Physical Performance, Self-Rated Health and Mortality among Older Adults in the US and England

Mary Beth Ofstedal, John Bound and Min Hee Kim
University of Michigan

Preliminary Draft
Prepared for presentation at the 2013 meeting of the Population Association of America

Abstract

The integration of physical measures and biomarkers in population-based surveys has become increasingly common in recent years. These measures are considered to be an important complement to self-reported health measures that are typically collected in large surveys. Yet the value of these measures has not been fully demonstrated. Using data from the Health and Retirement Study and the English Longitudinal Study of Ageing, we assess the value of three physical performance measures (grip strength, lung function and walking speed) as predictors of self-rated health and mortality among persons age 65 and older. We use ordered probit models to predict self-rated health and logistic regression to predict 4-year mortality controlling for demographic factors, self-reported functioning (Nagi, IADL, ADL) and self-reported diagnosed conditions. Results suggest that the performance measures capture a dimension of health that is not explained by the self-reported measures.

The integration of physical measures and biomarkers in population-based surveys has become increasingly common in recent years. Many surveys now routinely collect biological materials (e.g., blood, urine), anthropometric measures (e.g., height/weight, waist circumference), physical performance assessments (e.g., grip strength, peak expiratory flow), and genetic material. These measures are considered to be an important complement to self-reported health measures that are typically collected in large surveys and particularly valuable for cross-national comparisons. Yet few studies have investigated the value of physical measures and biomarkers over and above self-reported measures of health and functioning in any rigorous way. Because the measures are fairly expensive to collect and they add considerably to respondent burden, this is an important issue to address.

In this paper, we focus on three physical performance measures (grip strength, lung function and walking speed) and evaluate them in two different ways. First, we examine patterns of correlation between the self-reported measures of physical functioning, disability and disease and the performance measures, to assess the construct validity of the performance measures. Second, we examine whether the performance measures are important predictors of self-rated health and mortality, controlling for self-reported functioning and disease measures. We focus on self-rated health and mortality as outcomes because they are both, in a sense, global measures of health that are widely used in research. By incorporating the performance measures as predictors, we aim to estimate the added contribution of these measures toward explaining self-rated health and mortality. We are also interested in how the associations between physical performance, self-reported functioning and disease, self-rated health and mortality compare across the two countries and by gender.

Data and Methods

We use data from the 2006-2010 waves of the Health and Retirement Study (HRS) in the United States and the 2004-2008 waves of the English Longitudinal Study of Ageing (ELSA) in England.

Health and Retirement Study. The HRS is an ongoing, longitudinal panel study, surveying representative samples of Americans over 50 years of age every two years. The HRS started in 1992 as a longitudinal study of a cohort of individuals born in 1931-1941, and their spouses or partners of any age. The sample was subsequently augmented with additional cohorts in 1993 (AHEAD, born 1923 and earlier) and 1998 to represent the entire population age 51 and older in 1998 (b. 1947 and earlier). Since then, the steady-state design calls for refreshment every six years with a new six-year birth cohort of 51-56 year olds. This was done in 2004 with the Early Baby Boomers (b. 1948-53) and in 2010 with the Mid Boomers (b. 1954-59). The HRS sample includes oversampling of African Americans and Hispanics. The study covers various aspects of aging including such topics as employment and retirement, economic well-being, health status and use of health services, cognition, psychological well-being, and family structure and transfers.

HRS conducts biennial interviews either in-person or over the telephone. In 2006, HRS launched what the study refers to as an enhanced face-to-face interview, which includes a set of physical performance, anthropometric, and blood pressure measurements, blood and saliva samples, and a self-administered questionnaire (SAQ) on psychosocial topics that is left with respondents at the end of the interview and

they are asked to mail back upon completion. The enhanced face-to-face interview is administered to half of the sample in one wave, and the other half in the next wave; thus, each respondent receives it every other wave (i.e., every four years). Respondents who reside in nursing homes or those whose interview is completed by a proxy respondent are not eligible for the physical measures and biomarkers portion of the enhanced face-to-face interview.

English Longitudinal Study of Ageing. The ELSA study is the first panel study in the United Kingdom to connect the full range of topics necessary to understand the economic, social, psychological, and health elements of the aging process. ELSA is based on a nationally representative sample of persons age 50 and over and their partners of any age. The study, which began in 2002, conducts core face-to-face interviews every two years and a nurse visit every four years (during the even-numbered waves). The first nurse visit was conducted in 2004. The nurse visit generally takes place within a few weeks of the main interview and consists of a set physical performance, anthropometric and biological measures. Age-eligible respondents who complete a self-interview are eligible for the nurse visit. The measurements and protocols used in the two studies were designed to enhance comparability (Guyer et al. 2010).

Outcome measures. As global health measures, the two dependent variables used in this study are self-rated health and mortality. Self-rated health is measured by the following question in both surveys: "Would you say your health is excellent, very good, good, fair, or poor?" When self-rated health is modeled as an outcome, we retain the full five response categories (reference=poor). When it is used as a covariate (in the models predicting mortality) we collapse the categories into a dichotomous measure (fair/poor vs. excellent, very good or good).

Information about mortality is obtained from variables provided on the cross-wave tracker file for HRS and from the Index file for ELSA. We use a dichotomous variable indicating whether the respondent died between the 2006 and 2010 waves for HRS or between the 2004 and 2008 waves for ELSA (1=deceased, 0=not deceased).

Performance measures. As noted above, in addition to a detailed personal interview that is conducted biennially, each survey also collects a set of physical performance, anthropometric, and biological measures on respondents every four years. Our focus on this paper is on three performance measures: grip strength, lung function, and walking speed. We provide some information about each of these measures below. More detail on the protocols used to conduct these (and other) measurements in the HRS and ELSA studies are described in a paper by Guyer et al. (2010), in an on-line HRS documentation report (Crimmins et al., 2008), and an on-line ELSA documentation report (Nunn, 2011).

Grip strength was measured using a Smedley spring type hand dynamometer in both HRS and ELSA. The measure was conducted with the respondent standing and holding the dynamometer at a 90 degree angle. Two measurements were taken on each hand alternating between the left and right hand on HRS and three measurements were taken per hand on ELSA. Respondents who had recent hand surgery, pain or inflammation did not complete this measure. To be consistent with HRS, we used the average

of the first two measurements from the dominant hand as our measure of grip strength. If only one measurement was available (which was very rare), we used the value from that measurement.

Both HRS and ELSA obtain measurements of peak expiratory flow, which we use as our measure of *lung function*. HRS obtains this by using a peak flow meter, whereas ELSA uses a spirometer, which provides additional measures of respiratory functioning. Both HRS and ELSA conduct three measures of peak expiratory flow. We took the average of the first two measures to obtain average peak flow rate. If only one measurement was available, we used the value from that measurement.

Both studies also conducted a measure of *walking speed*. Respondents were timed as they walked a short distance at their normal walking pace. The length of the course was 98.5 inches (250 cm) in HRS and SHARE, and 96 inches (244 cm) in ELSA. Two measurements were conducted for each respondent. Walking speed was conducted with respondents aged 65 years or older on HRS and with those aged 60 or older on ELSA. In ELSA, unlike the other performance measures, walking speed was conducted by the interviewer during the main interview (rather than by the nurse during a separate visit). Respondents who were not able to stand or to walk without the aid of another person did not complete this measure. Respondents were allowed to use walking aids (e.g., canes, walkers) during the measurement. Invalid measures or extreme values of more than 30 seconds or below 0.54 seconds in such distance were recoded as a missing. The measures were converted to 'meters per second' for comparability; thus, high values are indicative of faster walking speeds. We averaged the values across the two measurements. In rare cases when one of the measurements had missing values, we used the value from the non-missing measurement.

Covariates. Age, gender, and education level were included as demographic controls. To allow for a non-linear effect of age, we coded age into 5-year groups (60-64 up through 85+). Given the different educational systems in the US and England, we grouped education into four levels, ranging from low to high. Grade levels corresponding with each category are defined in the footnotes for Tables 5 and 7.

Four types of self-report measures were included as covariates: activities of daily living (ADL), instrumental activities of daily living (IADL), physical functioning, and health condition/diseases. The ADL, IADL and physical functioning measures are indicators of whether the respondent has difficulty performing a range of self-care, household and physical activities. We selected items that were asked in both HRS and ELSA, which included 6 ADLs, 6 IADLs, and 10 physical functioning indicators. The specific items are listed in the Table 1 (and subsequent tables). We also included indicators for whether the respondent has been told by a doctor that they have high blood pressure, diabetes, cancer, heart disease, and stroke, and whether they have ever had or been told by a doctor that they have arthritis.

In HRS, each activity and health condition was read to the respondent by the interviewer and respondents answered in relation to each one individually. In ELSA, respondents were presented with show cards listing the activities in each group (ADL, IADL, physical functioning, health conditions), and they were asked to tell the interviewer which of the activities they had difficulty performing and which of the health conditions they had been diagnosed with.

Item missing data is very low for the individual ADL, IADL and condition measures in both studies. However, the physical functioning measures in HRS include a category for "don't do", and relatively large proportions of respondents report that they don't do some of these activities (e.g., 7% climbing stairs, 5% pushing/pulling large objects). Rather than exclude these responses or code them as having difficulty or not, we treated "don't know" responses as missing and imputed them for these variables using a basic imputation model that takes into account gender, age, and self-rated health.

Samples used in the analysis. The analysis samples are limited to participants age 65 or older who completed the physical performance components of the interview.[1] The participation rates in the physical performance measures were very high for both studies, ranging from 87-95% across the three measures in HRS and 87-89% across the three measures in ELSA (Guyer et al. 2010). The analysis is restricted to White participants, because the number of minority participants in ELSA is extremely small. Respondents who had missing values on any of the three performance measures, self-rated health, or the self-reported health measures (other than physical functioning, for which missing values were imputed) were excluded from the analysis sample (n=503 for HRS and n=804 for ELSA). The resulting sample sizes are 3,066 for HRS and 3,179 for ELSA.

Analysis Methods

As a first step in the analysis, we examine bivariate correlations between the performance measures and each of the self-reported functioning, IADL and ADL measures and health conditions. These correlations are derived from R-square statistics for bivariate OLS regression models predicting the score on each performance measure as a function of each self-reported health measure. We also examine R-square statistics from multivariate models predicting each performance measure as a function of all of the functioning, IADL and ADL measures combined, both for the total sample in each study and separately for men and women.

We then use ordered probit models to predict self-rated health and logistic regression to predict mortality over a 4 year period. In the ordered probit model predicting self-rated health, the reference category is 'poor '. Our key interest here is to assess whether the performance measures show significant associations with these health outcomes when self-reported health measures are controlled. Because we are interested in assessing the net effect of the performance measures on self-rated health and mortality, we include measures of *individual* Nagi, IADL and ADL difficulty in the model, as well as individual diseases, rather than using counts or some other type of summary variables for these health domains, as is often done in other studies.

We estimate two models for each outcome, a base model and a full model. The base model includes only demographic factors and the performance measures as controls; the full model adds the self-reported functioning, ADL, IADL and disease measures. (For the model predicting mortality, self-rated health was also included as a control.) This allows us to evaluate whether and how controlling for the self-reported measures moderates the effects of the physical performance tests on self-rated health and

[1] The reason for this age restriction is that walking speed was conducted only on individuals age 65 and over in the HRS (and age 60 and over in ELSA).

mortality. We also tested for interactions between gender and each of the performance measures and self-rated health. Interactions that were statistically significant were retained in the models.

As a final step in the analysis, we estimate a set of nested regression models to evaluate the change in model-chi square through the addition of the performance measures (both singly and as a group) and the sets of physical functioning, IADL and ADL measures. This allows us to assess the relative importance of each type of measure as a predictor of self-rated health and mortality, where "importance" is defined in terms of the amount of variance in the model that is explained by the measure or set of measures. All analyses are weighted and are conducted separately for each country.

Results

Health characteristics. Table 1 presents weighted distributions for the health outcome and control measures. Unadjusted four-year mortality rates are very similar in the two studies. Older adults in the US tend to report more positive self-ratings of their health compared to their British counterparts. The proportion reporting very good health is higher in the US than England and the proportion reporting fair or poor health is lower. In contrast, older adults in the US report higher levels of functional limitation than the British on most measures, with exceptions being lifting 10 pounds, shopping, dressing, bathing and getting in/out of bed. The proportions with any functional limitation and any IADL limitation are higher for the US than England. However, because of the much higher proportions reporting difficulty dressing and bathing in England than the US, the proportion with any ADL limitation is higher for England.[2] Finally, as has been shown previously by Banks and colleagues (2010), older Americans have higher prevalence of disease than their British counterparts for all diseases shown here. For diabetes and cancer the prevalence in the US is nearly double that in England.

Table 2 shows means and standard deviations for the three performance measures, by sex. Both men and women in the US have higher values on grip strength and lung function compared to their British counterparts. However, British men and women have slightly faster walking speeds.

Correlations between performance measures and self-reported measures. Table 3 presents Pearson correlation coefficients between the individual self-reported ADL, IADL, functional and health condition measures and the three performance measures. The statistics shown in the table are R-square statistics from bivariate OLS regression models. The dependent variable is the physical performance measure (continuous scale) and the independent variable is the self-report health indicator.

For most health indicators, the correlations are quite low, less than .05. In general, the self-report items correlate more highly with walking speed than with the other performance measures, although there are some exceptions to this pattern. For example, both pushing/pulling and lifting show slightly higher correlations with grip strength and lung function than with walking speed in the US. In addition, picking up a dime is more highly correlated with grip strength than with the other two performance measures in both countries. Climbing stairs, pulling/pushing and lifting are the individual self-report items that tend

[2] A study by Chan et al. (2012) that used differential item functioning analysis noted that dressing and bathing behaved differently than other ADL and IADL items in the ELSA study.

to correlate most highly with the performance measures. With regard to health conditions, the highest correlations are found between lung disease and lung function, heart disease and walking speed, and arthritis and both grip strength and walking speed. On balance, the correlations tend to be higher for England than for the US for grip strength and walking speed, and slightly higher for the US than England for lung function. Although none of the individual correlations are very high (the highest is 0.182 between lifting and walking speed in England), the pattern of correlations lends some construct validity to the performance measures.

Table 4 presents the R-squares from multivariate models that include all of the self-reported health indicators as predictors. Together the measures account for between 18 and 20 percent of the variance in the physical performance measures in the US and between 13 and 35 percent in England. The results reiterate some of the patterns noted above for the individual measures. As a group, the self-report measures tend to be more highly correlated with walking speed than with grip strength or lung function, and this is true in both countries. In addition, the amount of variance explained by the self-report measures is higher in England than the US for both grip strength and walking speed and lower for lung function. There are also some interesting gender differences. In England, the R-squares are higher for women than for men for all three performance measures, with the differences being most pronounced for grip strength and walking speed. For the US, the R-square for walking speed is somewhat higher for women than for men, but the reverse is true for grip strength and lung function.

Predicting self-rated health. Table 5 presents coefficients, standard errors and wald chi-square statistics for the associations between the demographic factors, performance measures, self-reported health indicators and higher ratings on the self-rated health measure, based on multivariate probit models. We estimate two models for each country: model 1 includes only demographic factors and the performance measures and Model 2 adds the self-reported health measures. We tested for gender differences in the associations between each of the physical performance measures and self-rated health. There were no significant interaction effects, suggesting that the association between physical performance and self-rated health is similar for men and women in the US and England.

Focusing first on the demographic factors, there are some interesting similarities and differences in associations between the United States and England. Controlling for other demographic and health factors, women tend to rate their health higher than men, as do more educated individuals compared to those with the lowest level of education. These associations are highly significant in both countries in both the reduced and full models. In contrast, age shows a different pattern across countries. Controlling for performance on the physical measures (plus sex and education), age is unrelated to self-rated health in the US. However, in England, older individuals tend to rate their health more positively than their younger counterparts.

All three of the performance measures show strong associations with self-rated health in Model 1, and this holds for both countries. When self-reports of functioning and disease are controlled, the effect of grip strength is reduced to non-significance in both the US and England. The effect of walking speed is also attenuated somewhat (the coefficients are reduced in half), but remains highly significant in both

countries. Controlling for the individual self-report measures has little impact on the effect of lung function, which remains significant in the full models for both countries.

None of the ADL measures (dressing through getting in/out of bed) are associated with self-rated health. With regard to the IADL measures, difficulty preparing meals and taking medications are associated with lower ratings of health in the US, whereas in England, difficulty shopping is associated with lower self-rated health. A number of the physical functioning measures show significant associations. Difficulty walking one block, climbing stairs, stooping/kneeling and pushing/pulling are associated with lower health ratings in both the US and England. In addition, difficulty sitting for 2 hours is associated with lower self-rated health in the US, and difficulty lifting is associated with lower ratings in England. Finally, with the exception of stroke and arthritis, all of the disease measures are significantly associated with self-rated health in the expected direction.

Table 6 presents summary results from a series of ordered probit models predicting self-rated health, specifically, chi-square statistics from incremental model tests. The likelihood ratio chi-square from the full model (including all predictors) is compared to those from models removing each of the specified variables or set of variables. The functional limitation, IADL and ADL measures are tested in blocks. By comparing the LR chi-square statistics across models we can get a sense of the relative importance of each variable or set of variables in terms of the marginal proportion of variance that each explains in self-rated health, controlling for other factors in the model.

As a group, the functional limitation measures show by far the largest marginal gain in the variance that is explained in self-rated health among the factors examined here, followed by the three performance measures combined. Of the individual performance measures, walking speed yields the largest difference in model chi-square, followed by lung function. Grip strength is not significantly associated with self-rated health. These patterns all hold for both the US and England. As a group, IADL limitation is significantly associated with self-rated health in the US (but not England), and the difference in chi-square is larger than that for lung function. As noted previously, ADLs are not significantly associated with self-rated health in either country.

Predicting mortality. Table 7 presents results from logistic regression models predicting mortality over a four year period. Again, for each country we estimate two models, a reduced and full model. We also tested for interactions between gender and each of the physical performance measures, as well as between gender and self-rated health. Interactions that were statistically significant were retained in the model and are presented in Table 7.

The effects of demographic factors on mortality are very similar in the US and England. As expected, there is a strong age gradient to mortality in both countries. Controlling for age and physical performance, however, gender and education are unrelated to mortality, and this holds in both countries.

The effects of physical performance on mortality are also quite similar across the two countries. Controlling for only demographic factors, performance on all three of the measures (grip strength, lung function, walking speed) is related to mortality. Respondents with higher values on each of these

measures were less likely to die than those with lower values during the four years after the measurements were taken. The effects of both grip strength and lung function on mortality were similar for men and women; there were no significant gender interactions for these measures. However, the effect of walking speed on mortality was significantly different for women than for men in both countries, such that the slope of the relationship is much steeper for women. Controlling for self-rated health plus the self-reported ADL, IADL, functioning and disease measures in Model 2 attenuates the effects for grip strength and lung function slightly for the US, but both of these measures remain strong predictors of mortality in the full model in both countries. The effect of walking speed on mortality is reduced to insignificance in the full model for men; however, the effect remains strong for women. Again, this pattern holds for both the US and England.

Self-rated health is also an important predictor of mortality in both countries (Model 2). Respondents who report fair or poor self-ratings were more likely to die during the next four years than those who reported excellent, very good or good ratings. The effect of self-rated health on mortality is similar for men and women in the US. However, in England, the effect is only significant for men, not for women. Other things being equal, few of the ADL, IADL and functioning measures are related to mortality and the associations that are observed are inconsistent across countries. Difficulty walking one block and sitting for two hours show moderate positive associations with mortality in the US. In England, difficulty using the phone and climbing stairs are positively associated with mortality, and difficulty getting in/out of bed is negatively associated with mortality.

Diseases show expected associations with mortality in the US for the most part. Diabetes, cancer, lung disease, and heart disease are all positively associated with mortality. Other things being equal, arthritis is negatively associated with mortality in the US. Somewhat surprisingly, cancer is the only disease that is significantly associated with mortality in England.

Table 8 provides incremental chi-square tests for the logistic models predicting 4-year mortality. The three performance measures together explain the largest amount of variance of the different factors examined here. Of the three performance measures, walking speed accounts for the largest difference in model chi-square. Lung function is next largest in the US, followed by grip strength. In England, the difference in chi-square is similar for grip strength and lung function. Second to the performance measures, diseases as a group explain the next largest amount of the variance in mortality in both countries.

Discussion

Our primary goal in this study was to assess the value of physical performance measures in studies of health and aging that collect extensive batteries of self-reported health measures. To address this we first examined correlations between the performance measures and self-reported measures of functioning and health conditions as an indication of construct validity, and then we examined the significance of the performance measures as predictors of two health outcomes, self-rated health and mortality, controlling for the self-report measures. Our findings suggest that physical performance

measures are indeed a valuable and important complement to self-reported health measures, at least for the two studies and the outcomes examined here.

The correlations between the performance measures and the self-reported measures of ADL, IADL, physical functioning and disease, though not as high as might have been expected, showed sensible patterns that suggest that the performance measures have construct validity. For example, activities that require upper body strength, such as pulling/pushing large objects and lifting 10 pounds were most highly correlated with grip strength, whereas those requiring lower body strength, such as walking one block and climbing stairs were most highly correlated with walking speed. In addition, lung disease was more highly correlated with lung function than with the other performance measures and arthritis was more highly correlated with both grip strength and walking speed.

The performance measures were also significant predictors of both self-rated health and mortality. For the most part, these associations remained significant after controlling for individual ADL, IADL, physical functioning and disease indicators. In addition, the patterns of associations between the performance measures and both outcomes were remarkably similar for the US and England. Both lung function and walking speed remained significantly associated with self-rated health in the full models in the US and England, and these effects were similar for men and women. Grip strength was not a significant predictor of self-rated health after controlling for the self-reported measures, and again this was the case in both the US and England. All three of the performance measures were significant predictors of mortality in both countries, net of the full set of covariates. However, for walking speed, the effect was significant only for women, not for men. Again, this interaction effect was observed in both the US and England.

The incremental model chi-square tests suggested that, as a group, the self-reported physical functioning measures outperformed each of the other groups of measures (performance measures, IADLs and ADLs) as predictors of self-rated health. Of the three performance measures, walking speed accounted for more of the explained variance in self-rated health than either lung function or grip strength in both countries. With regard to mortality, however, we observed a somewhat different pattern. The difference in model chi-square was larger for the three performance measures combined, than for any other group of measures (physical functioning, ADL, IADL, conditions/diseases, or self-rated health). Of the three performance measures, walking speed had the highest difference in model chi-square for mortality.

The study has some limitations. One relates to the selectivity of the individuals for whom we have complete data on the performance tests. Although participation in these measurements is fairly high among those who are eligible, the samples exclude the frailest respondents. In both HRS and ELSA, respondents who are interviewed by proxy or who reside in a nursing home are not eligible for the measurements. In addition, even among those who are eligible, participation is related to functional ability, at least in the US (Sakshaug, Couper and Ofstedal 2009). This is due in part to the safety exclusions that are used in the studies, and in part to less healthy respondents opting out of the measures. Exclusion of the frailest respondents may lead to an underestimate of the associations between both the performance measures and the self-reported health measures on one hand and self-

rated health and mortality on the other, though additional analysis is needed to evaluate the influence of selectivity in these samples.

A second limitation is the descriptive nature of the analyses. We have examined correlations or associations between measures without attempting to attribute causality or understand/elucidate the mechanisms through which physical performance is related to self-rated health and mortality. Furthermore, our analysis of self-rated health is cross-sectional, with the performance and functioning measures and the self-rating of health taken from the same interview.

In future extensions of this work we plan to examine the influence of physical performance on subsequent *transitions* in health and functioning in the US and England. We will also explore whether these associations vary by age and, in the US, by race and ethnicity. Finally, we plan to extend the analysis to other countries, including the 15 countries in Europe that are part of the Survey of Health, Aging and Retirement in Europe (SHARE), South Korea (the Korean Longitudinal Study of Aging, KLOSA), and Japan (the Japanese Study of Aging and Retirement, JSTAR) to examine whether the patterns we observed in the US hold up in these other settings.

Acknowledgements

Support for this research is provided by the National Institute on Aging through a pilot grant from the Michigan Center for the Demography of Aging (P30 AG0128467). The Health and Retirement Study is conducted by the Survey Research Center at the University of Michigan (David Weir, PI) and is funded by the National Institute on Aging (NIA U01AG009740) and the Social Security Administration. The English Longitudinal Study of Ageing is conducted by the National Centre for Social Research in London (Michael Marmot, PI) and is jointly funded by the National Institute on Aging in the US and a consortium of government departments in the UK (Department for Communities and Local Government, Department for Transport, Department for Work and Pensions, Department of Health, HM Revenue and Customs, and the Office for National Statistics).

References

Banks J., Muriel A., Oldfield Z., Smith J.P. (2010) Disease prevalence, disease incidence, and mortality in the United States and in England, Demography, 2010, Volume 47, Supplement 1, Pages S211-S231

Chan, K., J.D. Kasper, J. Brandt and L.E. Pezzin. (2012) Measurement Equivalence in ADL and IADL Difficulty Across International Surveys of Aging: Findings From the HRS, SHARE, and ELSA. *Journals of Gerontology* 67(1), 121–132.

Crimmins, E., H. Guyer, K. Langa, M.B. Ofstedal, R. Wallace and D. Weir. 2008. Documentation of physical measures, anthropometrics and blood pressure in the Health and Retirement Study. HRS Documentation Report, DR-11. Survey Research Center, University of Michigan. http://hrsonline.isr.umich.edu/sitedocs/userg/dr-011.pdf

Guyer, H., M.B. Ofstedal, C. Lessof, K. Cox and H. Jurges. (2009) Collecting Physical Measure and Biomarker Data on Cross-National Studies. Presented at the SHARE 2nd International User Conference, Mainz, Germany, Oct. 12-13, 2009.

Nunn, S. 2011. User guide for the ELSA wave 2 nurse visit data. English Longitudinal Study of Ageing, Institute for Fiscal Studies, University College London. (http://www.ifs.org.uk/ELSA/publicationDetails/id/5902)

Sakshaug, J.W., M.P. Couper and M.B. Ofstedal. 2009. Characteristics of physical measurement consent in a population-based survey of older adults. *Medical Care* 48(1): 64-71.

Table 1. Weighted frequency distributions for health covariates and outcomes: US and England

	United States	England
Health outcomes		
Mortality: died w/in 4 years	15.3	15.2
Self-rated health		
Excellent	10.3	10.6
Very good	33.1	26.4
Good	33.5	35.4
Fair	17.8	21.7
Poor	5.3	5.9
Health controls		
Functional limitation		
Walking one block	12.9	11.0
Sitting two hours	16.1	12.7
Getting up from chair	40.9	28.7
Climbing several flights stairs	48.7	43.7
Climbing one flight stairs	16.5	15.3
Stooping, kneeling	50.0	41.9
Reaching up	14.6	10.5
Pushing/pulling	24.4	18.4
Lifting 10 pounds	19.6	27.3
Picking up a dime	7.1	5.0
1+ functional limitation	73.6	66.0
IADL limitation		
Using a map	10.4	5.5
Meal preparation	3.5	2.7
Shopping	6.8	7.9
Using telephone	3.7	2.0
Taking medications	2.9	1.4
Managing money	4.6	2.0
1+ IADL limitation	19.1	14.0
ADL limitation		
Dressing	7.0	13.9
Bathing	4.2	12.6
Eating	1.6	1.3
Using the toilet	4.3	2.0
Walking across room	4.0	1.1
Getting in/out of bed	2.6	3.7
1+ ADL limitation	14.3	22.3
Diseases and conditions		
High blood pressure	60.1	49.7
Diabetes	19.2	9.6
Cancer	20.9	9.2
Lung disease	10.7	8.7
Heart disease	32.0	27.8
Stroke	6.9	5.9
Arthritis	67.4	42.6
Total %	100.0	100.0
Sample size	3,066	3,179

Table 2. Means and standard deviations (in parens) for performance measures, by gender

	United States		**England**	
	Male	Female	Male	Female
Grip strength	36.72 (7.99)	21.37 (5.03)	33.84 (8.45)	19.79 (5.83)
Lung function	380.99 (119.05)	246.06 (73.27)	358.51 (131.44)	225.76 (84.38)
Walking speed (meters/second)	0.83 (0.24)	0.74 (0.23)	0.87 (0.26)	0.79 (0.27)

Table 3. R-squared statistics for the association between performance measures and individual ADL, IADL, Nagi and disease measures (from bivariate OLS regressions)

Self-report health indicator	**United States**			**England**		
	Grip strength	Lung function	Walking speed	Grip strength	Lung function	Walking speed
Dressing	0.009	0.014	0.028	0.016	0.006	0.094
Bathing	0.022	0.030	0.048	0.058	0.028	0.132
Eating	0.012	0.014	0.020	0.021	0.007	0.010
Toileting	0.019	0.011	0.025	0.011	0.002	0.021
Walking across room	0.021	0.022	0.042	0.015	0.006	0.033
In/out bed	0.005	0.010	0.009	0.014	0.003	0.048
Using map	0.039	0.041	0.033	0.041	0.019	0.050
Meal prep	0.019	0.026	0.041	0.032	0.016	0.066
Shopping	0.043	0.050	0.062	0.085	0.050	0.155
Using phone	0.007	0.010	0.017	0.003	0.003	0.014
Taking meds	0.008	0.012	0.023	0.015	0.005	0.017
Managing money	0.026	0.031	0.044	0.020	0.009	0.039
Walking one block	0.039	0.056	0.077	0.037	0.021	0.151
Sitting 2 hrs	0.003	0.007	0.009	0.021	0.006	0.038
Getting up from chair	0.023	0.018	0.054	0.043	0.013	0.096
Climbing sev flt stairs	0.068	0.082	0.084	0.061	0.046	0.149
Climbing one flt stairs	0.050	0.059	0.080	0.064	0.037	0.164
Stooping, kneeling	0.044	0.028	0.054	0.068	0.026	0.139
Reaching up	0.018	0.020	0.026	0.045	0.017	0.066
Pushing/pulling	0.085	0.083	0.058	0.094	0.051	0.158
Lifting 10 lbs	0.097	0.085	0.080	0.136	0.088	0.182
Picking up dime	0.020	0.007	0.008	0.038	0.008	0.026
High blood pressure	0.006	0.003	0.007	0.003	0.003	0.012
Diabetes	0.000	0.000	0.006	0.000	0.001	0.012
Cancer	0.000	0.001	0.000	0.000	0.000	0.000
Lung disease	0.001	0.043	0.004	0.001	0.014	0.009
Heart disease	0.001	0.003	0.017	0.001	0.001	0.015
Stroke	0.005	0.008	0.019	0.005	0.005	0.016
Arthritis	0.026	0.008	0.016	0.065	0.010	0.054

Table 4. R-squared statistics for the association between performance measures and combined ADL, IADL, Nagi and disease measures (from multivariate OLS regressions), by gender

Respondent gender	**United States**			**England**		
	Grip strength	Lung function	Walking speed	Grip strength	Lung function	Walking speed
Both sexes	0.181	0.184	0.201	0.220	0.129	0.350
Female	0.149	0.153	0.218	0.250	0.105	0.386
Male	0.167	0.200	0.168	0.130	0.092	0.279

Table 5. Probit results predicting (higher) self-rated health among white persons age 65+, United States and England (weighted)

Predictors	United States (HRS)								England (ELSA)							
	Model 1				Model 2				Model 1				Model 2			
	Coeff	se	Wald Chisq		Coeff	se	Wald Chisq		Coeff	se	Wald Chisq		Coeff	se	Wald Chisq	
Intercept (exc)	-3.534	0.166		***	-1.836	0.185		***	-3.718	0.142		***	-2.220	0.159		***
Intercept (v good)	1.178	0.036		***	1.338	0.040		***	1.007	0.031		***	1.127	0.035		***
Intercept (good)	2.162	0.043		***	2.511	0.050		***	2.043	0.038		***	2.329	0.044		***
Intercept (fair)	3.132	0.056		***	3.725	0.069		***	3.132	0.051		***	3.625	0.060		***
Intercept (poor)	---				---				---				---			
Female	0.497	0.066	55.981	***	0.355	0.070	25.584	***	0.430	0.058	55.469	***	0.426	0.061	49.375	***
Age 65-69	---				---				---				---			
Age 70-74	0.069	0.058			0.042	0.060			0.089	0.052			0.110	0.053		
Age 75-79	0.081	0.064	5.168		0.104	0.065	9.059		0.120	0.056	48.890	***	0.136	0.058	50.507	***
Age 80-84	0.042	0.072			0.069	0.075			0.338	0.067			0.336	0.069		
Age 85+	0.176	0.082			0.248	0.086			0.535	0.087			0.589	0.090		
Educ 1[a]	---				---				---				---			
Educ 2	0.233	0.062	46.152	***	0.192	0.063	24.715	***	0.138	0.053	16.524	***	0.158	0.054	20.532	***
Educ 3	0.368	0.069			0.323	0.071			0.186	0.054			0.220	0.055		
Educ 4	0.434	0.069			0.293	0.071			-0.024	0.068			-0.001	0.070		
Grip strength	0.014	0.003	17.138	***	0.003	0.004	0.708		0.013	0.003	19.279	***	0.006	0.003	3.327	
Lung function	0.001	0.000	28.470	***	0.001	0.000	6.905	**	0.001	0.000	28.821	***	0.001	0.000	23.442	***
Walking speed	1.026	0.093	121.440	***	0.521	0.099	27.973	***	1.526	0.085	320.740	***	0.654	0.096	46.343	***
Dressing					-0.033	0.095	0.119						-0.098	0.067	2.136	
Bathing					-0.103	0.129	0.637						-0.047	0.070	0.450	
Eating					0.177	0.186	0.902						0.076	0.186	0.167	
Toileting					-0.066	0.120	0.302						-0.212	0.152	1.940	
Walking across room					0.062	0.131	0.227						-0.120	0.205	0.341	
In/out bed					-0.201	0.147	1.876						-0.007	0.119	0.003	
Using map					-0.072	0.076	0.885						0.048	0.093	0.271	

Table 5 (cont.)

Predictors	United States (HRS)								England (ELSA)							
	Model 1				Model 2				Model 1				Model 2			
	Coeff	se	Wald Chisq		Coeff	se	Wald Chisq		Coeff	se	Wald Chisq		Coeff	se	Wald Chisq	
Meal prep					-0.307	0.143	4.575	*					-0.028	0.140	0.041	
Shopping					0.032	0.112	0.081						-0.199	0.093	4.645	*
Using phone					-0.042	0.126	0.110						-0.170	0.148	1.327	
Taking meds					-0.451	0.150	9.070	**					0.145	0.183	0.627	
Managing money					0.058	0.128	0.207						-0.179	0.156	1.323	
Walking one block					-0.474	0.080	34.926	***					-0.220	0.077	8.283	**
Sitting 2 hrs					-0.190	0.064	8.783	**					-0.126	0.065	3.767	
Getting up from chair					-0.058	0.054	1.158						-0.081	0.052	2.374	
Climbing sev flt stairs					-0.335	0.052	41.707	***					-0.384	0.048	63.055	***
Stooping, kneeling					-0.216	0.053	16.906	***					-0.115	0.049	5.574	*
Reaching up					-0.033	0.067	0.239						-0.027	0.071	0.148	
Pushing/pulling					-0.423	0.063	45.369	***					-0.142	0.067	4.545	*
Lifting					-0.128	0.070	3.361						-0.291	0.058	24.922	***
Picking up dime					-0.050	0.087	0.332						-0.079	0.095	0.694	
High blood pressure					-0.139	0.045	9.480	**					-0.206	0.040	26.315	***
Diabetes					-0.223	0.057	15.560	***					-0.445	0.068	43.049	***
Cancer					-0.220	0.053	17.000	***					-0.283	0.068	17.537	***
Lung disease					-0.293	0.074	15.667	***					-0.479	0.072	44.554	***
Heart disease					-0.372	0.049	57.768	***					-0.407	0.045	80.867	***
Stroke					-0.067	0.088	0.587						-0.138	0.085	2.605	
Arthritis					-0.030	0.050	0.363						-0.070	0.044	2.546	
Log Likelihood (N)	-3419.807 (3066)				-3053.604 (3066)				-4169.816 (3179)				-3791.145 (3179)			

* p < .05 ** p < .01 *** p < .001

[a]Education is defined as follows in the two surveys:

HRS: 1=< HS, 2=HS grad, 3=some college, 4=college grad or higher

ELSA: 1=< O level equivalent, 2=O level equivalent or A level equivalent, 3=> A level equivalent, 4=foreign/other

Table 6. Summary of probit results predicting self-rated health: Incremental model chi-square tests

	United States				England			
	LR Chisq	Dif Chisq	df		LR Chisq	Dif Chisq	df	
Full model	1108.28				1344.75			
Remove grip strength	1107.58	0.71	1		1341.42	3.33	1	
Remove lung function	1101.38	6.91	1	**	1318.26	26.49	1	***
Remove walking speed	1080.26	28.02	1	***	1298.31	46.43	1	***
Remove 3 perf meas	1063.47	44.82	3	***	1248.26	96.49	3	***
Remove func lim	797.24	311.05	10	***	1124.49	220.26	10	***
Remove IADL	1088.71	19.58	6	**	1336.01	8.74	6	
Remove ADL	1103.72	4.56	6		1337.85	6.90	6	

Table 7. Logistic regression results predicting 4-year mortality among whites age 65+, United States and England (weighted)

Predictors	United States (HRS)								England (ELSA)							
	Model 1				Model 2				Model 1				Model 2			
	Coeff	se	Wald Chisq		Coeff	se	Wald Chisq		Coeff	se	Wald Chisq		Coeff	se	Wald Chisq	
Intercept	1.819	0.460	15.945	***	0.599	0.538	1.239		-0.189	0.413	0.210		-1.429	0.503	8.066	**
Female	-0.111	0.386	0.082		-0.155	0.418	0.138		-0.267	0.364	0.539		0.304	0.442	0.472	
Age 65-69	---				---				---				---			
Age 70-74	-0.016	0.202			-0.008	0.211			0.530	0.195			0.541	0.201		
Age 75-79	0.151	0.202	30.527	***	0.134	0.212	29.235	***	0.728	0.194	76.315	***	0.713	0.202	73.231	***
Age 80-84	0.370	0.208			0.386	0.220			1.052	0.204			1.115	0.213		
Age 85+	0.970	0.215			1.024	0.230			1.873	0.225			1.941	0.238		
Educ 1[a]	---				---				---				---			
Educ 2	-0.046	0.165	2.867		0.018	0.176	3.811		-0.097	0.170	6.245		-0.133	0.176	5.788	
Educ 3	0.138	0.184			0.311	0.196			0.045	0.169			0.009	0.178		
Educ 4	-0.180	0.193			0.008	0.205			0.446	0.195			0.426	0.203		
Grip strength	-0.038	0.010	14.123	***	-0.031	0.011	8.516	**	-0.024	0.009	7.247	**	-0.027	0.010	8.137	**
Lung function	-0.004	0.001	41.729	***	-0.003	0.001	22.947	***	-0.002	0.001	14.910	***	-0.002	0.001	7.682	**
Walking speed	-0.987	0.364	7.376	***	-0.700	0.385	3.307		-0.798	0.346	5.316	*	-0.093	0.384	0.059	
Interaction: Female*WalkSpeed	-2.411	0.544	19.646	***	-2.069	0.573	13.028	***	-1.326	0.470	7.959	**	-1.634	0.517	9.988	**
Fair/poor health					0.395	0.155	6.489	*					0.759	0.183	17.160	***
Interaction: Female*fair/poor hlth					---								-0.631	0.254	6.178	*
Dressing					-0.171	0.240	0.510						-0.216	0.185	1.363	
Bathing					0.470	0.284	2.734						-0.026	0.182	0.020	
Eating					0.013	0.416	0.001						0.031	0.481	0.004	
Toileting					0.017	0.301	0.003						0.223	0.403	0.306	
Walking across room					0.040	0.296	0.018						0.610	0.452	1.824	
In/out bed					-0.402	0.382	1.108						-1.208	0.393	9.472	**

Table 7 (cont.)

Predictors	United States (HRS)								England (ELSA)							
	Model 1				Model 2				Model 1				Model 2			
	Coeff	se	Wald Chisq		Coeff	se	Wald Chisq		Coeff	se	Wald Chisq		Coeff	se	Wald Chisq	
Using map					0.335	0.201	2.772						-0.293	0.253	1.340	
Meal prep					0.256	0.314	0.668						0.405	0.323	1.577	
Shopping					0.215	0.248	0.748						-0.002	0.228	0.000	
Using phone					0.182	0.295	0.383						1.096	0.315	12.081	***
Taking meds					0.144	0.330	0.189						-0.087	0.431	0.040	
Managing money					0.221	0.288	0.588						0.432	0.350	1.522	
Walking one block					0.393	0.193	4.145	*					0.105	0.190	0.304	
Sitting 2 hrs					-0.383	0.191	4.003	*					-0.048	0.183	0.067	
Getting up from chair					-0.205	0.157	1.711						-0.105	0.151	0.479	
Climbing sev flt stairs					0.103	0.158	0.424						0.436	0.145	9.113	**
Stooping, kneeling					0.043	0.160	0.073						0.184	0.145	1.594	
Reaching up					-0.078	0.183	0.182						-0.065	0.191	0.115	
Pushing/pulling					-0.245	0.176	1.947						0.114	0.180	0.402	
Lifting					0.310	0.185	2.799						-0.014	0.168	0.007	
Picking up dime					0.132	0.234	0.316						-0.491	0.275	3.189	
High blood pressure					-0.023	0.136	0.029						-0.099	0.123	0.644	
Diabetes					0.517	0.151	11.736	***					0.116	0.183	0.403	
Cancer					0.365	0.147	6.195	*					0.723	0.174	17.227	***
Lung disease					0.534	0.188	8.101	**					0.151	0.189	0.631	
Heart disease					0.356	0.135	6.971	**					0.144	0.128	1.262	
Stroke					-0.282	0.226	1.559						0.408	0.211	3.731	
Arthritis					-0.341	0.147	5.387	*					-0.242	0.133	3.311	
LR Chi-square (df)	381.396 (12)				494.350 (41)				339.765 (12)				446.557 (42)			

* p < .05 ** p < .01 *** p < .001

[a]Education is defined as follows in the two surveys:

HRS: 1=< HS, 2=HS grad, 3=some college, 4=college grad or higher

ELSA: 1=< O level equivalent, 2=O level equivalent or A level equivalent, 3=> A level equivalent, 4=foreign/other

Table 8. Summary of logistic results predicting 4-year mortality: Incremental model chi-square tests

	United States				England			
	Model Chisq	Dif Chisq	df		Model Chisq	Dif Chisq	df	
Full model	494.35				446.56			
Remove grip strength	485.71	8.64	1	**	438.42	8.14	1	**
Remove lung function	470.85	23.50	1	***	438.79	7.77	1	**
Remove walking speed	454.85	39.50	2	***	429.31	17.25	2	***
Remove 3 perf meas	398.39	95.96	4	***	407.46	39.10	4	***
Remove func lim	479.18	15.17	10		428.22	18.34	10	*
Remove IADL	482.44	11.91	6		429.25	17.31	6	**
Remove ADL	489.84	4.51	6		431.61	14.95	6	*
Remove conditions	455.42	38.93	7	***	420.69	25.87	7	***
Remove self-rated health	487.96	6.39	1	*	429.57	16.99	2	