such as Mexican dance at Latino congregate meal sites.

In the Puget Sound area, classes are offered as a benefit to Group Health Cooperative (GHC) Medicare enrollees; however, classes are also open to

other older adults. GHC pays each site U.S. $2.25 per enrollee/per class (2005 rate), while other attendees pay their own fee. The percentage of GHC participants in this data set is 27%.

Measures

Functional performance measures. Since the original 1998 study, EF has used the Functional Fitness Test (Jones & Rikli, 2002) for performance measurement. This is a seven-item test measuring strength (Arm Curl, 30-Second Chair Stand), flexibility (Chair Sit-and-Reach, Back Scratch), endurance (Six-Minute Walk, Two-Minute Step Test), and functional mobility (Eight Foot Up-and-Go). The Functional Fitness Test includes age and gender norms for adults age 60 to 94 years. Test validity and test-retest reliability (intraclass reliability estimates ranged from .81 to .95) have been established for the measures in older adults (Rikli & Jones, 1999) and indicate that the tests are psychometrically strong.

To test the feasibility of integrating the Functional Fitness Test into EF, pilot testing was carried out in three EF classes on two occasions, 4 months apart. Instructors and participants reported dissatisfaction with the complete Functional Fitness Test. Both reported that performance of all seven measures was too time-consuming to be included in routine class sessions. Instructors reported they could not complete all seven tests with all class participants in a single class period. Participants were unhappy that testing took an entire session, leaving no time for exercise. Several instructors asked for participant help in conducting the tests but raised concerns regarding the reliability of testing done by untrained participants. Instructors reported difficulty in (a) completing the 6-min walk because of lack of space and (b) performing and scoring the 2-min step test. Based on this feedback, four of the seven tests were eliminated: Back Scratch, Chair Sit-and-Reach, Six-Minute Walk, and Two-Minute Step. Three tests were retained for routine data collection: (a) Arm Curl, a measure of upper-extremity strength that assesses the number of times in which a weight can be lifted in a 30-sec time period, with women using a 5-pound and men using an 8-pound weight; (b) 30-Second Chair Stand, a measure of lower-extremity strength, in which participants are asked to stand up and sit down without using their arms from a 17-in high chair as many times as they can in a 30-sec period; and (c) Eight Foot Up-and-Go, a measure of balance and mobility in which participants are asked to stand up from a 17-in chair, walk as quickly as they safely can for a distance of 8 feet, turn, walk back, and sit down in the chair. All results are entered on optically scannable forms.

Health status. The Short Form-12 (SF-12) Health Survey summary scores for the physical component summary (PCS-12) and mental component summary (MCS-12) were used as measures of perceived health status (Ware, Kosinski, & Keller, 1996). The SF-12 was developed to be a much shorter but valid alternative to the Short Form-36 (SF-36; Ware et al., 1996) for use in large surveys and also for longitudinal studies of health outcomes. The SF-12 PCS and SF-12 MCS outcome scores are interchangeable with those from the SF-36 in general and specific populations (Ware et al., 1996; for scoring, see Ware, Kosinski, & Keller, 1998). Higher scores represent better health. The first item on the SF-12 asks participants to rate their health, in general, on a 5-point scale from *poor* to *excellent*.

Procedures

EF participants were older adults who resided in the community and at the time of entry into the study were already participating in a congregate meal program, attending a senior center, or had selected EF as part of a Health Action Plan. They were invited to join an EF class, and if they agreed they completed a Health Enrollment Form (health history) and Participant Information Form (demographics). The EF instructor sent a letter to each participant's primary care physician notifying them of their interest in the class. Participants were encouraged but not required to attend class 3 times a week. Attendance was taken every session by the EF instructor.

Participants also completed three performance measures at entry into the study and then 1 year later. EF instructors who had been trained in the testing procedures conducted performance-based testing within a normally scheduled EF class every 4 months (February, June, and October). Every 4 months a day was set aside for testing. Participants present on that day were tested; those absent for any reason are either offered testing on another day soon thereafter or not tested until a subsequent test date.

Participation in the testing was voluntary. Consent to allow data to be shared for quality improvement and research was obtained from class participants at the time of testing. Of participants, 85% agreed to have data used for these purposes. Data collection for the current study occurred during the period of April 2002 through September 2005. The initial data reported for participants in the current study is not synonymous with enrollment in an exercise class, but rather by the initiation of testing procedures within EF. Follow-up data for the initial 2,889 participants was gathered at 4 months for 1,258 of these participants and at 8 months for 880 participants. The Institutional Review Board of the University of Washington approved all data use, consent, and sharing procedures.

Statistical Analysis

Two-tailed *t* tests or chi-square tests were used to compare participants with and without follow-up data. Two-tailed, paired *t* tests and McNemar's chi-square tests were used to compare initial with follow-up measures. A participant's functional performance measures were included in the analyses if the performance outcome fell within the 99.5 percentile (Arm Curls: 0 to 50 repetitions, 30-Second Chair Stands: 0 to 30, Eight-Foot Up-and-Go: 3 to 42 seconds). The presence of follow-up data was defined as having at least one (of the three) performance measures or the SF-12 measure subsequent to initial testing.

Results

Initial Demographics, Health Characteristics, and Performance Status

Initial results were available for 2,889 participants (Table 1). The majority of participants were women with an average age of 75.5 years. The sample included Whites (n = 1,844), African Americans (n = 219), Asians (n = 209), Hispanics (n = 117), and Others (n = 418). A number of participants elected not to report their ethnic group affiliation (n = 82). More than three fourths of all participants rated their health as good or better. Of participants, 18% reported falling in the past 4 months. Participants attended, on average, 1.8 EF classes/week.

The number of participants with data at 4 and 6 months varied for each performance test. Of the participants, 44% did not have follow-up testing at 4 months, and another 35% did not have follow-up at 8 months. Among sites that recorded a reason for lack of testing, the most common reasons (n = 1,737) for lack of follow-up were dropped out of EF (551 or 31.7%), temporary absence (not attending class on the day of the testing; 395 or 22.7%), vacation (196 or 11.%), illness (173 or 9.9%), irregular attendance (57 or 3.3%), and other (365 or 20.6%; e.g., moved, caregiver issues, not specified). Initial scores for those with follow-up data were compared to those without follow-up data to determine if participants with follow-up data were representative of the group (i.e., all those attending EF) as a whole. There were no significant differences in initial scores between those with and without follow-up data on the demographics variables, self-report of falls, or initial scores on the three performance measures. Those without follow-up data at 4 months differed on MCS-12 and PCS-12 scores, scoring significantly worse on each of these measures, compared to those with follow-up data.

Table 1. Initial Participant Characteristics ($M \pm SD$ or percentage) by 4-Month Follow-Up Status

Characteristic	*Total*	*Without 4-Month Follow-Up*	*With 4-Month Follow-Up*	p *value*[a]
N	2,889	1,631	1,258	—
Gender				
Female (%)	80.5	80.3	80.8	.72
Ethnicity[b]				
White (%)	1,844	1,006 (55)	838 (45)	< .001
Black (%)	219	116 (53)	103 (47)	
Asian (%)	209	107 (51)	102 (49)	
Hispanic (%)	117	51 (44)	66 (56)	
Unknown (%)	418	293 (70)	125 (30)	
Other (%)	82	58 (71)	24 (29)	
Age (years)	75.5 ± 6.67	5.5 ± 6.5	75.5 ± 6.7	.38
Age (% ≥ 75)	52.7	53.5	51.7	.32
Self-rated health (% good/very good/ excellent)	84.8	84.2	85.6	.30
Any falls in past 4 months (%)	18.0	18.1	17.9	.92
BNL[c] subgroup (%)				
Eight-Foot Up-and-Go	41.6	42.4	40.6.	.37
30-Second Chair Stand	21.1	21.1	21.3	.89
Arm Curl	11.8	12.2	11.2	.42
Eight-Foot Up-and-Go (seconds)	8.1 ± 4.1	8.2 ± 4.3	8.0 ± 3.7	.15
30-Second Chair Stand (# in 30 seconds)	12.6 ± 4.6	12.6 ± 4.6	12.7 ± 4.5	.62
Arm Curl (# in 30 seconds)	17.0 ± 6.0	17.0 ± 6.2	16.9 ± 5.7	.72
SF-12, physical[d]	44.8 ± 10.3	44.2 ± 10.5	45.5 ± 10.0	.011
SF-12, mental[d]	54.1 ± 8.2	53.5 ± 8.6	54.7 ±7.7	.005

a. Significance was calculated using paired *t* tests for continuous variables and McNemar's chi-square for dichotomous variables.
b. Ethnicity: $N = 2,426$ (1,308 no follow-up; 1,118 with follow-up)
c. BNL: below normal limits
d. SF-12 (Medical Outcomes Study Short-Form-12): $N = 1,688$ (899 no follow-up, 789 with follow-up)

Initial performance on each of the three tests was compared to age- and gender-based published cut-points and used to classify participants as within (at or above) normal limits (WNL) or below normal limits (BNL) for age and gender (Jones & Rikli, 2002).

Table 2. Effects of Participation in EnhanceFitness on Performance Measures, Health Status, and Falls at 4 Months (*M* ± *SD* or %)

Test	N	*Initial*	*Month 4*	Δ	p *Value*[a]
Eight-Foot Up-and-Go (seconds)					
Total	1,128	7.5 ± 2.4	7.1 ± 2.2	-0.4	< .001
% BNL[b]	1,128	39.8	32.8	7.0	< .001
Initial BNL	449	9.6 ± 2.4	8.4 ± 2.4	–1.2	< .001
Initial WNL[c]	679	6.2 ± 1.2	6.3 ± 1.5	0.1	.024
Chair Stand (# stands in 30 seconds)					
Total	1,195	12.7 ± 4.5	14.3 ± 4.8	1.6	< .001
% BNL	1,195	21.0	11.7	9.3	< .001
Initial BNL	251	7.3 ± 2.5	10.4 ± 4.0	3.1	< .001
Initial WNL	944	14.1 ± 3.8	15.3 ± 4.4	1.2	< .001
Arm Curl (repetitions in 30 seconds)					
Total	1,136	16.9 ±5.6	19.4 ± 6.2	2.5	< .001
% BNL	1,136	11.1	4.6	6.5	< .001
Initial BNL	126	9.1 ± 3.0	15.8 ± 6.4	6.7	< .001
Initial WNL	1,010	17.9 ± 5.1	19.9 ± 6.0	1.9	< .001
Self-rated health					
% Good/very good/ excellent	1,097	85.8	87.2	1.4	.16
SF-12 (Medical Outcomes Study Short-Form-12), physical mean	553	45.4 ± 10.1	45.8 ± 10.0	–.28	.26
SF-12, mental mean	553	54.8 ± 7.7	54.7 ± 7.2	.13	.70
% Any falls in past 4 months	599	18.2	19.4	1.2	.61

a. Significance was calculated using paired *t* tests for continuous variables and McNemar's chi-square for dichotomous variables.
b. BNL: below normal limits.
c. WNL: within (at or above) normal limits.

Effect of EF on Measured and Perceived Function

Tables 2 and 3 compare initial and 4- and 8-month follow-up scores on performance-based measures of function by subgroup (BNL and WNL) and the results of the self-rated health, SF-12, and falls items.

Table 3. Effects of Participation in EnhanceFitness on Performance Measures, Health Status, and Falls at 8 Months (*M* + *SD* or %)

Test	N	*Initial*	*Month 8*	Δ	p *Value*[a]
Eight-Foot Up-and-Go (seconds)					
Total	830	7.4 ± 2.3	7.0 ± 2.3	–0.4	< .001
% BNL[b]	830	37.8	29.2	8.6	< .001
Initial BNL	314	9.5 ± 2.3	8.3 ± 2.6	–1.2	< .001
Initial WNL[c]	516	6.2 ± 1.1	6.2 ± 1.6	0.0	.36
30-Second Chair Stand (# stands in 30 seconds)					
Total	861	12.8 ± 4.6	14.7 ± 5.0	1.9	< .001
% BNL	861	20.8	11.4	9.4	< .001
Initial BNL	179	7.5 ± 2.3	11.1 ± 4.4	3.6	< .001
Initial WNL	682	14.2 ± 4.0	15.6 ± 4.7	1.4	< .001
Arm Curl (# repetitions in 30 seconds)					
Total	814	16.9 ± 5.5	20.0 ± 6.6	3.1	< .001
% BNL	814	11.7	3.3	8.4	< .001
Initial BNL	95	9.2 ± 2.7	18.0 ± 7.3	8.8	< .001
Initial WNL	719	18.0 ± 5.0	20.3 ± 6.4	2.3	< .001
Self-rated health % Good/very good/ excellent	800	86.6	90.5	3.9	.001
SF-12 (Medical Outcomes Study Short-Form-12), physical mean	377	45.8 ± 9.8	46.0 ± 9.8	–.22	.56
SF-12, mental mean	377	54.9 ± 7.4	54.7 ± 7.8	0.26	.53
% Any falls in past 4 months	363	17.9	16.8	1.1	.74

a. Significance was calculated using paired *t* tests for continuous variables and McNemar's chi-square for dichotomous variables.
b. BNL: below normal limits.
c. WNL: within (at or above) normal limits.

With two exceptions, there were significant improvements in the three performance measures at the 4- and 8-month testing in the BNL and WNL subgroups. At 4 months, the BNL group improved their Eight-Foot Up-and-Go time by decreasing, on average, from 9.6 (*SD* = 2.4) to 8.4 seconds (*SD* = 2.4,

$p > .001$). The number of chair stands in 30 seconds, on average, improved from 7.3 ($SD = 2.5$) to 10.4 ($SD = 4.0$, $p > .001$). The repetitions during the Arm Curl test improved, on average, from 9.1 ($SD = 3.0$) to 15.8 times ($SD = 6.4$, $p > .001$). The WNL group slightly worsened their Eight-Foot Up-and-Go time by increasing, on average, from 6.2 ($SD = 1.2$) to 6.3 seconds ($SD = 1.5$, $p = .024$). The number of chair stands in 30 seconds, on average, improved from 14.1 ($SD = 3.8$) to 15.3 ($SD = 4.4$; $p > .001$). The repetitions during the Arm Curl test improved, on average, from 17.9 ($SD = 5.1$) to 19.9 times ($SD = 6.0$, $p > .001$).

At 8 months, the BNL group improved their Eight-Foot Up-and-Go time by decreasing, on average, from 9.5 ($SD = 2.3$) to 8.3 sec ($SD = 2.6$, $p > .001$). The number of chair stands in 30 seconds, on average, improved from 7.5 ($SD = 2.3$) to 11.1 ($SD = 4.4$, $p > .001$). The repetitions during the Arm Curl test improved, on average, from 9.2 ($SD = 2.7$) to 18.0 times ($SD = 7.3$, $p > .001$). The WNL group did not improve on their Eight-Foot Up-and-Go test (from 6.2 [$SD = 1.1$] to 6.2 sec [$SD = 1.6$, $p = .36$]). The number of chair stands in 30 seconds, on average, improved from 14.2 ($SD = 4.0$) to 15.6 ($SD = 4.7$, $p > .001$). The repetitions during the Arm Curl test improved, on average, from 18.0 ($SD = 5.0$) to 20.3 times ($SD = 6.4$, $p > .001$).

Although at 4 months participants did not report a significant change in their self-rated health (85.8 to 87.2, $p = .16$), at 8 months participants did report a significant improvement on self-rated health (86.6 to 90.5, $p > .001$). There were no significant differences in physical or mental subscale scores from the initial to the 4- and 8-month testing.

Performance Results by Ethnicity

Analysis of data for White participants compared to participants of color showed similar effects with improvements on all performance measures, with one exception. The WNL White participant subgroup stayed the same on the Eight-Foot Up-and-Go test from initial to the 4-month testing ($p = .85$). Only the participants in the color subgroup (not-White subgroup) reported a significant improvement in self-rated health (data not shown, $p = .049$).

The biggest baseline difference across ethnic groups (African American, Asian, White, and Hispanic) was that the Asian subgroup had many more seniors that were within or above normal limits in the Chair Stand and Arm Curl. For example, only 6% of the Asians were BNL in the Chair Stand, compared to 22% BNL for other ethnic groups.

Hispanic seniors were most often BNL in the Up and Go and Arm Curl compared to the other ethnic groups. For example, 19% of the Hispanic seniors were BNL in the Arm Curl, compared to 11.5% for other ethnic

groups. Despite this finding, the Hispanic group is slightly, although significantly younger than the other ethnic groups (73.3 years vs. 75.6 years).

In summary, three main findings are evident when examining data across ethnic groups. First, Whites and Hispanics show significant improvement in all three performance measures (Chair Stand, Arm Curl, and Up and Go) from initial test to 4-month follow-up. Second, Asians display higher functioning at initial testing with the outcome measures maintained at 4-month follow-up. And third, an area for improvement could be with the African Americans who on the Up and Go at baseline are 52% BNL that only improves slightly but not significantly to 45% at 4-month follow-up.

Discussion

The current study was undertaken to determine whether participation in a community-based group exercise program would improve physical performance in community living older adults. Results suggest that EF has beneficial effects on physical function, as measured by physical performance tests assessing upper- and lower-extremity function and functional gait. EF improved physical performance in older adults who performed below age- and gender-based norms and in those who performed at or above normal limits for age and gender at initial testing and for White participants and participants of color.

The current effectiveness study demonstrates, in quantitative terms, the significant gains in physical performance that may be achieved through 4 to 8 months of EF participation. These findings extend the results from the initial efficacy trial of EF, which reported positive effects of the EF on health status (SF-36; Wallace et al., 1998). Wallace's sample and the current sample are similar in that our participants were on average age 75.5 years and were 80% female; participants in Wallace's study were on average age 72 years and were 73% female.

Numerous community-based programs that promote physical activity among older adults have been designed and evaluated (King, Rejeski, & Buchner, 1998). EF is one of a few group- and community-based programs, specifically designed for older adults, that meets many of the criteria that have been described as characteristics of a successful physical activity program: (a) designed within a social ecological framework, (b) tailored to the specific age-related challenges and concerns of the target population, (c) demonstrated to be effective in an evaluation that was rigorously designed, (d) has high reach and efficacy, (e) builds on existing community resources, and (f) uses participatory research methods (Prohaska et al., in press;

Smedley & Syme, 2001). Other exemplary community-based programs include Community Health Advice by Telephone (CHAT; King, Haskell, Taylor, Kraemer, & DeBusk, 1991), Community Healthy Activities Model Program for Seniors (CHAMPS II; Stewart et al., 2001), Active Living Every Day and Active Choices (Dunn et al., 1998; Wilcox et al., 2006), and Strong for Life (Jette et al., 1999). However, EF is one of a limited number of programs being widely disseminated to the public (Jette et al., 1999; Nelson & Wernick, 1997). In 2006, EF was added to the CDC Arthritis Program's list of recommended physical activity interventions appropriate for people with arthritis (www.cdc.gov/arthritis/index.htm).

Of interest is the finding that scores on PCS-12 and MCS-12 were unchanged despite improvements on performance-based testing. One possible explanation for this finding is the SF-12 scores were high to begin with among these participants (ceiling effect; Pettit et al., 2001). Ware published national means for the PCS-12 of 43.4 (our initial sample mean was 44.8) and for the MCS-12 of 55.2 (our sample mean was 54.1) for individuals age 65 years and older (Ware et al., 1996). Alternatively, Simon, Revicki, Grothaus, and Vonkorff (1998) found that SF-36 scores may not accurately reflect changes in health over time in longitudinal studies. In addition, there may be a lag between changes in physical performance and change in ratings of perceived physical and mental health. This finding suggests that researchers should exercise caution in relying solely on measures such as the SF-12 if they wish to capture the full range of effects of physical activity interventions, especially in evaluations of community-living older adults.

Our joint experience as a community agency–academic partnership to develop sustainable community programs has shown that collecting performance-based measures as a component of any broad-based dissemination initiative is possible but presents challenges for the community agency(s), researchers, instructors, and participants. Our lessons learned include the following: (a) Decisions need to be made as to why the data are being collected and who will have access to the data (participants, partnering agencies, instructors, and research partners); (b) When selecting the appropriate measures, consideration needs to be made for the rationale for data collection, making an explicit link to the anticipated outcome(s), cost, and storage of equipment, staff time in learning how to conduct the measures, correctly enter data and actually conducting the testing, and setting (is appropriate space available at the required time to conduct the measures); (c) Collecting performance-based measures requires attention to ongoing assessment of interrater reliability; (d) Participant cooperation is critical; (e) Decisions need to be made as to the optimal time points for data collection based on practical and scientific perspectives; (f) When the initial training is conducted, decisions need to be made as to who will pay

for refresher training programs, interrater reliability checks, ongoing data collection, and costs to maintain a database; (g) Variations in exercises need to be part of the standardized program to accommodate various functional abilities such as participants with differences in mobility; (h) Users of the data should expect to encounter significant gaps in data for any given participant and therefore should decide how valuable more complete data capture will occur at their individual site. If crucial, much greater attention to individual tracking and testing may be required than can be expected under otherwise usual operating conditions; (i) Finally, diverse community agencies across the nation may themselves have research partners or funders requiring other data than those in common use. Therefore, data collection systems should provide for some flexibility, such as addition of an arm curl or Eight-Foot Up-and-Go should they be needed and if reliable testing can be done in a given setting.

There were several study limitations. The results are susceptible to biases associated with observational studies: bias related to loss to follow-up data (testing nonparticipation) and selection bias (recruitment was not population based). In regard to loss of follow-up data, those who did not complete follow-up testing reported worse physical and mental health (SF-12) initially and thus would have had more room to improve; this would have tended to result in larger improvements between initial and follow-up testing if they had been included in the follow-up testing. Another limitation is the uncertainty as to the generalizability of the findings to those older adults who are less functional, as the current study involved primarily higher functioning older adults living in the community. Blind assessments were not conducted, and as such there might be concern about internal validity. The use of an instructor to administer the performance tests may have a biasing effect. Of note, the full benefits of the program were not measured as EF has an aerobic component yet an aerobic measure was not included.

These limitations notwithstanding, the current study has several strengths, which include its demonstration of the benefits on physical performance of a community-based group exercise intervention that involves all the components (endurance, strength training, balance, and flexibility) that are needed to preserve physical function and independence with advancing age. In conclusion, the current study provides new evidence about the benefits of participation in EF, a group exercise program especially designed for older adults and currently in the process of being widely disseminated. It is possible to develop and broadly disseminate and incorporate standardized performance measures for ongoing program evaluation and effectiveness. EF is a dynamic program and evolves as evidence becomes available and participants' needs change.

References

Ades, P. A., Waldmann, M. L., Meyer, W. L., Brown, K. A., Poehlman, E. T., Pendlebury, W. W., et al. (1996). Skeletal muscle and cardiovascular adaptations to exercise conditioning in older coronary patients. *Circulation*, *94*(3), 323-330.

Batty, G. D., & Lee, I. M. (2004). Physical activity and coronary heart disease. *British Medical Journal*, *328*(7448), 1089-1090.

Borg, G. A. (1998). *Borg's rating of perceived exertion and pain scales*. Champaign, IL: Human Kinetics.

Centers for Disease Control and Prevention. (2002). *National Center for Health Statistics, National Health Interview Survey*. Retrieved April 24, 2006, from www.cdc.gov/nchs/

Chang, J. T., Morton, S. C., Rubenstein, L. Z., Mojica, W. A., Maglione, M. Suttorp, M. J., et al. (2004). Interventions for the prevention of falls in older adults: Systematic review and meta-analysis of randomized clinical trials. *British Medical Journal*, *328*, 680-683.

Chavannes, N., Vollenberg, J. J., van Schayck, C. P., & Wouters, E. F. (2002). Effects of physical activity in mild to moderate COPD: A systematic review. *British Journal of General Practice*, *52*(480), 574-578.

Clemson, L., Cumming, R. G., Kendig, H., Swann, M., Heard, R., & Taylor, K. (2004). The effectiveness of a community-based program for reducing the incidence of falls in the elderly: A randomized trial. *Journal of the American Geriatrics Society*, *52*, 1487-1494.

Coleman, E. A., Buchner, D. M., Cress, M. E., Chan, B. K., & de Lateur, B. J. (1996). The relationship of joint symptoms with exercise performance in older adults. *Journal of the American Geriatric Society*, *44*(1), 14-21.

Cress, M. E., Buchner, D. M., Questad, K. A., Esselman, P. C., de Lateur, B. J., & Schwartz, R. S. (1999). Exercise: Effects on physical functional performance in independent older adults. *Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *54*(5), M242-248.

Dunn, A. L., Garcia, M. E., Marcus, B. H., Kampert, J. B., Kohl, H. W., & Blair, S. N. (1998). Six-month physical activity and fitness changes in Project Active, a randomized trial. *Medicine and Science in Sports and Exercise*, *30*(7), 1076-1083.

Fiatarone, M. A. (2002). Exercise in the oldest old: Some new insights and unanswered questions. *Journal of the American Geriatric Society*, *50*, 2089-2091.

Fried, L. P., & Guralnik, J. M. (1997). Disability in older adults: Evidence regarding significance, etiology, and risk. *Journal of the American Geriatric Society*, *45*(1), 92-100.

Gardner, M. M., Robertson, M. C., & Campbell, A. J. (2000). Exercise in preventing falls and fall related injuries in older people: A review of randomised controlled trials. *British Journal of Sports Medicine*, *34*(1), 7-17.

Gillespie, L. D., Gillespie, W. J., Robertson, M. C., Lamb, S. E., Cumming, R. G., & Rowe, B. H. (2003). Interventions for preventing falls in elderly people (CD000340). Retrieved from Cochrane Database Systematic Review 4 database.

Hirvensalo, M., Rantanen, T., & Heikkinen, E. (2000). Mobility difficulties and physical activity as predictors of mortality and loss of independence in the community-living older population. *Journal of the American Geriatric Society*, *48*(5), 493-498.

Jette, A. M., Lachman, M., Giorgetti, M. M., Assmann, S. F., Harris, B. A., Levenson, C., et al. (1999). Exercise—It's never too late: The Strong-For-Life Program. *American Journal of Public Health*, *89*(1), 66-72.

Jones, J., & Rikli, E. (2002, March/April). Measuring functional fitness of older adults. *Journal of Active Aging*, *1*, 24-30.

King, A. C., Haskell, W. L., Taylor, C. B., Kraemer, H. C., & DeBusk, R. F. (1991). Group- vs home-based exercise training in healthy older men and women: A community-based clinical trial. *Journal of the American Medical Association*, *266*(11), 1535-1542.

King, A. C., Rejeski, W. J., & Buchner, D. M. (1998). Physical activity interventions targeting older adults. A critical review and recommendations. *American Journal of Preventive Medicine*, *15*(4), 316-333.

Kostka, T., Drai, J., Berthouze, S. E., Lacour, J. R., & Bonnefoy, M. (2000). Physical activity, aerobic capacity and selected markers of oxidative stress and the anti-oxidant defense system in healthy active elderly men. *Clinical Physiology*, *20*(3), 185-190.

Nelson, M. E., & Wernick, S. (1997). *Strong women stay young*. New York: Bantam Books.

Pettit, T., Livingston, G., Manela, M., Kitchen, G., Katona, C., & Bowling, A. (2001). Validation and normative data of health status measures in older people: The Islington Study. *International Journal of Geriatric Psychiatry*, *16*(11), 1061-1070.

Prohaska, T., Belansky, E., Belza, B., Buchner, D., Marshal, V., McTigue, K., et al. (in press). Physical activity, public health, and aging: Critical issues and research priorities. *Journal of Gerontology: Social Sciences*.

Rikli, R., & Jones, J. (1999). Development and validation of a functional fitness test for community-residing older adults. *Journal of Aging and Physical Activity*, *7*, 127-159.

Simon, G. E., Revicki, D. A., Grothaus, L., & Vonkorff, M. (1998). SF-36 summary scores: Are physical and mental health truly distinct? *Medical Care*, *36*(4), 567-572.

Singh, M. A. (2004). Physical activity and bone health. *Australian Family Physician*, *33*(3), 125.

Smedley, B. D., & Syme, S. L. (2001). Promoting health: Intervention strategies from social and behavioral research. *American Journal of Health Promotion*, *15*(3), 149-166.

Stewart, A. L., Verboncoeur, C. J., McLellan, B. Y., Gillis, D. E., Rush, S., Mills, K. M., et al. (2001). Physical activity outcomes of CHAMPS II: A physical activity promotion program for older adults. *Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *56*(8), M465-470.

Stuck, A., Walthert, J., Nikolaus, T., Bula, C. J., Hohmann, C., & Beck, J. C. (1999). Risk factors for functional status decline in community living elderly people. *Social Science and Medicine*, *48*, 445-469.

U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion. (1996). *Physical activity and health: A report of the surgeon general*. Atlanta, GA: Author.

Van Den Brink, C. L., Picavet, H., Van Den Bos, G. A., Giampaoli, S., Nissinen, A., & Kromhout, D. (2005, March 18). Duration and intensity of physical activity and disability among European elderly men. *Disability and Rehabilitation*, *27*(6), 341-347.

Wallace, J. I., Buchner, D. M., Grothaus, L., Leveille, S., Tyll, L., LaCroix, A. Z., et al. (1998). Implementation and effectiveness of a community-based health promotion program for older adults. *Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *53A*(4), M301-306.

Ware, J., Jr., Kosinski, M., & Keller, S. D. (1996). A 12-item short-form health survey: Construction of scales and preliminary tests of reliability and validity. *Medical Care*, *34*(3), 220-233. Retrieved April 24, 2006, from www.sf-36.org/tools/SF36.shtml

Ware, J. E., Kosinski, M., & Keller, S. D. (1998, September). *How to score the SF-12 Physical and Mental Health Summary Scales* (3rd ed.). Lincoln, RI, and Boston: Quality Metric and Health Assessment Lab.

Wilcox, S., Dowda, M., Griffin, S., Rheaume, C., Ory, M., Leviton, L., et al. (2006). Results of the first year of Active for Life: Translation of 2 evidence-based physical activity programs for older adults into community settings. *American Journal of Public Health*. Prepublished May 30, 2006, DOI: 10.2105/AJPH.2005.074690

Article accepted April 24, 2006

Basia Belza, PhD, RN, is an associate professor in biobehavioral nursing and health systems and the Aljoya Endowed Professor in Aging at the University of Washington. Her research is focused on translating and disseminating evidence-based physical activity programs to ethnically diverse older adults.

Anne Shumway-Cook, PhD, PT, is an associate professor in rehabilitation medicine at the University of Washington. Her research is focused on understanding the neural basis of balance and mobility problems in neurologic and geriatric populations, and the clinical application of this research to improving function and participation in these populations.

Elizabeth A. Phelan, MD, MS, is an assistant professor in medicine/gerontology and geriatric medicine at the University of Washington. Her research is focused on improving quality of primary care of older adults; linking health care systems and communities to promote health, and preventing disability among community-dwelling older adults.

Barbara Williams, PhD, is a senior computer specialist, Health Promotion Research Center, at the University of Washington. Her interest and expertise is in the area of data analysis associated with measuring health outcomes in older adults. More recently, this analysis has tested the hypothesis that healthy behaviors, such as exercise and nutrition, can decrease rates of disease and disability.

Susan J. Snyder, MS, is the current director of Project Enhance of Senior Services of Seattle/King County. Her interests are in the areas of building the evidence base to determine effectiveness of EnhanceFitness in varied settings and populations, and determining how to effectively and efficiently disseminate EnhanceFitness (while maintaining fidelity).

James P. LoGerfo, MD, MPH, is a professor of medicine and health services and the director of UW Health Promotion Research Center at the University of Washington. His research focuses on community-based approaches to promoting healthy aging.