The Journal of Nutrition, Health & Aging©
Volume 13, Number 6, 2009

WHAT IS A MEANINGFUL CHANGE IN PHYSICAL PERFORMANCE? FINDINGS FROM A CLINICAL TRIAL IN OLDER ADULTS (THE LIFE-P STUDY)

S. KWON[1,2,3], S. PERERA[4,5,6], M. PAHOR[3,7], J.A. KATULA[8,9], A.C. KING[10], E.J. GROESSL[11,12], S.A. STUDENSKI[4,6,13]

1. College of Pharmacy, University of Florida, Gainesville, FL; 2. Rehabilitation Outcomes Research Center NF/SG VHA, Gainesville, FL; 3. Institute on Aging at University of Florida. Gainesville, FL; 4. Division of Geriatric Medicine University of Pittsburgh, Pittsburgh, PA; 5. Department of Biostatistics, University of Pittsburgh, Pittsburgh, PA; 6. Institute on Aging at University of Pittsburgh, Pittsburgh, PA; 7. Department of Aging and Geriatric Research, College of Medicine, University of Florida, Gainesville, FL; 8. Wake Forest University, Winston-Salem, NC; 9. Institute on Aging at Wake Forest University; 10. Stanford University School of Medicine, Palo Alto, CA; 11. VA San Diego Health Care System, San Diego, CA; 12. University of California San Diego, San Diego, CA; 13. Geriatric Research Education and Clinical Center, Department of Veterans Affairs Hospital, Pittsburgh, PA. Corresponding author: Sooyeon Kwon, R.Ph., Ph.D. Dr Kwon was at University of Florida during the manuscript development. Currently Dr. Kwon is affiliated to South Carolina College of Pharmacy – MUSC campus, Charleston, SC. Department of Clinical Pharmacy and Outcome Sciences, South Carolina College of Pharmacy – MUSC campus, 280 Calhoun Street, P.O. Box MSC140, Charleston, SC 29425, Office Tel: (843) 792-6076, kwons@musc.edu

Abstract: *Objective:* Performance measures provide important information, but the meaning of change in these measures is not well known. The purpose of this research is to 1) examine the effect of treatment assignment on the relationship between self-report and performance; 2) to estimate the magnitude of meaningful change in 400-meter walk time (400MWT), 4-meter gait speed (4MGS), and Short Physical Performance Battery (SPPB) and 3) to evaluate the effect of direction of change on estimates of magnitude. *Design:* This is a secondary analysis of data from the LIFE-P study, a single blinded randomized clinical trial. Using change over one year, we applied distribution-based and anchor-based methods for self-reported mobility to estimate minimally important and substantial change in 400MWT, 4MGS and SPPB. *Setting:* Four university-based clinical research sites. *Participants:* Sedentary adults aged 70-89 whose SPPB scores were less than 10 and who were able to complete a 400MW at baseline (n=424). *Interventions:* A structured exercise program versus health education. *Measurements:* 400MWT, 4MGS, SPPB. *Results:* Relationships between self-report and performance measures were consistent between treatment arms. Minimally significant change estimates were 400MWT: 20-30 seconds, 4MGS: 0.03-0.05m/s and SPPB: 0.3 – 0.8 points. Substantial changes were 400MWT: 50-60 seconds, 4MGS: 0.08m/s, SPPB: 0.4 – 1.5 points. Magnitudes of change for improvement and decline were not significantly different. *Conclusions:* The magnitude of clinically important change in physical performance measures is reasonably consistent using several analytic techniques and appears to be achievable in clinical trials of exercise. Due to limited power, the effect of direction of change on estimates of magnitude remains uncertain.

Key words: Aging, physical performance, meaningful change.

Introduction

Given the power of physical performance measures to reflect concurrent and future health, functioning and health care utilization among older adults, such measures are increasingly incorporated into many types of studies of aging (1-4). As novel interventions are developed to improve health in aging, physical performance measures have the potential to serve as primary indicators of benefit in future clinical trials. However, to be accepted as outcome measures, the clinical meaning of change in these measures must be understood. Previous reports have begun to estimate the magnitude of meaningful change in order to understand the meaning of change over time in performance measures (5).

However, important gaps in knowledge about meaningful change in performance remain. First, many interventions that are important for the health of older adults, such as exercise, are not amenable to participant blinding (6-9), so that knowledge of treatment assignment might influence the relationship between self-reported measures of change and physical performance estimates of change. Second, prior estimates have been based on either observational studies or small clinical trials, and no estimate has been provided for meaningful change in the 400-meter walk. Data from larger clinical trials are needed to provide robust and more precise estimates. Third, the direction of change might influence the magnitude of what is meaningful. For example, an important improvement might be larger or smaller than an important decline. While standard distribution-based methods assume symmetry of response, anchor-based methods of estimating meaningful change can be used to compare magnitudes in each direction (10, 11).

Due to its size, use of performance measures and wide range of change effects, the Lifestyle Interventions and Independence for Elders Pilot Study (LIFE-P) provides a unique opportunity to address these important gaps in knowledge and help prepare for future clinical trials that use performance measures as endpoints. The purpose of this analysis is to 1) examine the consistency of relationships between self-reported and performance measures between intervention groups, 2) estimate the magnitude of meaningful change in 400-meter walk time, gait speed, and Short Physical Performance Battery (SPPB), and 3) evaluate the effect of direction of change on estimates of magnitude.

Received October 8, 2008
Accepted for publication October 29, 2008

The Journal of Nutrition, Health & Aging©
Volume 13, Number 6, 2009

Methods

Data Source

We used data from the Lifestyle Interventions and Independence for Elders Pilot (LIFE-P) study. LIFE-P was a multi-center, single-blind, randomized trial of a physical activity intervention versus health education in 424 sedentary older adults aged 70 to 89 years. Participants were required to demonstrate increased risk of future mobility disability by having a Short Physical Performance Battery (SPPB) score of 9 or less (12), but also to retain adequate mobility at baseline as demonstrated by capacity to complete a 400-meter walk in 15 minutes or less (13). The study design, protocol, inclusion and exclusion criteria, the contents of physical activity and health education interventions, and baseline characteristics of the subjects were described in detail elsewhere (14-16). This study was reviewed and approved by Institutional Review Board at University of Florida.

Performance Measures

Performance measures assessed include 400-meter walk time (2, 17), 4-meter walk speed (13, 18), and Short Physical Performance Battery (SPPB) (12, 13). Four-hundred-meter usual pace walk time and 4 meter gait speed were calculated in 'seconds' and 'meters per second', respectively. Four meter gait speed was measured from a standing start. The SPPB score consists of three domains: standing balance, walking speed, and repeated chair rises and yields an integer score ranging from 0-12, with each domain contributing 0-4 points. Higher scores indicate higher levels of functioning (12). For these analyses, baseline to 12-month change was calculated.

Anchors, Self-Reported Mobility

For this study, we selected self-reported indicators of mobility as anchors, because they represent the participant's perspective of a construct closely related to lower extremity performance measures, one of the essential criteria for a valid anchor for estimating meaningful change (10). Self-reported mobility status was assessed with the following three separate questions from Disability Questionnaire; "Because of your health, how much difficulty do you have walking a quarter of a mile, which is about 3 or 4 blocks?", "Because of your health, how much difficulty do you have walking several blocks?", and "Because of your health, how much difficulty do you have climbing one flight of stairs?" Participants responded using a five level Likert scale: 'No difficulty', 'a little difficulty', 'some difficulty', 'a lot of difficulty', and 'unable to do the activity'. Participants who answered 'did not do for other reasons', and 'don't know/refused', were not included in the analyses. We operationally defined 5 levels of change over time in self-reported mobility: a) no change; b) small decline (a decrease of one point), c) substantial decline (decrease of 2 or more points); d) small improvement (an increase of one point); and e) substantial improvement (increase of 2 or more points). Participants whose baseline and 12-month responses were both at the ceiling (no difficulty) or floor (unable to do the activity) were removed from the analyses because it would not be possible to detect change beyond the ceiling or floor.

Meaningful Change Analysis

We used both distribution-based and anchor-based methods to obtain estimates of meaningful change in each of the three physical performance measures.

Distribution-Based Methods

We used the effect size method and standard error of measurement (SEM). The effect size is defined as $\delta = (\mu_{12\text{-month}} - \mu_{baseline})/\sigma_{baseline}$, where μ is the mean and σ is the standard deviation of each performance measure. An effect size of 0.2 is considered small, or the minimal value for meaningful change and 0.5 is considered moderate, or substantially meaningful (19, 20). By inverting this formula, mean differences over time corresponding to small and moderate effect sizes were obtained as $0.2\times\sigma_{baseline}$ and $0.5\times\sigma_{baseline}$. SEM was computed as $\sqrt{1-\gamma}$, where γ is the test-retest reliability of the performance measure (21). SEM only yields a single estimate which can be considered a reflection of meaningful change. Test-retest reliability estimates were obtained from the literature for 4m gait speed (0.94) and SPPB (0.9) (22, 23), and from personal communication (Dr. Pahor and Dr. Cesari) for the 400m walk time in seconds (0.904). Since distribution-based methods assume symmetry, they were used to estimate magnitudes of meaningful change without respect to direction of change.

Anchor-Based Method

Consistency of relationships between self-report and performance measures

Because lack of blinding might differentially affect the relationship between change in self-reported and performance measures between the two intervention groups, we first assessed whether estimates of meaningful change using anchor based methods differed by treatment arm. We fitted a two-way analysis of variance model with each performance measure change as the response variable, and treatment group, self-reported anchor change and their interaction as factors of interest. Evidence that the performance measure-anchor association varied between the two treatment groups was based on assessment of the statistical significance of the interaction term.

Calculation of meaningful change

We estimated the mean performance change for each of the three performance measures for each of the 5 levels of self-reported anchor change. We then calculated the difference between the magnitudes of performance change for those self-reporting "no change" to each of the other four anchor change groups. These differences yield estimates of the anchor-based magnitude of substantial decline, minimally meaningful decline, minimally meaningful improvement and substantial

improvement in physical performance.

Symmetry test

In order to determine whether anchor-based estimates of meaningful change were similar by direction of change, we fitted a one way analysis of variance model for each performance change measure using the self-reported anchor change as the main factor of interest with appropriately constructed contrasts of means to compare magnitudes across directions of change.

Results

The baseline characteristics of the total study sample are described in table 1. Sample sizes for individual analyses vary based on the methods and anchors used. The population was diverse in gender and ethnicity. Based on study eligibility, at baseline all had physical performance limitations but were able to walk 400 meters in 15 minutes or less. Over half of participants reported no baseline difficulty with any of the three self reported mobility anchors. Of the 424 who were assessed, at baseline, 4m gait speed was missing in 4 participants. At 12 months 68 (16%) did not have 400 meter walk data, 40 (9%) did not have gait speed, and 26 (6%) did not have SPPB data. Subjects without 12-month data were not significantly different in age or gender from those who did. Baseline 400-meter walk time was shorter for participants who had data at 12 months (475.9 ± 102 seconds) compared to those who did not (566.3 ± 139 seconds) (p<.0001), but baseline 4-meter gait speed (0.74 ± 0.16 vs 0.73 ± 0.12 meters per second) and SPPB (7.5 ± 1.4 versus 7.3 ± 1.4 points) were not statistically different.

Table 1
Baseline Characteristics of Participants in the LIFE-P study by Treatment Arm

	Physical Activity N= 213	Successful Aging N= 211	All
Age (n and %)			
< 80 years	160 (75%)	149 (71%)	309 (73%)
≥ 80 years	53 (25%)	62 (29%)	115 (27%)
Gender, female (n and %)	146 (69%)	146 (69%)	292 (69%)
Race (n and %)			
White	160 (75%)	155 (74%)	315 (74%)
African American/Black	37 (17%)	40 (19%)	77 (18%)
Others	16 (8%)	16 (8%)	32 (8%)
Performance measures (mean ± SD)			
400m walk time (seconds) (n=424)	489.74 ± 113.78	491.09 ± 113.82	490 ± 113.67
4m walk speed (m/s) (n=420)	0.73 ± 0.15	0.75 ± 0.17	0.74 ± 0.16
SPPB score (n=424)	7.57 ± 1.45	7.46 ± 1.38	7.52 ± 1.42
Anchors (n)			
Walk a quarter mile			
no difficulty	111	110	221
a little difficulty	52	42	94
some difficulty	26	32	58
a lot of difficulty	12	10	22
unable to do the activity	2	2	4
Walk several blocks			
no difficulty	116	115	231
a little difficulty	36	27	63
some difficulty	22	27	49
a lot of difficulty	15	12	27
unable to do the activity	3	2	5
Climbing stairs			
no difficulty	115	106	221
a little difficulty	49	57	106
some difficulty	31	32	63
a lot of difficulty	14	12	26
unable to do the activity	0	2	2

Characteristics of change in physical performance and self-reported mobility

Figure 1 shows the distribution of change over 12 months in each of the three performance measures. Figure 2 illustrates how the 5 levels of self reported change were operationally defined and provides sample sizes for each level of self-reported change for each anchor. Participants who reported 'no difficulty' or 'unable to do the activity' at both baseline and 12 months were excluded from our analyses because no further change can be detected when the self-report scale lacks further discrimination, e.g., ceiling and floor effects. Note that many participants reported mobility at the ceiling both at baseline and 12 months, and therefore could not be analyzed. Participants who did not complete the 400m walk at 12 months were excluded from 400m walk time analysis because no change could be calculated. A small number of individuals with missing data for main anchors or other performance measures were excluded from relevant analyses. Despite these exclusions, there were adequate numbers (n=18-68 per cell) to proceed with estimates of change.

Figure 1
Distribution of Change over 12 months in Three Measures of Physical Performance

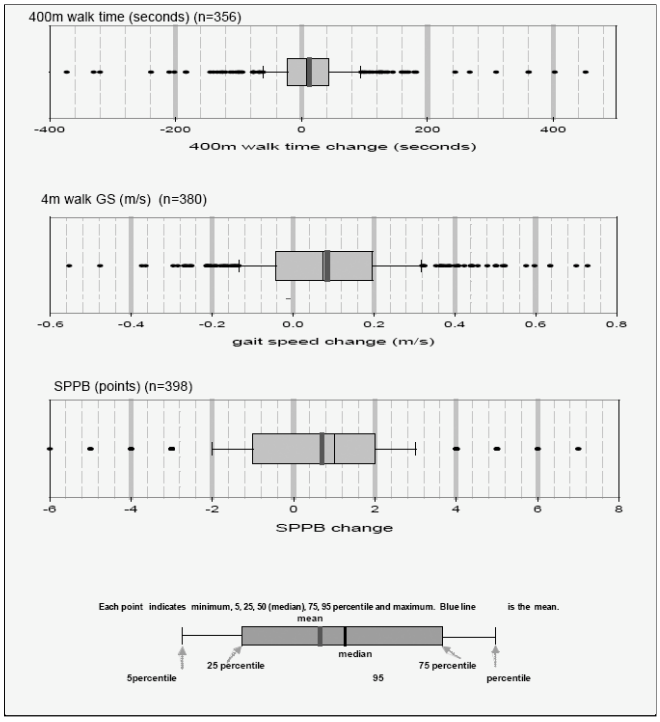
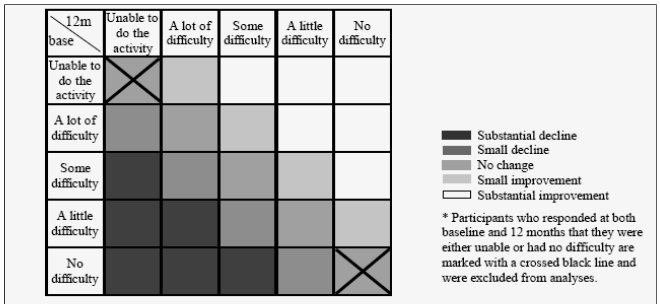


Figure 2
Definitions and Distribution of Change over 12 months in Self-reported Mobility Anchors



	Total	floor	substantial decline	small decline	no change	small improvement	substantial improvement	ceiling	Valid N for analyses
Quarter mile	368	1	**41**	**44**	**47**	**59**	**21**	155	**212**
Several blocks	336	1	**25**	**48**	**38**	**37**	**20**	167	**168**
Stairs	388	0	**24**	**60**	**68**	**60**	**18**	158	**230**

Effect of lack of participant blinding on anchor based estimates of change

There were no significant interaction effects between treatment group and the relationship between self-reported and performance change. The p-values for the nine interaction terms (combinations of three performance measures and three anchors) were 0.1 – 0.9. Since effects were similar across treatment groups, all subsequent analyses used pooled data from both treatment arms.

Meaningful change estimates by distribution-based methods

See Table 2 for estimates of small (minimally meaningful) and moderate (substantially meaningful) change in physical performance using the effect size method and meaningful change using the SEM method. The estimates using the SEM methods consistently fell in between the estimates for the two magnitudes of change using the effect size method.

Table 2
Meaningful Change Estimates in Three Physical Performance Measures using Two Distribution-based Methods

	Standard Error of Measurement method meaningful change	Effect size method	
		small (minimally meaningful) change	substantial change
400m walk time (seconds)	35.22	22.73	56.83
4m gait speed (m/s)	0.04	0.03	0.08
SPPB (points)	0.45	0.28	0.71

test retest reliability: 400 meter walk time: 0.904 (personal communication- see text); 4 meter gait speed: 0.94 (Ostir et al., J Clin Epidemiol, 2002); SPPB: 0.9 (Rolland et al., J Am Geriatr Soc., 2004)

Meaningful change estimates by anchor-based methods

In Figure 3, the average magnitude of change from baseline to 12 months for each of the three physical performance measures is presented for each of the three mobility anchors. Within each of the 9 analyses, the mean performance change for each of the 5 self-reported change levels (for example, substantial decline, small decline, no change, small improvement, substantial improvement) is represented by a single bar. In general, as anchors proceeded from substantial decline to substantial improvement, the height of the bars trended consistently in the expected direction, with isolated exceptions. We noticed that the average magnitude of change in physical performance among persons who self-reported 'no change' was usually not zero, and represented slight improvements for gait speed and SPPB, and slight worsening for 400 meter walk time. Table 3 presents the estimates for the magnitude of meaningful change in each performance measure corrected for the value of physical performance associated with self-report of no change, so that each estimate represents the mean for a level of self-reported change minus the mean for those with a self-report of no change. These estimates based on anchor-based approaches can be compared to the estimates based on distribution-based approaches in table 2. Estimates from the two approaches fell into similar ranges with the exception of the SPPB where substantial decline was much larger with the anchor-based method than with the distribution-based method. Taken together, the best estimates for minimally (or small) meaningful change appeared to be 20-30 seconds for 400m walk time, 0.03-0.05m/s for 4m gait speed, and 0.3 – 0.8 points for SPPB. For substantial change, estimates were in the range of 50-60 seconds for 400m walk time, 0.08m/s for 4 m gait speed and 0.4 – 1.5 points for SPPB.

Figure 3
Absolute Magnitude of Change over 12 Months in Three Physical Performance Measures by Level of Self-reported Change in Three Mobility Anchors

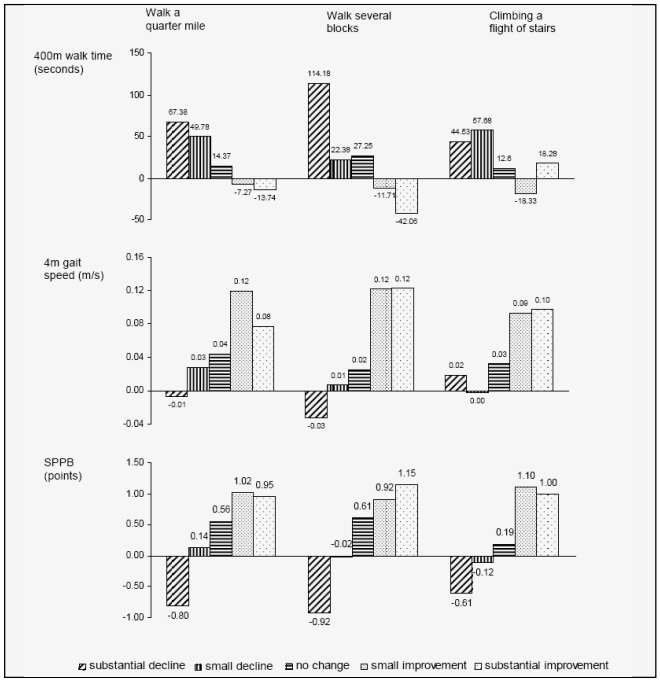


Table 3
Anchor Based Estimates of the Magnitude of Meaningful Change in Three Performance Measures Corrected for Performance Change Among Persons who Self-report No Change

	No Change vs Substantial Decline	No change vs Small Decline	Small Improvement vs No Change	Substantial Improvement vs No Change	Symmetry test, p-value Small Improvement vs Decline	Symmetry test, p-value Substantial Improvement vs Decline
400m walk time (sec)						
Quarter mile	53.01	35.42	-21.64	-28.11	0.71	0.56
Several blks	86.93	-4.87	-38.96	-69.31	0.31	0.72
Stairs	45.08	31.92	5.67	-30.93	0.66	0.35
4m GS (m/s)						
Quarter mile	-0.05	-0.02	0.07	0.03	0.36	0.80
Several blks	-0.06	-0.02	0.10	0.10	0.28	0.62
Stairs	-0.01	-0.03	0.06	0.07	0.64	0.47
SPPB (points)						
Quarter mile	-1.36	-0.42	0.46	0.40	0.95	0.22
Several blks	-1.53	-0.63	0.31	0.54	0.69	0.27
Stairs	-0.80	-0.31	0.91	0.81	0.35	1.00

Symmetry of meaningful change between decline and improvement

In table 3, symmetry can be assessed by comparing estimates of the magnitude of performance change associated with self-reported mobility change between states of decline and improvement. Formal statistical tests of symmetry did not detect significant differences. The statistical tests may have lacked sufficient power due to inadequate sample size.

Discussion

The clinical meaning of change in measures must be understood in order to interpret effects over time in observational studies and clinical trials. Since physical performance measures are becoming preferred indicators of health and function in older adults, it is essential to develop supportive evidence for their use as measures of change. This study contributes to this purpose with the novel finding that in a single blinded clinical trial, the relationship between self-report and performance measures was consistent between two intervention groups. Furthermore, this study provided estimates of the magnitudes of performance change based on clinical trial data. Estimates for gait speed and SPPB appeared to be in the same range as earlier calculations from smaller studies (5). This report provided the first estimates for meaningful change in 400 meter walk time.

This study had several strengths. It was based on longitudinal rather than cross sectional data, making estimates of change more valid and reliable (21). The data came from a large multi-site clinical trial that targeted a population of older adults with mild to moderate mobility limitations. This was a population at high risk of future disability, and therefore likely to be the target of future clinical trials. Research on this population is likely to provide estimates that may be generalizable and useful in future studies. In contrast to observational studies, where performance tends to decline over time and improvement is uncommon, the LIFE-P intervention increased the potential to improve performance, and allowed both the magnitudes of improvement and decline to be estimated. The LIFE-P trial also used meticulous training protocols and quality assurance methods to produce highly reliable performance measures, reducing noise and error in the data. The analyses performed here use state of the art approaches to estimates of meaningful change and the sample size allowed us to compare effects in subgroups based on treatment arm and direction of change.

We used two analytic methods, two magnitudes of change and multiple indicators of self perceived mobility in our work. Anchor-based methods have strong face validity for clinical meaning because they are based on a clinical perception of change while distribution-based methods have optimal capacity to maximize precision (24). The combined use of both methods has been recommended as the best approach to balance clinical meaning and precision (24). We calculated two levels of change; minimally important and substantial. The minimally (clinically) important difference (M(C)ID), has been used traditionally to estimate power and sample size (10, 11, 24). We believe, however, that a larger magnitude of change; one that is considered substantial by patients or moderate by effect size estimates, is valuable. In the clinical arena, changes in health and function are perceived as smaller or larger; and larger changes might be more valued or worth more effort than smaller changes (25). A clinical trial that achieves a substantial rather than a minimally important change might be considered to have had a greater effect. We used multiple anchors, as has been recommended by others, in order to seek consistency across individual items and increase the robustness of our conclusions (24, 25).

This study has limitations. Our self-reported mobility anchors appeared to have significant ceiling effects since over

half of participants had to be excluded from some of the analyses because they reported "no difficulty". This led to reduced sample sizes and lower precision in some cases. Newer self-reported mobility items that include degree of ease as well as degree of difficulty, such as those used in the Health, Aging and Body Composition Study (Health ABC) (2), may expand the ability to detect change, especially improvement, in performance measures. The study had small to moderate rates of missing data at 12-month follow up. It is possible that our estimates might be biased by this censoring. Despite the large sample size of the LIFE-P trial, it was still inadequate to reliably estimate the magnitude of improvement versus decline and we are unable to state with certainty whether symmetry can be assumed.

We made assumptions about the magnitude of change in ordinal anchor measures that could be considered minimally detectable or substantial. This problem is inherent when using ordinal measures which have no defined magnitude between levels. We arbitrarily defined a one level change in degree of difficulty as minimally detectable change and a change of two or more levels as substantial change. This approach has been recommended by others to create more than one level of change (25). We acknowledge that we did not account for differences in baseline degree of difficulty; for example, a one level change could occur from "no difficulty" to "a little", or from "a little" to "some". Further insights into the effect of baseline status on estimates of change require much larger sample sizes. Rasch analysis of ordinal data could help calibrate the distance between ordinal points, as has been done with other mobility scales (26).

Interestingly, the relationship between self reported "no change" and the magnitude of change in the three performance measures was not consistent. When subjects reported "no change", the mean 400m walk time was slightly worse but the mean gait speed and SPPB were slightly better. (Figure 3) We do not know why these effects were discordant.

Finally, this study, like a prior study of meaningful change, yielded estimates of minimally significant meaningful change in SPPB that were smaller than a one unit change in the score (5). While a meaningful change of less than one point could not be detected in an individual, it could be used for groups. For example, differences in between-group mean change can be used for power estimates or for interpreting the importance of intervention effects.

In order to enhance the utility of performance measures in research and clinical care of older adults, further work on meaningful change is needed. Future studies could examine the consistency of estimates of meaningful change across subgroups based on demographics, initial health or functional status, as well as in subgroups with intercurrent events such as beneficial treatments or adverse health events.

While clinically meaningful change has traditionally referred to change that is detectable to patients, significant others, or providers, subclinical change can be important when it predicts future clinically relevant states. For this reason, additional research should assess the effect of change on future events such as hospitalization, mobility disability and survival. We hope that the estimates provided here will be of assistance to both the developers and users of physical performance data so that the mobility, health and function of older adults is improved.

Conclusion

Clinical trials without blinded participants can assume the magnitude of meaningful change is consistent between intervention groups. The first estimates of important change for the 400-meter walk are presented and estimates for gait speed and SPPB are consistent with prior studies. The effect of direction of change on estimates of magnitude remains uncertain.

Acknowledgments: Research Investigators for Pilot Phase of LIFE. **Cooper Institute, Dallas, TX:** Steven N. Blair, P.E.D. – Field Center Principal Investigator; Timothy Church, M.D., Ph.D., M.P.H. – Fielding Center Co-Principal Investigator; Jamile A. Ashmore, Ph.D.; Judy Dubreuil, M.S.; Georita Frierson, Ph.D.; Alexander N. Jordan, M.S.; Gina Morss, M.A.; Ruben Q. Rodarte, M.S.; Jason M. Wallace, M.P.H. **National Institute on Aging:** Jack M. Guralnik, M.D., Ph.D. – Co-Principal Investigator of the Study; Evan C. Hadley, M.D.; Sergei Romashkan, M.D., Ph.D. **Stanford University, Palo Alto, CA:** Abby C. King, Ph.D. – Field Center Principal Investigator; William L. Haskell, Ph.D. – Field Center Co-Principal Investigator; Leslie A. Pruitt, Ph.D.; Kari Abbott-Pilolla, M.S.; Karen Bolen, M.S.; Stephen Fortmann, M.D.; Ami Laws, M.D.; Carolyn Prosak, R.D.; Kristin Wallace, M.P.H.; **Tufts University:** Roger Fielding, Ph.D.; Miriam Nelson, Ph.D.; Dr. Fielding's contribution is partially supported by the U.S. Department of Agriculture, under agreement No. 58-1950-4-401. Any opinions, findings, conclusion, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the U.S. Dept of Agriculture. **University of California, Los Angeles, Los Angeles, CA:** Robert M. Kaplan, Ph.D., M.A. **VA San Diego Healthcare System and University of California, San Diego, San Diego, CA:** Erik J. Groessl, Ph.D. **University of Florida, Gainesville, FL:** Marco Pahor, M.D. – Principal Investigator of the Study; Michael Perri, Ph.D.; Connie Caudle; Lauren Crump, M.P.H; Sarah Hayden; Latonia Holmes; Cinzia Maraldi, M.D.; Crystal Quirin; Dr. Pahor is partially supported by the Geriatric Research, Education and Clinical Center (GRECC) of the Malcom Randall Veteran's Affairs Medical Center, North Florida/South Georgia Veterans Health System, Gainesville, FL. **University of Pittsburgh, Pittsburgh, PA:** Anne B. Newman, M.D., M.P.H. – Field Center Principal Investigator; Stephanie Studenski, M.D., M.P.H. – Field Center Co-Principal Investigator; Bret H. Goodpaster, Ph.D., M.S.; Nancy W. Glynn, Ph.D.; Erin K. Aiken, B.S.; Steve Anthony, M.S.; Sarah Beck (for recruitment papers only); Judith Kadosh, B.S.N., R.N.; Piera Kost, B.A.; Mark Newman, M.S.; Jennifer Rush, M.P.H. (for recruitment papers only); Roberta Spanos (for recruitment papers only); Christopher A. Taylor, B.S.; Pam Vincent, C.M.A.; The Pittsburgh Field Center was partially supported by the Pittsburgh Claude D. Pepper Center P30 AG024827. **Wake Forest University, Winston-Salem, NC:** Stephen B. Kritchevsky, Ph.D. – Field Center Principal Investigator; Peter Brubaker, Ph.D.; Jamehl Demons, M.D.; Curt Furberg, M.D., Ph.D.; Jeffrey A. Katula, Ph.D., M.A.; Anthony Marsh, Ph.D.; Barbara J. Nicklas, Ph.D.; Jeff D. Williamson, M.D., M.P.H.; Rose Fries, L.P.M.; Kimberly Kennedy; Karin M. Murphy, B.S., M.T. (ASCP); Shruti Nagaria, M.S.; Katie Wickley-Krupel, M.S. **Data Management, Analysis and Quality Control Center (DMAQC):** Michael E. Miller, Ph.D. – DMAQC Field Principal Investigator; Mark Espeland, Ph.D. – DMAQC Co-Principal Investigator; Fang-Chi Hsu, Ph.D.; Walter J. Rejeski, Ph.D.; Don P. Babcock, Jr., P.E.; Lorraine Costanza; Lea N. Harvin; Lisa Kaltenbach, M.S.; Wei Lang, Ph.D.; Wesley A. Roberson; Julia Rushing, M.S.; Scott Rushing; Michael P. Walkup, M.S.; The Wake Forest University Field Center is, in part, supported by the Claude D. Older American Independence Pepper Center #1 P30 AG21332. **Yale University:** Thomas M. Gill, M.D.; Dr. Gill is the recipient of a Midcareer Investigator Award in Patient-Oriented Research (K24AG021507) from the National Institute on Aging. The Lifestyle Interventions and Independence for Elders (LIFE-P) Pilot Study is funded by a National Institutes on Health/National Institute on Aging Cooperative Agreement #UO1 AG22376 and sponsored in part by the Intramural Research Program, National Institute on Aging, NIH.

This material was presented at the 61st Annual Meeting of Gerontological Society of America, National Harbor, MD, November 21-25, 2008.

Funding source: Claude D. Pepper Older Americans Independence Center (OAIC) NIH at University of Florida (P30 AG028740-01), and the Claude D. Pepper Older

Americans Independence Center at University of Pittsburgh (P30 AG024827). The Lifestyle Interventions and Independence for Elders (LIFE-P) Pilot Study is funded by a National Institutes on Health/National Institute on Aging Cooperative Agreement #U01 AG22376 and sponsored in part by the Intramural Research Program, National Institute on Aging, NIH.

Disclosure: We certify that no party having a direct interest in the results of the research supporting this article has or will confer a benefit on us or on any organization with which we are associated AND, if applicable, we certify that all financial and material support for this research (eg, NIH or NHS grants) and work are clearly identified in the title page of the manuscript. No conflict of interest and no financial benefits to the authors. No prior publications or presentations made for this specific content.

Disclosure of Corporate sponsorship: Dr Perera is receiving research support from Merck Research Labs to do observational research, and have received such support in the past from Eli Lilly & Co and Ortho Biotech LLC. Also, Dr Perera has been a consultant to Teva Neuroscience in the past.

References

1. Enright PL, McBurnie MA, Bittner V, Tracy RP, McNamara R, Arnold A et al. The 6-min walk test: a quick measure of functional status in elderly adults. Chest 2003;123(2):387-98.
2. Newman AB, Simonsick EM, Naydeck BL, Boudreau RM, Kritchevsky SB, Nevitt MC et al. Association of long-distance corridor walk performance with mortality, cardiovascular disease, mobility limitation, and disability. Jama 2006;295(17):2018-26.
3. Rogers ME, Rogers NL, Takeshima N, Islam MM. Methods to assess and improve the physical parameters associated with fall risk in older adults. Prev Med 2003;36(3):255-64.
4. Topolski TD, LoGerfo J, Patrick DL, Williams B, Walwick J, Patrick MB. The Rapid Assessment of Physical Activity (RAPA) among older adults. Prev Chronic Dis 2006;3(4):A118.
5. Perera S, Mody SH, Woodman RC, Studenski SA. Meaningful change and responsiveness in common physical performance measures in older adults. J Am Geriatr Soc 2006;54(5):743-9.
6. Arai T, Obuchi S, Inaba Y, Nagasawa H, Shiba Y, Watanabe S et al. The effects of short-term exercise intervention on falls self-efficacy and the relationship between changes in physical function and falls self-efficacy in Japanese older people: a randomized controlled trial. Am J Phys Med Rehabil 2007;86(2):133-41.
7. Avraham F, Aviv S, Ya'akobi P, Faran H, Fisher Z, Goldman Y et al. The efficacy of treatment of different intervention programs for patellofemoral pain syndrome--a single blinded randomized clinical trial. Pilot study. ScientificWorldJournal 2007;7:1256-62.
8. Bradshaw BG, Richardson GE, Kumpfer K, Carlson J, Stanchfield J, Overall J et al. Determining the efficacy of a resiliency training approach in adults with type 2 diabetes. Diabetes Educ 2007;33(4):650-9.
9. Sullivan KJ, Brown DA, Klassen T, Mulroy S, Ge T, Azen SP et al. Effects of Task-Specific Locomotor and Strength Training in Adults Who Were Ambulatory After Stroke: Results of the STEPS Randomized Clinical Trial. Phys Ther 2007.
10. Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR. Methods to explain the clinical significance of health status measures. Mayo Clin Proc 2002;77(4):371-83.
11. Wyrwich KW, Bullinger M, Aaronson N, Hays RD, Patrick DL, Symonds T. Estimating clinically significant differences in quality of life outcomes. Qual Life Res 2005;14(2):285-95.
12. Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG 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(2):M85-94.
13. Guralnik JM, Ferrucci L, Pieper CF, Leveille SG, Markides KS, Ostir GV et al. Lower extremity function and subsequent disability: consistency across studies, predictive models, and value of gait speed alone compared with the short physical performance battery. The journals of gerontology 2000;55(4):M221-31.
14. Katula JA, Kritchevsky SB, Guralnik JM, Glynn NW, Pruitt L, Wallace K et al. Lifestyle Interventions and Independence for Elders pilot study: recruitment and baseline characteristics. J Am Geriatr Soc 2007;55(5):674-83.
15. Pahor M, Blair SN, Espeland M, Fielding R, Gill TM, Guralnik JM et al. Effects of a physical activity intervention on measures of physical performance: Results of the lifestyle interventions and independence for Elders Pilot (LIFE-P) study. The journals of gerontology 2006;61(11):1157-65.
16. Rejeski WJ, Fielding RA, Blair SN, Guralnik JM, Gill TM, Hadley EC et al. The lifestyle interventions and independence for elders (LIFE) pilot study: design and methods. Contemp Clin Trials 2005;26(2):141-54.
17. Newman AB, Haggerty CL, Kritchevsky SB, Nevitt MC, Simonsick EM. Walking performance and cardiovascular response: associations with age and morbidity--the Health, Aging and Body Composition Study. The journals of gerontology 2003;58(8):715-20.
18. Studenski S, Perera S, Wallace D, Chandler JM, Duncan PW, Rooney E et al. Physical performance measures in the clinical setting. J Am Geriatr Soc 2003;51(3):314-22.
19. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum; 1988.
20. Yost KJ, Eton DT. Combining distribution- and anchor-based approaches to determine minimally important differences: the FACIT experience. Eval Health Prof 2005;28(2):172-91.
21. Wyrwich KW, Nienaber NA, Tierney WM, Wolinsky FD. Linking clinical relevance and statistical significance in evaluating intra-individual changes in health-related quality of life. Medical care 1999;37(5):469-78.
22. Ostir GV, Volpato S, Fried LP, Chaves P, Guralnik JM. Reliability and sensitivity to change assessed for a summary measure of lower body function: results from the Women's Health and Aging Study. Journal of clinical epidemiology 2002;55(9):916-21.
23. Rolland YM, Cesari M, Miller ME, Penninx BW, Atkinson HH, Pahor M. Reliability of the 400-m usual-pace walk test as an assessment of mobility limitation in older adults. J Am Geriatr Soc 2004;52(6):972-6.
24. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. Journal of clinical epidemiology 2003;56(5):395-407.
25. Wyrwich KW, Wolinsky FD. Identifying meaningful intra-individual change standards for health-related quality of life measures. J Eval Clin Pract 2000;6(1):39-49.
26. MacKnight C, Rockwood K. Rasch analysis of the hierarchical assessment of balance and mobility (HABAM). Journal of clinical epidemiology 2000;53(12):1242-7.