Gender differences in measures of health in an English population survey of older people: A latent variable modelling approach

George B. Ploubidis[1] & Emily Grundy[1]
[1]London School of Hygiene and Tropical Medicine

Abstract

Our primary objective was to introduce a multimethod measurement model to survey based health assessment that combines information from both self report and observed/objective health indicators in a latent global health index. A secondary objective was to use this index to examine gender and socio-economic differentials in the health status of the older population. Three objective/observed and three self report health indicators were combined in a latent variable model where valid health, systematic error due to measurement method induced bias, and random error were empirically estimated, utilizing data from the second wave of the English Longitudinal Study of Ageing (ELSA), N= 5,965. Objective/observed and self report health indicators were equally biased by method and measurement error, with the exception of self report of functional limitations which was the most reliable health indicator. The gender gap in health of older people was confirmed, with women having worse health compared to men, when measurement error was controlled. The widely reported socioeconomic gradients in health were confirmed, as well as a tendency of the less advantaged and the least educated to over-report health problems. The multimethod measurement model is offered as a tool for future researchers, that may wish to use the latent global health index and the latent factors that quantify measurement method induced error.

Paper prepared for presentation at the Population Association of America in New Orleans (LA), April, 2008. This research was conducted as part of a research programme on Major Ageing and Gender Issues in Europe (MAGGIE) funded under the EU Sixth Framework.

Introduction

The reliable and accurate measurement of population health is fundamental to the development of evidence for health policy and for the evaluation and planning of health systems and intervention programs. However, health as recognized in the standard World Health Organization definition is a multi-dimensional concept encompassing physical, social and mental well-being who as such is difficult to measure consistently across populations and population sub-groups. The considerable controversy surrounding the issue of trends in indicators of the health status of populations, especially older populations, to a large extent arises from measurement problems and the difficulties involved in making comparisons between health indicators derived in different ways (Robine et al, 1992) (Wolf et al, 2005). Similar problems have bedeviled attempts to make international comparisons of health status as, even if questions are harmonised, it is clear that the ways people respond to them are not (Robine, Jagger and Euro, 2003).

Efforts to measure health in population surveys have generated a plethora of indicators, which can be broadly classified into self reported or subjective measures and observed – or objective- measures. Self reported measures typically include self-rated general health status, reported presence of longstanding illness; self reported problems with functional activities and, in some studies, self report of particular conditions or symptoms (Banks et al, 2006). Observed/objective indicators include assessments of physical functioning, such as balance tests, timed chair stands, grip strength and lung function, as well as biological information such as blood pressure and results from analyses of blood analytes and in some cases, physician assessment of specific morbidities (Banks et al, 2006). Both types of indicators have been used as outcomes in health related research, sometimes with contradictory results (Barsky, 1988). This seems to be a particular issue in older populations as subjective evaluation of health is markedly influenced by people's willingness to accept impairment or disability as being normal for their age (Svanborg, 1988); this 'willingness' is, however, likely to vary between populations and population subgroups.

Other studies which have examined differentials in a range of self reported indicators of health status have found varying associations suggesting that, perhaps not surprisingly, different indicators capture different dimensions of health (Grundy and Sloggett, 2003). Investigations of surveys in the UK concluded that the direction and magnitude of gender differences in health vary according to the particular health indicator used (Blaxter, 1990). Previous studies have also found considerable discrepancies between self-reported and observed/objective indicators (Elam et al, 1991) (Ferrer et al,1999). For example, McDowell and Newell reported a discrepancy between self reported limitations in function in instrumental activities of daily living and actual physical impairment. Similarly large differences between reported difficulties with Activities of Daily Living and observed performance based indicators have been reported (Myers et al.). It appears that a significant portion of variance in self-reported disability can be explained by demographic, cultural, social and psychological variables such as gender, familiarity with scale activities and depression (McDowell, 1996).

Response styles, social desirability bias, extremity scoring and certain personality characteristics have also been implicated as sources of reporting bias (Hebert et al, 2001) (Adams et al, 2005).

This variance in the reliability of measuring health with different indicators is of significant importance and needs urgent attention. In a similar manner, the identification of sources of bias in survey based health measurement is of equal importance, since understanding biases and response shifts in the measuring of health has considerable implications for health services research, clinical care, and health policy applications. In this paper we use principles and methods of measurement borrowed from psychometrics, to develop a measurement model that combines information from both self reported and objective health indicators. Psychometric (latent variable) models assess the common variance across multiple indicators of a specific construct and are often used when measurement error may bias parameter estimates. The variance in survey based health measures can be decomposed into true health variance and error variance. Error variance can be further subdivided into measurement method induced error (systematic error) and random error variance (Eid et al, 2006). Mathematically this relationship is expressed in Equation 1

$$\sigma_o = \sigma_\iota + \sigma_m + \sigma_r \quad (1)$$

where σ_o is the observed variance of the health indicators, σ_ι the valid health variance, σ_m the systematic error due to the measurement instrument and σ_r the random error variance. Within psychometric theory this is commonly referred to as a multimethod measurement model, where two or more methods are used to measure a single latent characteristic. This model assumes that health is a single latent construct; measured by several distinct methods (the model can easily be extended to accommodate any number of health measurement methods, including the use of biomarkers), with multiple indicators within each. Multimethod measurement models have been widely used in psychology (Tram and Cole 2006), (Beresnevaite et al, 2007), (Majdandzic and van den Boom, 2007), in the assessment of smoking (Palmer, Dwyer and Semmer 1994), in psychiatry (Prinstein et al, 2001) and paediatrics (Nichter et al,1995), among others. Latent variable models have been previously used in health measurement related research, for example Boniface (Boniface and Tefft, 2001) used structural equation modeling to derive a latent health index and similarly, Lillard and Swaminathan (Lillard and Swaminathan, 2000) proposed a unidimensional model for health, where multiple health indicators were combined into a conceptual latent measure. However, none of the previous studies attempted to combine objective and self report health indicators in a single model. By employing a multimethod measurement model the primary goal in this paper is to estimate a latent global health measure that combines information from both self reported and observed/objective health indicators and to use this to examine gender and socio-economic differentials in the health status of the older population.

Methods

Data and measures

We use data from the second wave (2004) of the English Longitudinal Study of Ageing (ELSA), a nationally representative multi-purpose sample of the population aged 50 and over living in England. The ELSA sample was drawn from households that responded to the 1998, 1999 or 2001 rounds of the Health Survey for England (HSE), a stratified random sample of all households in England. Response rates to these HSE rounds were 69%, 70% and 67% respectively (Marmot et al, 2002). A total of 19,924 individuals in households responded to the HSE who would have been aged 50 years by 2002. Of these, 11,392 (65.7%) became ELSA respondents (core participants). Analyses of socio-demographic characteristics against census results indicated that the ELSA sample remained representative of the population (Marmot, et al 2002). In the second wave of the ELSA 8,780 core participants were available. In our analysis we included participants that had complete data in all six health indicators, leaving us with an analysis sample of 5,965.

Measures

Health indicators

The three observed indicators we used were grip strength; a measure of respiratory function (Forced Vital Capacity – FVC), and chair rise speed (all measured by nurses). Grip strength was measured three times for each hand using a dynamometer in the form of a handle. Participants were instructed by the nurse to squeeze the handle as hard as they could for two or three seconds and then let go. We calculated the mean grip strength of the dominant hand and used this in further analysis. FVC is the total amount of air that can forcibly be blown out after full inspiration, measured in liters. The highest technically satisfactory reading was used in the analysis. With respect to chair rise speed, the time it took participants to complete five chair rises having their arms folded at their chest was recorded. In the analysis we used the inverse of this variable so high scores indicate good health. The self report measures were self-assessed general health, presence of long standing – chronic- illness, and the presence of one or more Functional Limitations (FL) such as walking a 100 yards, climbing a flight of stairs without resting and lifting/carrying over 10 lbs (like a heavy bag of groceries) among others. In Table 1 we present descriptive statistics and the response formats of all health indicators as they were entered in the analysis. All indicators were recoded so as high values represent good health.

INSERT TABLE 1 ABOUT HERE

Co-variates

Health is known to be strongly associated with demographic characteristics and with various indicators of socio-economic status (Huisman, 2003) (Banks et al, 2006). We accordingly included age, gender, marital status and five indicators of socio-economic status in the analyses, both to control their effect and obtain adjusted estimates of the associations between socio-demographic characteristics and health. Age was used as a continuous variable (mean = 65.4 s.d = 10.1), whereas dummy variable coding was applied to all other variables. The marital status variable was based on current legal status and distinguished five groups: those in first marriage; those in a second or subsequent marriage; the never-married; divorced or legally separated respondents, and the widowed. Socio-economic status was measured using the occupationally based National Statistics socio-economic classification (NS-SEC); highest educational level; housing tenure; income; and non-housing wealth. We used the five category version of NS-SEC which allocates people to managerial and professional; intermediate; small employees and own account workers; lower supervisory and technical workers; and those in semi-routine occupations with those in the highest status category (managerial and professional) serving as the reference group. Social class allocation was based on own most recent (or current where available) occupation.

Similarly, five groups were derived to reflect the participants' educational qualifications, The reference group comprised participants with a degree or equivalent qualification, the second participants with GCE A level or equivalent qualifications (exams normally taken around age 18), the third respondents with O level or CSE grade qualifications (exams taken at age 16), the fourth with foreign qualifications and the fifth those without a formal educational qualification. Housing tenure was recoded to a binary variable distinguishing owners (reference category) from non-owners. Finally, gross total non housing wealth and equivalised household income were recoded to quintiles, with the quintile presenting the highest income/wealth serving in both cases as the reference group.

INSERT FIGURE 1 ABOUT HERE

<u>Statistical modelling</u>

In the first stage of the modelling we derive estimates of a latent variable 'health' and two method factors representing systematic measurement error using the six indicators previously identified. In subsequent analyses, we fit a MIMIC model test the association of the socioeconomic status and demographic characteristics of the participants to the three latent factors. The proposed health measurement model is depicted in Figure 1a. Six health indicators – three observed (lung function, grip strength and chair rise speed) and three self reported (general health status, self reported presence of long lasting illness and an FL summary binary variable) were employed. In accordance with Equation 1, the latent factor "health" represents the proportion of valid health variance present in the health indicators. The latent factor "observed/objective" represents variance due to measurement method induced error (systematic error) attributed to the observed/objective health indicators and the latent factor "self report" represents systematic error attributable to the self report health indicators. With respect to self report indicators, their corresponding method factor reflects bias in the indicators due to factors other than the health status of the participants influencing their- responses.). We note that the three objective indicators we employed represent three different measurement methods, rather than one. A model with a latent method for factor for each of the indicators is not mathematically identified and thus was not estimated. The consequence of this is that this method factor can not be used for evaluation of the measurement methods used, but it can provide us with useful insights on the measurement properties of the objective indicators.

In the proposed model continuous and categorical/ordinal indicators of continuous latent variables are combined. The part of the model where ordinal or binary indicators are linked with the continuous latent variables is a normal ogive item response model, similar to the graded responses model (Samejima, 1969). The part of the model where continuous health indicators are linked with continuous latent factors is a traditional structural equation model with linear regressions between observed and latent variables. We statistically compared the proposed multimethod measurement model with two other competing models: a) a unidimensional model, where a single heath latent factor accounts for variation in all health indicators (Figure 1b); b) a hierarchical model with two first order latent factors, each corresponding to the self report and observed indicators and a second order health factor (Figure 1c).

In the second stage of the analysis, the multimethod measurement model was regressed on external covariates that were used as predictors of valid health as well as method variance, thus extending the measurement model to a Multiple Causes Multiple Indicators (MIMIC) model. All models were estimated with the robust maximum likelihood estimator (MLR) estimator with adaptive (Gauss) quadrature with 15 integration points. The best model selected was the one that returned minimum values for information criteria. We report values for the Akaike's Information Criterion (Nichter et al.), the Bayesian Information Criterion (BIC), and the sample-size adjusted BIC. All analyses were carried out with the Mplus 5 software (Mutthen and Muthen, 1998-2007).

INSERT TABLE 2 ABOUT HERE

Results

Association between objective and self report health indicators

We estimated the associations between the objective and the self report health indicators, with the three self report indicators serving as outcomes. An ordinal logistic regression was employed when health status was the outcome, whereas binary logistic regression was used for chronic illness and functional limitations, with all models being estimated simultaneously. In Table 2 we present the odds ratios from a fully adjusted model including gender, age, social class, wealth, educational status, marital status, housing tenure and equivalised household income. All objective health indicators were significantly positively associated with all self report indicators. Greater grip strength, good lung function and higher speed in chair rises were associated with better self reported health status, the absence of chronic illness and functional limitations, with lung function being the best predictor of the self report indicators. The overlap was not excessive though, as the highest odds ratio (1.22) suggests.

INSERT TABLE 3 ABOUT HERE

Description of selected measurement model

According to the information criteria (see Table 3), the multimethod measurement model was superior to either the unidimensional and two factor models. In Table 2 we present the resulting factor loadings of the best fitting health measurement model on the three latent factors. All health indicators significantly loaded on the health latent factor, and the magnitude of all loadings was satisfactory (>0.47). Self reported functional limitations was the single indicator which loaded most strongly (r = 0.71). The loadings of the remaining health indicators were of roughly similar magnitude (0.47 – 0.52). Grip strength had the highest loading (r = 0.86) of the objective health indicators method factor with 74% of this indicator being due to systematic measurement error. With respect to the self report indicators method factor the general health status item had the highest loading (0.62). High values on this method factor indicate a tendency to report better health (see Graph 1). Conversely low scores suggest a tendency to over report health related problems.

INSERT GRAPH 1 ABOUT HERE

MIMIC model

In Table 5 we report the parameter estimates derived from the MIMIC model.

Health

Gender was significantly associated with the latent health factor (r = -0.61, p<0.001), with men scoring higher compared to women, whereas age had an inverse significant association with health (r =-0.43, p<0.001). Wealth and income were both inversely associated with health, with the least affluent having the lowest score on the health latent variable, with wealth having a stronger overall effect. On the contrary occupational social class was not associated with health. Education had an inverse association with health. Participants without any formal educational qualification scored significantly less on the health latent factor, compared to participants with a degree. Marital status was associated with health, the effect being to due to never married respondents scoring significantly less well than married respondents and to those who were remarried that scored better than those in first marriages. Finally housing tenure was associated with the health latent factor. Owners scored significantly higher than non owners.

INSERT TABLE 4 ABOUT HERE

Self report health indicators method factor

Examining differentials in associations with the self reported health indicators method factor allows us to identify differences in reporting patterns and groups who appear to 'overstate' or 'understate' their true level of health as estimated using the combined method. A high score on this method factor indicates a tendency to report positive health outcomes, for example a tendency to endorse the extremely good response option in the general health question, whereas a low score indicates a tendency to over state health problems (see Graph 1). Gender was significantly associated with this method factor ($r = 0.43$, $p<0.001$), with women scoring higher compared to men. This suggests that men may 'overstate' health related problems or limitations. Age had also a significant association with the self reported health method factor ($r =0.39$, $p<0.001$). Wealth and income were both inversely associated with this method factor, with the least affluent receiving the lowest score, indicating that they tend to over report health problems or limitations. Occupational social class was not associated with this method factor. On the contrary education was inversely associated with the self report method factor, with the least educated scoring less compared to participants with higher degree educational qualifications. Marital status was associated with the self report indicators method factor, the effect being to due to remarried participants scoring significantly more on the method factor, compared to married (only marriage) participants. Furthermore widowed participants scored significantly less compared to married participants. Finally we observed a significant association between tenure and the self report indicators method factor ($r = -0.11$).

INSERT TABLE 5 ABOUT HERE

Observed/Objective health indicators method factor

Gender was significantly associated with the objective health indicators method factor ($r = -0.80$, $p<0.001$), with men scoring higher compared to women. Age had an inverse significant association with the objective health indicators method factor ($r =-0.44$, $p<0.001$). Wealth and income were both inversely associated with this method factor, with the least affluent receiving the lowest score on the latent variable. Occupational social class was not associated with the method factor, whereas education had an inverse association with the objective indicators method factor, with the least educated scoring less compared to participants with higher education. Marital status was also associated with the objective indicators method factor, with the effect being to due to remarried participants scoring significantly more on the method factor, compared to married (only marriage) participants. Furthermore widowed participants scored significantly more compared to married participants. On the contrary, singe/never married participants scored significantly less on this method factor. Finally we observed a significant inverse association between tenure and the objective indicators method factor ($r = 0.04$).

INSERT TABLE 6 ABOUT HERE

Discussion

It was argued recently that the most effective method to decrease controversy over health statistics is to encourage better primary data collection and the development of better analytical methods (Murray, 2007). In this paper we introduce a novel analytic method in the form of a multimethod measurement model for the assessment of health in population based surveys. The model successfully combined information from both self report and objective indicators, which were found to be moderately associated in this sample. We empirically decomposed six health indicators – three based on self report and three observed/objective- into three components: a) valid health variance, b) measurement method induced variance (systematic error) and c) random error. This model was tested against competing measurement models and was found to be superior. Based on the multimethod measurement model we offer a measurement error free latent global index of individual health status as this is reflected by several aspects of health that are captured by the six indicators that could not be assessed by a single indicator alone. Aside form this global health index the outcomes of the model extend to the quantification with latent factors of measurement method induced bias. Another attractive feature of this model is the opportunity to compare health indicators with respect to their reliability in indexing health status. Contrary to what might have been expected, we found that objective indicators were not superior compared to self report indicators and self report of functional limitations was the most reliable health indicator, being the most influenced by the participants' health status, rather than systematic or random error.

A secondary goal of the present study was to use the multimethod health measurement model as an outcome in order to test to what extend the observed health inequalities with respect to socioeconomic status and gender, are due to systematic and/or random measurement error. Simultaneous analysis using the multimethod health measurement model along with covariates (MIMIC model), confirmed the gender gap in the health of older people, with women having less optimal health compared to men, when external bias in the form of method variance and random error is controlled. However, both genders are prone to bias when responding to self rated health questions. Women tend to endorse the positive response options (indicating good health). On the contrary men, tend to endorse the negative response options (indicating ill health). For both males and females these responding tendencies are being influenced by exogenous to their health status factors. Furthermore, older men appear to be prone to 'under-reporting' of health problems when self report measures of health are used, whereas women appear to be prone to over report positive health outcomes. This indicates that the longevity of women despite their worse health can not be attributed to response bias. With respect to the observed health indicators method factor, the finding that men scored higher than women reflects measurement bias possibly due to an excess in physical strength and/or in height that does not necessarily reflect the true health status of the individual, but influence the measurement of grip strength and FVC.

In terms of socioeconomic status, the widely reported health gradients were confirmed. Wealth was the stronger predictor of health compared to social class, income and house tenure, a finding in accordance with previous research (Pollack et al, 2007). Furthermore, education was the strongest predictor of health compared to the other SES indicators, a finding in accordance with previous research (Winkleby, 1992). Confirming previous findings (Bennett, 2006), we found a weak association between marital status and health, with the single and/or never married participants having less optimal health. The observed association between socioeconomic status and the self report method factor provides evidence that self rated health indicators function differentially between socioeconomic groups, a finding in accordance with previous research (Dowd and Zajacova, 2007). We observed a social gradient evident in all the socioeconomic status indices we used. It appears that the disadvantaged participants and the least educated are prone to over reporting health problems compared to their educated counterparts. As previous findings suggest, high socioeconomic status individuals more often use multi-dimensional information when assessing their health compared to those from disadvantaged backgrounds (Calnan and Johnson, 1985). With respect to the observed/objective health indicators method factor, similar socioeconomic gradients were observed, but the effect was considerably weaker. It appears that socioeconomic status is more related to systematic measurement error in self report indicators than in observed/objective indicators.

Future studies are warranted to further test the multimethod health measurement model in population samples that comprise of all age groups, as well as utilising additional health indicators to the ones employed here. The identified latent factors may prove to be very useful tools in studies where health is utilised either as an outcome or a predictor. For example, a comparison of the predictive power on mortality between the error and method free health latent factor and either/or self report and objective health indicators would shed further light on the health – mortality association. The prediction of health services use is another outcome of great public health importance. The latent global health index developed here will provide researchers with a more accurate estimation of the association between general health status and service use compared to estimations using single health indicators as predictors of services use. Another potential use of the latent health global index is in longitudinal designs where health trends over time are of interest, especially since conflicting results with respect to trends have been observed in the past (Barsky, 1988) (Dunnell et al, 1999). Future researchers may also concentrate on the understanding of the nature and the causes of systematic measurement error in health indicators as well as in measurement method evaluation. Strengths of this study include the application for the first time of a multimethod latent variable measurement model on survey based health indicators and the availability of a representative population based sample. A notable limitation, is that in terms of health indicators, the ones employed in this study represent a sample of the possible universe of health indicators (objective/observed or self reported) and future research is needed to test the replicability of our findings. Furthermore the results presented here are based on complete case analysis, that possibly excluded the most frail participants of the ELSA, in which the time to complete five chair rises was not measured. We have estimated our models with partial incomplete data using the full information maximum likelihood method which is available in Mplus 5 and the results we obtained were very similar with results presented here (results form missing data analysis not presented here, available from corresponding author).

References

Adams, S.A., C.E. Matthews, C.B. Ebbeling, C.G. Moore, J.E. Cunningham, J. Fulton, and J.R. Hebert. 2005. "The effect of social desirability and social approval on self-reports of physical activity." *American Journal of Epidemiology* 161(4):389-398.

Banks, J., M. Marmot, Z. Oldfield, and J.P. Smith. 2006. "Disease and disadvantage in the United States and in England." *Jama-Journal of the American Medical Association* 295(17):2037-2045.

Barsky, A.J. 1988. "The paradox of health." *New England Journal of Medicine* 318(7):414-418.

Bennett, K.M. 2006. "Does marital status and marital status change predict physical health in older adults?" *Psychological Medicine* 36(9):1313-1320.

Beresnevaite, M., G.J. Taylor, and R.M. Bagby. 2007. "Assessing alexithymia and type A behavior in coronary heart disease patients: A multimethod approach." *Psychotherapy and Psychosomatics* 76(3):186-192.

Blaxter, M. 1990. *Health and Lifestyles*. London: Routledge.

Boniface, D.R.and M.E. Tefft. 1997. "The application of structural equation modelling to the construction of an index for the measurement of health-related behaviours." *Statistician* 46(4):505-514.

Calnan, M.and B. Johnson. 1985. "Health, health risks and inequalities - an exploratory study of women's perceptions." *Sociology of Health & Illness* 7(1):55-75.

Dowd, J.B.and A. Zajacova. 2007. "Does the predictive power of self-rated health for subsequent mortality risk vary by socioeconomic status in the US?" *International Journal of Epidemiology* 36:1214-1221.

Dunnell, K., J. Fitzpatrick, and J. Bunting. 1999. "Making use of official statistics in research on gender and health status: recent British data." *Social Science & Medicine* 48(1):117-127.

Eid, M., T. Lischetzke, and F. Nussbeck. 2006. "Structural Equation models for multitrait-multimethod data." in *Handbook of multimethod measurement in psychology*, edited by M. Eid and E. Diener. Washington DC: American Psychological Association.

Elam, J.T., M.J. Graney, T. Beaver, D. Elderwi, W.B. Applegate, and S.T. Miller. 1991. "Comparison of subjective ratings of function with observed functional ability of frail older persons." *American Journal of Public Health* 81(9):1127-1130.

Ferrer, M., R. Lamarca, F. Orfila, and J. Alonso. 1999. "Comparison of performance-based and self-rated functional capacity in Spanish elderly." *American Journal of Epidemiology* 149(3):228-235.

Grundy, E.and A. Sloggett. 2003. "Health inequalities in the older population: the role of personal capital, social resources and socio-economic circumstances." *Social Science & Medicine* 56(5):935-947.

Hebert, J.R., K.E. Peterson, T.G. Hurley, A.M. Stoddard, N. Cohen, A.E. Field, and G. Sorensen. 2001. "The effect of social desirability trait on self-reported dietary measures among multi-ethnic female health center employees." *Annals of Epidemiology* 11(6):417-427.

Huisman, M., A.E. Kunst, and J.P. Mackenbach. 2003. "Socioeconomic inequalities in morbidity among the elderly; a European overview." *Social Science & Medicine* 57(5):861-873.
Lillard, L.and M. Swaminathan. 2000. "Health in Bangladesh, Indonesia and the United States: A MIMC Model with Threshold Responses and Person, Family and
Community Components." in *Population Association of America annual meeting*. Los Angeles.
Majdandzic, M.and D.C. van den Boom. 2007. "Multimethod longitudinal assessment of temperament in early childhood." *Journal of Personality* 75(1):121-167.
Marmot M, B.J., Blundell R, Lessof C and Nazroo J. 2002. "Health, wealth and lifestyles of the older population in England: THE 2002 ENGLISH LONGITUDINAL STUDY OF AGEING." Institute for Fiscal Studies.
Murray, C.J.L. 2007. "Health statistics 2 - Towards good practice for health statistics: lessons from the Millennium Development Goal health indicators." *Lancet* 369(9564):862-873.
Mutthen, L.K.and B.O. Muthen. 1998-2007. *Mplus User's Guide. Fourth Edition*. Los Angeles, CA.
Myers, A.M., P.J. Holliday, K.A. Harvey, and K.S. Hutchinson. 1993. "Functional performance measures - are they superior to self assessments." *Journals of Gerontology* 48(5):M196-M206.
Nichter, M., C. Ritenbaugh, N. Vuckvoic, and M. Aickin. 1995. "Dieting and watching behaviours among adolescent females - report of a multimethod study." *Journal of Adolescent Health* 17(3):153-162.
Palmer, R.F., J.H. Dwyer, and N. Semmer. 1994. "A measurement model of adolescent smoking." *Addictive Behaviors* 19(5):477-489.
Pollack, C.E., S. Chideya, C. Cubbin, B. Williams, M. Dekker, and P. Braveman. 2007. "Should health studies measure wealth? A systematic review." *American Journal of Preventive Medicine* 33:250-264.
Prinstein, M.J., M.K. Nock, A. Spirito, and W.L. Grapentine. 2001. "Multimethod assessment of suicidality in adolescent psychiatric inpatients: Preliminary results." *Journal of the American Academy of Child and Adolescent Psychiatry* 40(9):1053-1061.
Robine, J.M., C. Jagger, and R.G. Euro. 2003. "Creating a coherent set of indicators to monitor health across Europe - The Euro-REVES 2 project." *European Journal of Public Health* 13:6-14.
Robine, J.M., J.P. Michel, and L.G. Branch. 1992. "Measurement and utilization of healthy life expectancy - conceptual issues." *Bulletin of the World Health Organization* 70(6):791-800.
Samejima, F. 1969. "Estimation of latent trait ability using a response pattern of graded scores." *Psychometrika Monograph Supplement* 17.
Svanborg, A. 1988. "Practical and functional consequenses of aging." *Gerontology* 34:11-15.
Tram, J.M.and D.A. Cole. 2006. "A multimethod examination of the stability of depressive symptoms in childhood and adolescence." *Journal of Abnormal Psychology* 115(4):674-686.

Winkleby, M.A., D.E. Jatulis, E. Frank, and S.P. Fortmann. 1992. "Socioeconomic status and health - how education, income and occupation contribute to risk factors for cardiovasular disease." *American Journal of Public Health* 82(6):816-820.
Wolf, D.A., K. Hunt, and J. Knickman. 2005. "Perspectives on the recent decline in disability at older ages." *Milbank Quarterly* 83(3):365-395.

Table 1. Descriptive statistics of the six health indicators

		f	%
Functional limitations			
FL problems reported		5258	59.9
No limitations		3522	40.1
Self rated general health			
Poor		679	7.8
Fair		1794	20.7
Good		2755	31.7
Very good		2391	27.5
Excellent		1063	12.2
Self reported long standing (chronic) illness			
Yes		5062	57.7
No		3713	42.3
Objective Indicators	**Mean**		**Standard deviation**
Grip strength (N = 7508)	29.18		11.51
Chair rise (N= 6368)	52.28		4.33
Lung function (N = 7025)	3.15		1.10

Table 2 Descriptive statistics of co-variates

	f	%
Gender		
Male	3949	45.0
Female	4831	55.0
Wealth		
Wealth 1st quintile (highest)	1713	19.8
Wealth 2nd quintile	1736	20.1
Wealth 3d quintile	1727	20.0
Wealth 4th quintile	1751	20.2
Wealth 5th quintile (lowest)	1724	19.9
Social class		
Social class 1 (highest)	2415	28.4
Social class 2	1533	18.0
Social class 3	859	10.1
Social class 4	968	11.4
Social class 5 (lowest)	2736	32.1
Income		
Income 1st quintile (highest)	1643	18.7
Income 2nd quintile	1697	19.3
Income 3d quintile	1739	19.8
Income 4th quintile	1785	20.3
Income 5th quintile (lowest)	1916	21.8
Marital status		
Married (first marriage)	4816	54.9
Remarried	921	10.5
Never married	455	5.2
Legally separated/divorced	934	10.6
Widowed	1653	18.8
Education		
Degree/Higher education	2093	23.9
A level	575	6.5
O level/CSE grade	1870	21.3
Foreign/other	764	8.7
No qualifications	3468	39.5
Tenure		
Owners	7168	81.8
Tenants etc	1595	18.2

Table 3. Fully adjusted odds ratios of objective and self report indicators association

	Health status	*Absence of chronic illness*	*No Functional limitations*
Grip strength	**1.02***	**1.02**	**1.02**
Chair rise	**1.10**	**1.08**	**1.13**
Lung function	**1.22**	**1.15**	**1.16**

*Highlighted odd ratios are significant (p<0.05). Associations with health status odds ratios derived from ordinal logistic regression, all other odds ratios from binary logistic regression

Table 4. Information criteria for the three competing measurement models (smaller values indicate better fit)

	AIC	*BIC*	*ssaBIC*
Unidimensional model	126336.477	126456.963	126399.764
Hierarchical model	124695.040	124828.913	124765.359
Multimethod model	124079.648	124246.989	124167.546

Table 5. Standardised factor loadings of health indicators to latent health, self report method and objective method factors. Percentage of variance in parentheses

	Health	*Self report indicators method factor*	*Observed indicators method factor*	*Random error*
Grip strength	**0.51** (26%)		**0.86** (74%)	4%
Chair rise	**0.52** (27%)		-0.03 (1%)	72%
Lung function	**0.50** (25%)		**0.42** (17%)	68%
Functional limitations	**0.71** (50%)	**0.22** (8%)		42%
General health	**0.52** (27%)	**0.62** (40%)		33%
Long standing (chronic) illness	**0.47** (23%)	**0.50** (27%)		50%

* highlighted parameters are significant

Table 6. MIMIC model standardised parameters

	Health	Self report indicators method factor	Observed indicators method factor
Gender	**-0.61**	**0.42**	**-0.80**
Age	**-0.43**	**0.39**	**-0.44**
Wealth			
Wealth 1st quintile (highest)	0	0	0
Wealth 2nd quintile	-0.01	**-0.15**	-0.01
Wealth 3d quintile	-0.03	**-0.10**	**-0.03**
Wealth 4th quintile	**-0.05**	**-0.16**	**-0.05**
Wealth 5th quintile (lowest)	**-0.06**	**-0.32**	**-0.06**
Social class			
Social class 1 (highest)	0	0	0
Social class 2	-0.01	-0.03	-0.01
Social class 3	0.01	-0.01	0.01
Social class 4	0.01	0.03	-0.02
Social class 5 (lowest)	-0.02	0.06	-0.01
Income			
Income 1st quintile (highest)	0	0	0
Income 2nd quintile	**-0.01**	**-0.12**	-0.01
Income 3d quintile	**-0.03**	**-0.13**	**-0.03**
Income 4th quintile	**-0.03**	**-0.19**	**-0.04**
Income 5th quintile (lowest)	**-0.04**	**-0.21**	-0.01
Marital status			
Married (first marriage)	0	0	0
Remarried	**0.04**	**0.08**	**0.03**
Never married	**-0.02**	0.03	**-0.02**
Legally separated/divorced	0.01	0.01	-0.01
Widowed	-0.01	**-0.09**	**0.02**
Education			
Degree/Higher education	0	0	0
A level	**-0.03**	**-0.08**	**-0.03**
O level/CSE grade	**-0.04**	**-0.07**	**-0.04**
Foreign/other	**-0.05**	**-0.11**	**-0.03**
No qualifications	**-0.08**	**-0.15**	**-0.07**
Tenure (owners vs others)	**-0.04**	**-0.11**	**-0.04**

* highlighted parameters are significant

Figure 1a. Multimethod health measurement model

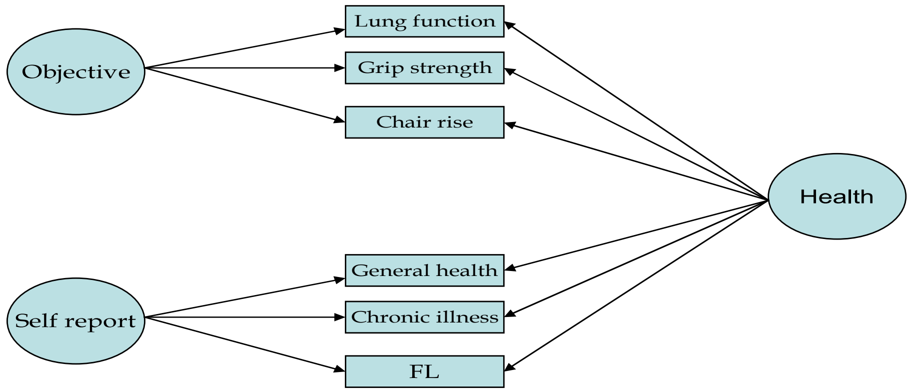


Figure 1b. Unidimensional measurement model

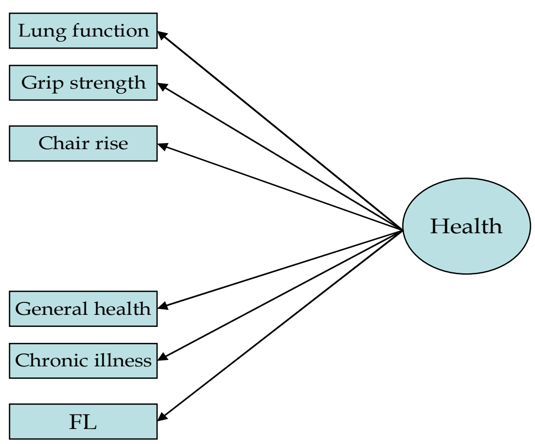
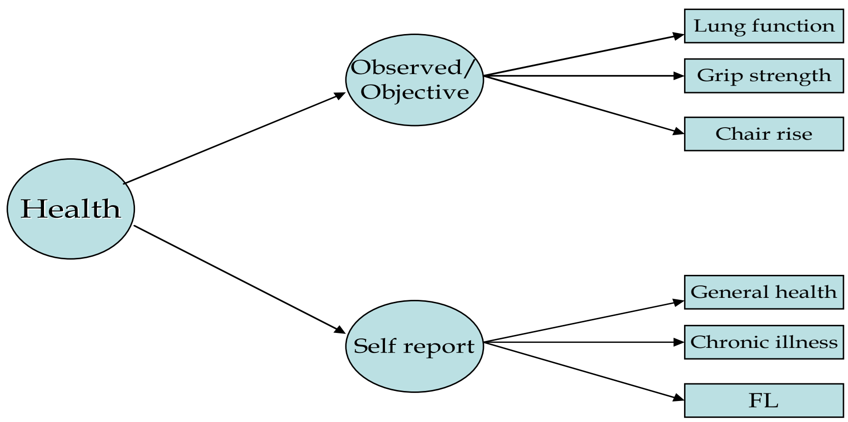


Figure1c. Hierarchical measurement model



Graph 1. Self report method factor means with respect to self reported general health status