Published in final edited form as:
J Nutr Health Aging. 2009 January ; 13(1): 57–62.

The SF-36: a simple, effective measure of mobility-disability for epidemiological studies

Holly E Syddall, MSc[1], **Helen J Martin, BSc**[1], **Rowan H Harwood, MSc MD FRCP**[2], **Cyrus Cooper, MA DM FRCP**[1], and **Avan Aihie Sayer, MSc PhD FRCP**[1,3]

1*MRC Epidemiology Resource Centre, University of Southampton, UK*

2*Health Services for Older People, Nottingham City Hospital, Nottingham, UK*

3*University Geriatric Medicine, University of Southampton, UK*

Abstract

Background—Mobility disability is a major problem in older people. Numerous scales exist for the measurement of disability but often these do not permit comparisons between study groups. The physical functioning (PF) domain of the established and widely used Short Form-36 (SF-36) questionnaire asks about limitations on ten mobility activities.

Objectives—To describe prevalence of mobility disability in an elderly population, investigate the validity of the SF-36 PF score as a measure of mobility disability, and to establish age and sex specific norms for the PF score.

Methods—We explored relationships between the SF-36 PF score and objectively measured physical performance variables among 349 men and 280 women, 59-72 years of age, who participated in the Hertfordshire Cohort Study (HCS). Normative data were derived from the Health Survey for England (HSE) 1996.

Results—32% of men and 46% of women had at least some limitation in PF scale items. Poor SF-36 PF scores (lowest fifth of the gender-specific distribution) were related to: lower grip strength; longer timed-up-and-go, 3m walk, and chair rises test times in men and women; and lower quadriceps peak torque in women but not men. HSE normative data showed that median PF scores declined with increasing age in men and women.

Conclusion—Our results are consistent with the SF-36 PF score being a valid measure of mobility disability in epidemiological studies. This approach might be a first step towards enabling simple comparisons of prevalence of mobility disability between different studies of older people. The SF-36 PF score could usefully complement existing detailed schemes for classification of disability and it now requires validation against them.

Keywords

epidemiology; physical functioning; disability; ageing; mobility

Key Points

- Mobility disability is a major problem in older people but its prevalence cannot be easily compared between study groups or populations.

Correspondence and reprint requests to: Mrs Holly E Syddall, MRC Epidemiology Resource Centre, University of Southampton, Southampton General Hospital, Southampton SO16 6YD, UK, Tel: +44 (0)23 8077 7624, Fax: +44 (0)23 8070 4021, Email: hes@mrc.soton.ac.uk.

UKPMC Funders Group Author Manuscript

- The physical functioning (PF) domain of the Short Form-36 (SF-36) questionnaire could serve as a measure of mobility disability which is easily compared across epidemiological studies.
- In the Hertfordshire Cohort Study, poor SF-36 PF scores were related to poorer physical performance (lower grip strength, longer timed-up-and-go, 3m walk, and chair rises test times, and lower quadriceps peak torque) consistent with the SF-36 PF score being a valid measure of mobility disability.
- The SF-36 PF score could usefully complement existing detailed schemes for classification of disability and it now requires validation against them.

Introduction

Mobility disability (locomotor activity limitation) is a major problem in older people and its prevalence is rising.[1,2] It is defined as the inability to perform locomotor tasks in a normal manner. Numerous schemes exist for the classification of functioning and disability. The World Health Organisation developed the International Classification of Impairments, Disabilities and Handicaps defining these concepts and linking them together conceptually, in an attempt to define an international standard. This has recently been revised as the International Classification of Functioning, disability and health (ICF)[3] which provides a comprehensive framework for classification. However, it is not easily operationalised as a tool for use in epidemiological research.[4] There are many empirically developed disability scales, including the Barthel Activities of Daily Living Index[5], the Townsend's Disability Scale[6] and the Office of Population Censuses and Surveys (OPCS) Disability Scale[1] but the disparate classification systems do not permit comparison of disability levels between study groups.[7, 8]

The Short Form-36 (SF-36) questionnaire is an established and widely used health-related quality of life measure (HRQL).[9,10] It has been used extensively in observational and randomised studies for a range of illness conditions and validated across a range of ages and participant characteristics.[11,12] Its use extends beyond people with specific disease states, to determine HRQL in populations.[10,13] The SF-36 comprises eight domain scores including 'physical functioning'. This domain asks respondents to report limitations on ten mobility activities, such as walking specified distances, carrying groceries and bathing or dressing. This raises the possibility that an established and widely-used instrument can be mapped to the ICF, and validated as a standard measure of mobility disability. The association between poor physical performance using a variety of objective performance measures, and an increased risk of mobility disability, has been well researched.[14,15]

Our objective was to assess the construct validity of the SF-36 PF score as a measure of mobility disability by exploring its relationship with a panel of physical performance measures. We also present normative data for this score.

Methods

We investigated the performance of the SF-36 PF domain as a measure of mobility disability using data from the Hertfordshire Cohort Study (HCS) of men and women 59-72 years of age. The study has been described previously.[16] In brief, from 1911-1948, midwives collected detailed records on infants born in Hertfordshire, UK. In 1998, the NHS central registry used these records to trace 3822 men and 3284 women born in Hertfordshire between 1931-1939 and still living in the county and registered with a General Practitioner. Fieldwork was conducted over a period of five years phased by region of Hertfordshire (East, North, West).

Initial findings were based on participants from East Hertfordshire[17,18]; the present study is based on participants from West Hertfordshire who underwent a range of assessments of physical performance.

Permission to contact 792 men and 647 women resident in West Hertfordshire was obtained from their General Practitioners. 396 (50%) men and 321 (50%) women agreed to participate in a home interview where trained nurses collected information on medical and social history including the SF-36 questionnaire[19], self-reported walking speed (taken from questions asking about the number of minutes spent walking outdoors in the previous day)[20], alcohol intake, smoking habit and social class (classified using the 1990 OPCS Standard Occupational Classification scheme). 349 (88%) men and 280 (87%) women subsequently attended a clinic for a number of investigations including measurement of height and weight[21] and physical performance. The battery of physical performance assessments comprised: hand grip strength measured three times on each side using a Jamar hydraulic dynamometer (Promedics, Blackburn, U.K.)[22]; timed up and go test (ability to rise from a seated position, walk 3m, turn, and walk back returning to a seated position, at customary speed)[23]; 3m customary walking test[24]; chair rises test(ability to rise and sit 5 times as quickly as possible)[25]; timed one legged balance test with eyes open up to a maximum of 30 seconds[26]; and supine quadriceps torque measured twice on each side using a novel portable hand held dynamometer (HHD) (Lafayette Instruments Company Inc, Leicestershire U.K.).[27] Intra- and inter-observer studies were carried out regularly during the fieldwork to ensure comparability of measurements within and between observers. The study had ethical approval from the Hertfordshire and Bedfordshire Local Research Ethics Committee and all subjects gave written informed consent.

Statistical methods

SF-36 data were mapped to eight domain scores, including the physical functioning domain (PF), using the published coding algorithms and imputation methods for missing items.[19] The best of the six grip measurements was used to characterise maximum muscle strength. Weight and chair rises variables were transformed to normal distributions using the $\log_e$ transformation (geometric means and standard deviations, and proportional differences between groups, were therefore presented for these variables). Height and weight were highly correlated ($r=0.46$, $p<0.001$ for men; $r=0.29$, $p<0.001$ for women); to avoid multi-collinearity problems in modeling analyses we calculated a sex-specific standardised residual of weight-adjusted-for-height.

Variables were summarised using means and standard deviations, medians and inter-quartile ranges, and frequencies and percentage distributions. Prevalence of mobility disability was described using each of the 10 individual items comprising the PF domain and the overall PF score.

The internal validity of the PF domain was investigated using Cronbach's alpha. The construct validity of the PF score as a disability measure was investigated by using linear and logistic regression models to explore the relationships between "low/poor" SF-36 physical functioning scores (defined as scores in the lowest sex-specific fifth of the distribution i.e. ≤ 60 for men, and ≤ 75 for women) and the objective measures of physical performance. We hypothesised that lower PF scores would be associated with poorer objective physical performance.[14,15] Linear regression was used to measure associations between the continuously distributed physical performance variables (dependent variables) and the PF score (independent variable), and to derive estimates of the mean difference in dependent variable values between low and high PF groups. Logistic regression was used for the categorical balance time variable and yielded odds ratios for achieving maximal balance time between low and high PF groups. Analyses were conducted without and with adjustment for the potential confounding influences

UKPMC Funders Group Author Manuscript

of age, height, weight-adjusted-for-height, walking speed, social class, smoking habit and alcohol intake.

As a separate exercise, normative summary statistics for the SF-36 PF score were produced by gender and five year age-bands using data from the Department of Health's large and nationally representative Health Survey for England (HSE).[10] In 1996 this survey included the SF-36, from which we re-analysed the PF scores. This was accomplished by accessing the 1996 HSE dataset[28] from the ESRC UK data archive (www.data-archive.ac.uk). The HCS was too small to provide these data itself, but the availability of nationally representative norms for the PF domain adds considerably to its usefulness as an epidemiological tool.

All analyses were carried out for men and women separately using Stata, release 8.0 (Stata Corporation 2003).

Results

The SF-36 questionnaire was nurse-administered in HCS; it was acceptable to all study participants and the rate of missing data was extremely low. Only 16 (4.6%) men and 10 (3.6%) women had missing data for any of the questionnaire items, but all had sufficient data to enable computation of the eight SF-36 domain scores using the imputation approach recommended in the SF-36 manual[19]. The internal validity of the 10-item physical functioning domain was high for both men (Cronbach's alpha = 0.89) and women (alpha=0.90).

Women were more likely than men to report that their health limited them a lot in their performance of activities e.g. 43% of women and 27% of men reported that their health currently limited them a lot in their performance of vigorous activities (odds ratio 2.00 [95% CI 1.43, 2.80], p<0.001 - table 1) and 46% of women and 32% of men were limited a lot in their performance of at least one of the PF scale items (odds ratio 1.77 [95%CI 1.28, 2.45], p=0.001 - table 1). Correspondingly, women also had poorer overall PF scores (median and IQR for men 90[80,95] and 85[65,95] for women, p<0.0001 for gender difference). Given that this analysis was conducted on the West Hertfordshire sub sample of men and women who were participating in the countywide Hertfordshire Cohort Study, we compared their characteristics with men and women from East and North Hertfordshire. There were no significant differences between study participants from West compared with East and North Hertfordshire in terms of social class, smoking, alcohol, walking speed, SF-36 scores, individual SF-36 PF items, and height or weight (data not shown), although West Hertfordshire participants were slightly older owing to the phasing of the fieldwork (average ages of: East and North men vs West men, 65.1 (sd 2.7, n=1230) vs 67.8 (sd 2.5, n=349), p<0.0001; East and North women vs West women, 66.3 (sd 2.6, n=1138) vs 68.1 (sd 2.5, n=280), p<0.0001). We have previously shown[16] that participants in the East phase of the HCS were broadly comparable with those in the nationally representative Health Survey for England (in terms of their social class, smoking, alcohol, anthropometry and cardiovascular disease) although HCS participants had somewhat better SF-36 self-reported general health than their counterparts in the Health Survey for England. We have also previously analysed response bias patterns within HCS[16] and, unsurprisingly, identified a "healthy" or "health aware" responder bias, with men and women who progressed from the home interview to clinic tending to smoke less, and have better self-reported function and general health than those only participating in the home interview.

Older age was associated with a poorer PF score in men; the odds of having a poor PF score were increased 13% (95%CI [2%, 26%], p=0.02) per year increase in chronological age. The relationship between age and PF score in women was statistically uncertain (odds ratio for poor PF score per year increase in age: 1.04 (95%CI [0.93, 1.16], p=0.52)

Poorer SF-36 PF scores were related to worse outcomes across the range of physical performance tests in men and women. Specifically, poor SF-36 PF scores were associated with lower grip strength, longer timed-up-and-go, 3m walk and chair rises test times in men and women, and lower quadriceps peak torque in women, with and without adjustment for age, height, weight for height, walking speed, social class, smoking habit and alcohol consumption (Table 2). There was no relationship between SF-36 PF score and quadriceps peak torque in men. Poor SF-36 scores were also associated with diminished performance on the timed one-legged balance test in men and women, although these relationships were removed by adjustment (Table 2).

Table 3 presents normative data for the SF-36 PF score from older men and women who participated in the nationally representative Health Survey for England (1996). Percentiles of the PF score are presented by five year age-bands and show, in a dataset that includes subjects of a wide age range, that PF scores decline with increasing age in both men and women.

Discussion

We have demonstrated that the SF-36 questionnaire was acceptable to participants in the Hertfordshire Cohort Study. The rate of missing items was extremely low and, consistent with previous studies[12,19], the internal-validity of the SF-36 physical functioning (PF) domain was high. SF-36 PF scores were lower in women than men, and declined with older age among men. SF-36 PF scores were related to physical performance across a range of tests. Specifically, poor SF-36 PF scores were associated with lower grip strength, longer timed-up-and-go, 3m walk and chair rises test times in men and women, and lower quadriceps peak torque in women. We therefore propose that the physical functioning domain of the SF-36 questionnaire could be a useful instrument for the assessment of mobility disability in epidemiological studies.

We have presented summary statistics for the SF-36 PF score from the Health Survey for England; these data provide a nationally representative source of age- and gender-specific norms for an SF-36 PF measure of mobility disability. Age- and gender-specific definitions of mobility disability, based on thresholds derived from these data, could enable a consistent definition of mobility disability to be used in all studies which included SF-36 PF data. Such an approach would facilitate comparison of the prevalence of mobility disability between study populations; this is not currently possible owing to the wide range of classification schemes for functioning and disability.[1,3,5,6]

Although the SF-36 has been used extensively as a HRQL outcome measure[9-13], only one study to date has used the SF-36 as a marker of disability. In conjunction with official life table data, Bobak *et al* used the self-rated health and PF domains of the SF-36 to compare disability levels between middle aged and elderly people in Russia and Sweden.[29] However, this study did not propose that specific domains of the SF-36 might serve as widely useful measures of disability. Kuh *et al*[30] explored the relationships between objective measures of physical performance and a questionnaire-based assessment of functional limitations (difficulties walking, stair climbing, gripping and falls), and obtained broadly similar results to those presented in this study, however their functional limitations marker was not based on the SF-36 questionnaire. To our knowledge, no previous studies have proposed, and assessed the validity, of the SF-36 PF score as a marker of mobility disability.

This study had some limitations. Firstly, we found no relationship between SF-36 PF score and quadriceps peak torque in men. This may be due to limitations in the methodology used to measure quadriceps strength which demanded that the observer was able to resist the force exerted by the study participant.[27] It is possible that men, who were stronger than women on average, overpowered the observer and that consequently these data were less discriminating as a measure of quadriceps strength among men. Secondly, older age was associated with poorer SF-36 PF scores in Hertfordshire men, but not women. The absence of a relationship for women could be a reflection of the narrow age range of participants in HCS. Normative data from the HSE, which studied men and women of a wide age range, clearly show a strong relationship between increasing age and poorer SF-36 PF scores for men and women (Table 3). Thirdly, we could only assess the internal and construct validity of the SF-36 PF score as a marker of mobility disability. However, the SF-36 questionnaire has been extensively validated in previous work, and its repeatability has also been shown to be good, even if these studies did not explicitly assess the repeatability of the PF score as a marker of mobility disability.[12,19] Fourthly, the population contained individuals of a range of abilities, but was not especially frail, as reflected by the better SF-36 self-reported general health among HCS participants in comparison with their counterparts in the Health Survey for England [16] and the presence of a “healthy” or “health aware” responder bias within HCS [16]. The SF36 has previously been shown to be difficult for frailer groups (in self administration format at least), and was primarily designed for more well/less frail populations. However, it is important to have scales that measure at this end of the ability spectrum as well. Transitions from able to unable need to be detected, and any preventative or therapeutic intervention, especially those operating at the population level, are likely to be most effective in this ability range. The population was, however, representative of a home-dwelling ‘young elderly’ population. Lastly, we could only utilise data from the West phase of HCS. However we have shown that participants in the West phase of the study were comparable with those in East and North Hertfordshire, and have previously shown that participants in the East phase of HCS were broadly comparable with those in the nationally representative Health Survey for England [16].

Our study also had many strengths. Firstly, the SF-36 is an established and widely used questionnaire which has been extensively validated as a measure of health-related quality of life in a range of settings.[12,19] Secondly, we were able to assess the construct validity of the SF-36 PF score as a measure of mobility disability using a wide range of directly measured physical performance items. All of these were measured according to strict protocols by trained research nurses and doctors who completed inter- and intra-observer studies throughout the HCS fieldwork. Practical and time constraints in our clinics meant that we were unable to measure an even wider range of physical performance measures, but it would clearly be of interest to replicate our findings in other studies, and to widen the battery of objective measures where possible. Thirdly, we are confident that our results are generalisable to the wider population of older men and women in England, because the cohort have been shown to be broadly comparable with participants in the nationally representative HSE.[16] Lastly, rather than presenting SF-36 PF normative data from HCS, we have produced normative statistics using data from older men and women who participated in the nationally representative HSE in 1996 (the only year that the SF-36 was included in the survey). We elected not to use previously published sources of UK SF-36 normative data[12,31] because: these norms were derived from local studies which were not representative of England and Wales; the sample sizes for older people were relatively small; appropriate summary statistics for the skewed SF-36 PF score[32] were not presented; and normative data were not provided by gender and age group.

In conclusion, our results are consistent with the SF-36 PF score being a valid marker of mobility disability in epidemiological studies of older people. This approach might be a first step towards enabling simple comparisons of prevalence of mobility disability between different studies. We do not suggest that the SF-36 PF score could replace more detailed disability scales; consideration of the individual components of the SF-36 PF domain clearly shows that it could only serve as a direct marker of mobility disability. Many other aspects of disablement such as loss of cognition or vision would need to be captured by more detailed disability scales. However, the SF-36 PF score could usefully complement existing detailed schemes for classification of disability and it now requires validation against them.

Acknowledgements

We thank the men and women who participated in the Hertfordshire study, the Hertfordshire General Practitioners, and the nurses and doctors who conducted the fieldwork.

We acknowledge: the Joint Health Surveys Unit of Social and Community Planning Research and University College London, Department of Epidemiology and Public Health, as the original creators of the HSE 1996 dataset; the Social and Community Planning Research group as the depositor of this dataset with the UK data Archive; and the Department of Health as the original source of funding for the 1996 HSE. These organisations bear no responsibility for our analysis and interpretation of the 1996 HSE Dataset.

Funding

The Medical Research Council was the principal source of funding, but this multi-stakeholder study has also received grants from the British Heart Foundation, Arthritis Research Campaign, National Osteoporosis Society, Wellcome Trust, and University of Southampton.

References

1. Ebrahim S, Wannamethee SG, Whincup P, et al. Locomotor disability in a cohort of British men: the impact of lifestyle and disease. Int J Epidemiol 2000;29:478–486. [PubMed: 10869320]
2. Harwood RH, Prince MJ, Mann AH, et al. Associations between diagnoses, impairments, disabilities and handicaps in a population of elderly people. Int J Epidemiol 1998;27:261–268. [PubMed: 9602408]
3. International classification of functioning, disability and health. WHO Health Organisation; 2001.
4. Ebrahim S. Clinical and public health perspectives and applications of health-related quality of life measurement. Soc Sci.Med 1995;41:1383–94. [PubMed: 8560306]
5. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. Md State Med J 1965;14:61–5. [PubMed: 14258950]
6. Bond, J.; Carstairs, V. Services for the elderly. Scottish Home and Health Department. Scottish Health Service Studies; 1982.
7. Harwood RH, Sayer AA, Hirschfeld M. Current and future worldwide prevalence of dependency, its relationship to total population, and dependency ratios. Bull World Health Organ Apr;2004 82(4):251–8. [PubMed: 15259253]
8. Harwood, RH.; Ebrahim, S.; Kalache, A. Locomotor Disability in The epidemiology of Old Age. London: BMJ/WHO; 1996. p. 378-388.
9. Ware JE Jr. Gandek B. Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) Project. J Clin Epidemiol 1998;51:903–12. [PubMed: 9817107]
10. Prescott-Clarke, P.; Primatesta, P. Health Survey for England ‘96. Vol I: Findings and Vol II: Methodology and Documentation. London: HMSO; 1998.
11. Ferrucci L, Guralnik JM, Studenski S, et al. Designing randomized, controlled trials aimed at preventing or delaying functional decline and disability in frail, older persons: a consensus report. J Am Geriatr.Soc 2004;52:625–34. [PubMed: 15066083]
12. Brazier JE, Harper R, Jones NM, et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. BMJ 1992;305:160–4. [PubMed: 1285753]
13. Singh JA, Borowsky SJ, Nugent S, et al. Health-related quality of life, functional impairment, and healthcare utilization by veterans: veterans’ quality of life study. J Am Geriatr.Soc 2005;53:108–13. [PubMed: 15667386]
14. Guralnik JM, Ferrucci L, Pieper CF, 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. J Gerontol A Biol Sci Med Sci 2000;55:M221–M231. [PubMed: 10811152]

15. Ostir GV, Markides KS, Black SA, et al. Lower body functioning as a predictor of subsequent disability among older Mexican Americans. J Gerontol A Biol Sci Med Sci 1998;53:M491–M495. [PubMed: 9823755]
16. Syddall HE, Sayer AA, Dennison EM, et al. Cohort Profile: The Hertfordshire Cohort Study. Int J Epidemiol. 2005
17. Sayer AA, Syddall HE, Gilbody HJ, et al. Does sarcopenia originate in early life? Findings from the Hertfordshire Cohort Study. J Gerontol 2004;59A:930–4.
18. Sayer AA, Syddall HE, Dennison EM, et al. Birth weight, weight at one year and body composition in older men: findings from the Hertfordshire Cohort Study. Am J Clin Nutr 2004;80:199–203. [PubMed: 15213049]
19. Ware, JE.; Kosinski, M.; Gandek, B. SF-36 Health Survey: Manual and Interpretation Guide. Lincoln, RI: Quality Metric Incorporated; 2000.
20. Dallosso HM, Morgan K, Bassey EJ, et al. Levels of customary physical activity among the old and the very old living at home. J Epidemiol Community Health 1988;42:121–7. [PubMed: 3221161]
21. Lohman, TG.; Roche, AF.; Martorell, R., editors. Anthropometric standardization reference manual. Champaign, Illinois: Human Kinetic Books; 1988.
22. Mathiowetz V, Weber K, Volland G, et al. Reliability and validity of grip and pinch strength evaluations. J Hand Surg [Am] 1984;9:222–6.
23. Podsiadlo D, Richardson S. The timed "up & go": A test of basic functional mobility for frail elderly persons. Journal of American Geriatric Society 1991;39:142–148.
24. Tinetti ME, Inouye SK, Gill TM, et al. Shared risk factors for falls, incontinence, and functional dependence. Unifying the approach to geriatric syndromes. JAMA 1995;273:1348–53. [PubMed: 7715059]
25. Bassey EJ, Fiatarone MA, O'Neill EF, et al. Leg extensor power and functional performance in very old men and women. *Clin Sci* (Lond) 1992;(3):321–327. [PubMed: 1312417]
26. Briggs RC, Gossman MR, Birch R, et al. Balance performance among noninstitutionalized elderly women. Phys.Ther 1989;69(9):748–756. [PubMed: 2772037]
27. Martin HJ, Yule V, Syddall HE, et al. Is hand held dynamometry useful for the measurement of quadriceps strength in older people? A comparison with gold standard Biodex dynamometry. Gerontology 2006;52:154–9. [PubMed: 16645295]
28. Joint Health Surveys Unit of Social and Community Planning Research. University College London. Health Survey for England. Vol. 3rd ed.. Colchester, Essex: UK Data Archive; Mar 26. 2001 1996 [computer file]distributorSN: 3886
29. Bobak M, Kristenson M, Pikhart H, et al. Life span and disability: a cross sectional comparison of Russian and Swedish community based data. BMJ 2004;329(7469):767. [PubMed: 15377571]
30. Kuh D, Bassey EJ, Butterworth S, et al. Grip strength, postural control, and functional leg power in a representative cohort of British men and women: associations with physical activity, health status, and socioeconomic conditions. J Gerontol A Biol Sci Med Sci Feb;2005 60(2):224–31. [PubMed: 15814867]
31. Jenkinson C, Coulter A, Wright L. Short form 36 (SF36) health survey questionnaire: normative data for adults of working age. BMJ 1993;306:1437–40. [PubMed: 8518639]
32. Julious SA, George S, Campbell MJ. Sample sizes for studies using the short form 36 (SF-36). JECH 1995;49:642–644.

UKPMC Funders Group Author Manuscript

Table 1

Characteristics of study participants

Mean and SD unless stated otherwise	Men (n=349)	Women (n=280)
Age in (years)	67.8 (2.5)	68.1 (2.5)
Weight in (kg)[a]	81.2 (1.2)	70.5 (1.2)
Height in (cm)	174.1 (6.6)	160.9 (6.0)
N(%) Current manual social class (IIIM-V) [b]	202 (57.9)	169 (60.4)
N(%) Current smoker	40 (11.5)	25 (9.0)
N(%) Men >21/Women >14 units alcohol per week	66 (18.9)	13 (4.6)
N(%) Walking speed: very slow/stroll at easy pace	91 (26.1)	69 (24.6)
N(%) Walking speed: normal	146 (41.8)	129 (46.1)
N(%) Walking speed: fairly brisk/fast	112 (32.1)	82 (23.5)
SF-36 Physical Functioning score (median [IQR][c])	90 (85,95)	85 (65,95)
Individual SF-36 physical functioning (PF) items		
N (%) limited a lot by health in performance of;		
Vigorous activities	94 (26.9)	119 (42.5)
Moderate activities	12 (3.4)	16 (5.7)
Lifting/carrying groceries	15 (4.3)	33 (11.8)
Climbing several flights of stairs	18 (5.2)	45 (16.1)
Climbing one flight of stairs	8 (2.3)	7 (2.5)
Bending, kneeling, stooping	31 (8.9)	46 (16.4)
Walking >1 mile	27 (7.7)	41 (14.6)
Walking half a mile	15 (4.3)	23 (8.2)
Walking 100 yards	9 (2.6)	4 (1.4)
Bathing/dressing oneself	7 (2.0)	2 (0.7)
At least one of the ten PF scale items	113 (32.4)	129 (46.1)
Objective measurements of physical performance		
Grip strength (kg)	44.3 (7.6)	26.0 (5.8)
TUG (secs)[d]	10.6 (1.9)	11.2 (2.4)
3m walk (secs)[d]	3.3 (0.5)	3.5 (0.7)
Chair rises (secs)[a]	15.2 (1.2)	16.7 (1.3)
Quadriceps peak torque (Nm)[e]	95.1 (19.1)	68.6 (16.4)
N(%) achieving maximal balance time of 30 seconds[f]	156 (46.4)	73 (26.8)

[a] Geometric mean and SD

[b] IIIM-V denotes classes three (manual) to five of the 1990 OPCS Standard Occupational Classification scheme for occupation and social class. Social class was identified on the basis of own current or most recent full-time occupation for men and never-married women, but on the basis of the husband's occupation for ever-married women.

[c] Inter-quartile range

[d] Excludes 4 men and 4 women who completed the TUG and 3m walk with the assistance of a walking aid

[e] Measurement of quadriceps strength was only introduced part way through the West Herts phase of the Hertfordshire Cohort Study and therefore data were only available for 125 men and 193 women

[f]Excludes 13 men and 8 women who did not participate in the balance test

UKPMC Funders Group Author Manuscript

Table 2

Relationships between SF-36 physical functioning (PF) scores and objective physical performance measures

	MEN		WOMEN	
Mean (SD)	**High PF**	**Low PF[a]**	**High PF**	**Low PF[a]**
Grip strength (kg)	45.2 (7.5)	41.0 (7.4)	26.8 (5.2)	23.0 (6.7)
unadjusted difference (95%CI)[b]	-4.2 (-6.1, -2.2), p<0.001		-3.8 (-5.4, -2.2), p<0.001	
adjusted[c] difference (95%CI)	-2.0 (-4.0, 0.1), p=0.06		-2.4 (-4.2, -0.7), p=0.007	
TUG (secs)	10.2 (1.5)	11.9 (2.6)	10.7 (1.7)	13.5 (3.3)
unadjusted difference (95%CI)[b]	1.7 (1.2, 2.2), p<0.001		2.8 (2.2, 3.5), p<0.001	
adjusted[c] difference (95%CI)	0.9 (0.4, 1.5), p<0.001		1.6 (1.0, 2.3), p<0.001	
3m walk (secs)	3.2 (0.4)	3.6 (0.6)	3.3 (0.5)	4.1 (0.9)
unadjusted difference (95%CI)[b]	0.4 (0.3, 0.5), p<0.001		0.8 (0.6, 1.0), p<0.001	
adjusted[c] difference (95%CI)	0.2 (0.1, 0.3), p=0.003		0.5 (0.3, 0.7), p<0.001	
Chair rises (secs)[d]	14.8 (1.2)	17.0 (1.3)	15.8 (1.2)	20.6 (1.4)
unadjusted difference (95%CI)[b]	1.2 (1.1, 1.2), p<0.001		1.3 (1.2, 1.4), p<0.001	
adjusted[c] difference (95%CI)	1.1 (1.0, 1.1), p=0.03		1.2 (1.1, 1.3), p<0.001	
Quadriceps peak torque (Nm)	95.4 (19.3)	92.5 (18.2)	70.5 (15.4)	52.5 (16.5)
unadjusted difference (95%CI)[b]	-3.0 (-13.7, 7.8), p=0.59		-18.0 (-25.2, -10.8), p<0.001	
adjusted[c] difference (95%CI)	1.3 (-9.6, 12.3), p=0.81		-16.7 (-23.7, -9.7), p<0.001	
Proportion				
Maximal balance time (30secs)	49.3%	34.4%	29.8%	14.8%
unadjusted odds ratio (95%CI)[b]	0.5 (0.3, 0.9), p=0.03		0.4 (0.2, 0.9), p=0.03	
adjusted[c] odds ratio (95%CI)	1.1 (0.6, 2.2), p=0.72		0.7 (0.3, 1.7), p=0.44	

CI = confidence interval; PF = physical functioning

[a] Low physical functioning scores defined as scores in the lowest fifth of the gender-specific distribution i.e. ≤ 60 for men, and ≤ 75 for women

[b] Differences/odds ratios are presented for low PF subjects in comparison with high PF subjects

[c] Adjustment factors were age, height, weight for height, walking speed, social class, smoking habit and alcohol consumption.

[d] Geometric mean (SD) and proportional differences in low compared with high PF subjects

UKPMC Funders Group Author Manuscript

Table 3

Percentiles for the SF-36 physical functioning (PF) score among older participants in the nationally representative Health Survey for England (HSE 1996)

Percentile of SF-36 PF score	Age band (years)					
	55-59	**60-64**	**65-69**	**70-74**	**75-79**	**80+**
MEN						
Minimum	0	0	0	0	0	0
p5	10	10	10	5	5	3
p10	25	20	20	20	15	10
p15	40	35	30	30	25	20
p25	70	55	55	50	40	30
p50	90	85	85	80	67	55
p75	100	95	95	90	85	78
p85	100	100	95	95	90	85
p90	100	100	100	95	95	90
p95	100	100	100	100	100	95
Maximum	100	100	100	100	100	100
N	*459*	*501*	*461*	*401*	*273*	*249*
WOMEN						
Minimum	0	0	0	0	0	0
p5	10	10	8	5	5	0
p10	28	25	20	10	5	5
p15	40	39	30	18	15	5
p25	60	55	50	30	30	13
p50	85	80	75	64	60	40
p75	95	94	90	85	80	70
p85	100	95	95	90	90	85
p90	100	100	100	95	95	90
p95	100	100	100	100	100	95
Maximum	100	100	100	100	100	100
N	*521*	*536*	*571*	*502*	*373*	*452*

p# denotes percentiles of the physical functioning distribution. p50 is the median of the distribution, and p25 to p75 define the inter-quartile range.