

# Discriminative power of the health status questionnaire 12 in relation to age, sex, and longstanding illness: findings from a survey of households in Great Britain

Ann Bowling, Joy Windsor

## Abstract

**Study objective**—To assess the ability of the health status questionnaire 12 (HSQ-12) to discriminate between older and younger age groups, its appropriateness for use with an older population in terms of the spread of responses across categories, floor or ceiling effects, and its ability to discriminate between those with and without a reported longstanding illness and type (sensitivity and specificity).

**Design and setting**—The vehicle for the study was the Office for National Statistics (ONS) omnibus survey in Great Britain. The sampling frame was the British post-code address file of “small users”, stratified by region, and socioeconomic factors. This file includes all private household addresses. The postal sectors were selected with probability proportional to size. Within each sector 30 addresses were selected randomly. The number of selected addresses was 3000.

**Participants**—Altogether 1912 adults aged 16 and over were interviewed in person in their own homes, giving a response rate of 72%.

**Measures**—The HSQ-12, and the ONS general household survey questions on longstanding illness; the ONS omnibus standard sociodemographic items.

**Main results**—There were exceptionally high rates of item response in all age groups. The score differences by construct (eg, age group, sex, longstanding illness) were in the expected directions with statistically significant age gradients. Age was associated with most of the HSQ-12 domains, although this association had interactions with longstanding illness or sex. The differences in HSQ-12 scores with reported longstanding illness and type of longstanding illness made theoretical sense, which supports the discriminative power of the scale. The frequency distributions for HSQ-12 items in relation to age and sex, and by reporting of longstanding illness are also presented here in order to demonstrate ceiling effects. Most respondents in all age groups achieved high (good) scores on the “social functioning” subscale. The HSQ-12 had good results for specificity when tested against reporting of a long-

standing illness, although this was at the expense of sensitivity.

**Conclusions**—The results support the use of the HSQ-12 with older populations, particularly for those with chronic illnesses, although it will reveal relatively few problems among younger populations. The results presented here indicate that it will require supplementation with more sensitive disease and/or domain specific scales in the areas of interest or intervention, but it provides an acceptable, brief, core measure of health related quality of life. This paper presents the first British normative data using the HSQ-12.

(*J Epidemiol Community Health* 1997;51:564-573)

The current emphasis in the measurement of health outcomes and health status is to incorporate the effects of health on physical, psychological, and social wellbeing, commonly referred to as health related quality of life.

There is much debate about appropriate measures of health related quality of life for elderly people. It cannot be assumed that instruments which have been reported to be psychometrically sound with younger populations are also suitable for use with people aged 65 and over and 75 and over. At the same time there is frequently a need to use the same instrument with all age groups because patient groups in clinical trials, and populations in retrospective or prospective surveys, are not necessarily restricted to a limited range of age groups. Comparative data are also required to assess trends in health status and health outcomes in relation to age. While broader health status or health related quality of life instruments have been widely used with younger populations, studies of elderly people have often focused more narrowly on measuring physical functioning and mental health. The latter tend to be based on the use of separate measurement scales for each dimension, rather than one scale comprising appropriate subscales. While the broad but lengthy sickness impact profile (136 items)<sup>1</sup> has been recommended for use with older people, alongside other domain specific scales,<sup>2</sup> there is no consensus about which packages of scales are appropriate.

One recently developed scale which might be suitable as a core instrument for use with

Centre for Health Informatics and Multiprofessional Education, University College London Medical School, The Whittington Hospital, 4th floor Archway Wing, Highgate Hill, London N19 4NF  
A Bowling  
J Windsor

Correspondence to:  
Professor A Bowling.

Accepted for publication  
June 1997

elderly, as well as younger, populations is the health status questionnaire-12 (HSQ-12). This was derived from the Health Outcomes Trust's 39 item health status questionnaire (HSQ) version 2.0 (an extended version of their 36 item scale, in turn based on the original Rand short-form 36 (popularly known as the SF-36)).<sup>3,4</sup> The Trust's 36 item version has been validated for use in the UK, and is the recommended version in this country. The 12 item version maintains the conceptual integrity and retains the eight health status constructs of the original. The original Rand SF-36, and the latest versions of it (which are almost identical), was extensively tested for its psychometric properties and is in international use in cross sectional, longitudinal, and clinical intervention studies.<sup>5</sup> An alternative short version of the SF-36 is the short form-12 health survey questionnaire (SF-12). This was developed by Ware *et al.*<sup>6</sup> The 12 items used to represent the eight subdomains of both the HSQ-12 and the SF-12 include self assessed (perceived) health (HSQ-12 and SF-12, 1 item), physical functioning (HSQ-12, 3 items; SF-12, 2 items), physical role limitation (HSQ-12, 1 item; SF-12, 2 items), mental role limitation (HSQ-12, 1 item; SF-12, 2 items), social functioning (HSQ-12 and SF-12, 1 item), mental health (HSQ-12, 3 items; SF-12, 2 items), energy/fatigue (HSQ-12 and SF-12, 1 item), and pain (HSQ-12 and SF-12, 1 item). While the developers of both the HSQ-12 and SF-12 selected several different items from the original scale, both developers claim to have selected the 12 items which reproduce at least 90% of the variance in the longer versions, while retaining the eight domains.<sup>3,6</sup> However, the aim underlying the development of the SF-12 was not only to produce the eight dimensions of the SF-36, but also to construct the two summary scores from the SF-36 - the *physical component summary* and the *mental component summary*. In this, it differs from the alternative HSQ-12. For the SF-12, then, Ware *et al.*<sup>6</sup> identified 10 items from six of the eight SF-36 scales that reproduced at least 90% of the variance in both the physical and mental component summaries, as defined using the SF-36 scales. The addition of two more items created a 12-item short form—the SF-12—that yielded satisfactory estimates of the physical and mental component summaries and made it possible to reproduce the profile of the eight SF-36 domains.<sup>6</sup>

While preliminary results have been published in the handbooks for each scale, the psychometric properties of both of the shorter 12-item versions are under investigation, coordinated by the scale developers. Ware *et al.*<sup>6</sup> reported that it is valid to use the published SF-36 summary measures (subscale scores) in order to interpret the (alternative short version) SF-12 scores. However, the studies which have reported population norms and the psychometric properties of the SF-36 have been based on inadequate sample sizes of very elderly people or the samples of elderly people have been unrepresentative.

An important research question is whether the 12-item versions can be used with an older,

#### KEY POINTS

- This paper presents the first British normative data for the health status questionnaire-12 (HSQ-12).
- There was some evidence of ceiling effects. For example, respondents in all age groups achieved high (good) scores on the "social functioning" subscale.
- Differences in HSQ-12 scores with reported long standing illness, and type, made theoretical sense. This supports the discriminative power of the scale. The specificity of the scales was achieved at the expense of sensitivity.
- The results reported here support the use of the HSQ-12 with older populations, and particularly with specific chronic illnesses, although it will reveal relatively few problems among younger populations.
- The HSQ-12 will require supplementation with more sensitive disease and/or domain specific scales in the areas of interest.

as well as a younger, population. Problems with item response and respondents' perceptions of the relevance of items have been reported with older people in England<sup>7-9</sup> in relation to the original Rand SF-36 and the SF-36 health surveys (both popularly called the SF-36, but distributed from different centres).<sup>5,10</sup> However, of the two samples of the English population on which the SF-36 was tested, one did not include people aged 65 and over<sup>11</sup> and the other did not include people aged 75 and over.<sup>7</sup> Moreover, the main English study investigating its suitability for use with an population aged 75 and over was based on participants in a clinical trial for treatment for bone disease, rather than a random sample of the population.<sup>8</sup>

#### Aims and methods

The study aimed to assess the ability of the 12 items contained in the HSQ-12<sup>3,4</sup> to discriminate between older and younger age groups, its appropriateness for use with an older population in terms of the spread of responses across categories (rather than floor or ceiling effects), and its ability to discriminate between those with (sensitivity) and without (specificity) a reported longstanding illness and type.

The study design was an interview survey based on a national random sample of people aged 16 and over in Great Britain, taken by the Office for National Statistics (ONS) for their August omnibus survey. Respondents were interviewed in person in their own homes. The sampling frame for the survey was the postcode address file of "small users", which includes all private household addresses. It was stratified by region, housing tenure, and socioeconomic group. One hundred postal sectors were selected with probability proportionate to size, and, within each sector, 30 addresses were selected randomly. If an address contained more than one household, the interviewer used

Table 1 Sociodemographic characteristics of the Office of National Statistics (ONS) omnibus sample in comparison with the general household survey sample (GHS)

Characteristic	ONS omnibus sample (aged $\geq 16$ y) % (No)	GHS sample (aged $\geq 16$ y) % (No)
Age (y):		
16<45	51 (987)	51 (9449)
45<65	29 (550)	29 (5360)
> 65	20 (375)	21 (3884)
Sex:		
Male	48 (910)	47 (8812)
Female	52 (1002)	53 (9881)
Marital status:		
Married/cohabiting	64 (1231)	64 (11971)
Single	21 (402)	21 (3861)
Widowed/divorced/separated	15 (279)	15 (2847)
Ethnic group:		
White	94 (1799)	95 (17675)
Black/other	6 (112)	5 (911)
Housing tenure:		
Owner occupier/mortgagee	73 (1384)	67 (6544)
Rented from local authority/housing association	19 (359)	25 (2426)
Rented privately or with job/other	8 (162)	9 (853)
Economic activity:		
Working (full or part time)	55 (1056)	54 (9945)
Unemployed or inactive	45 (853)	46 (8443)
Health status:		
Reported longstanding illness, disability of infirmity	40 (756)	39 (7238)
None reported	60 (1150)	61 (11222)
No of respondents	(1905-1912)*	(9852-18693)*

\* Varying totals reflect item non-response, and also the use of different bases (households for housing tenure or individuals for other items in the GHS sample). Numbers in tables may not equal total due to weighting.

a standard procedure to select just one household randomly. Within households with more than one adult member, just one person aged 16 or over was selected with the use of a Kisch grid. Because only one household member was interviewed, people in households that contained few adults had a better chance of selection than those in households with many. A weighting factor was applied to correct for this unequal probability. The individual adult, rather than the household, was the unit of analysis in the results presented here.

#### SAMPLE SIZE

The number of selected addresses was 3000; this contained 329 (10%) ineligible addresses (eg non-domestic). At the remaining 2671 addresses, 470 (18%) people refused to take part, 289 (10%) were non-contactable, and 1912 people were interviewed at home, giving a response rate of 72%. No data on the characteristics of non-responders was collected by ONS in order to assess response bias, although comparisons of the sample of responders can be made with census data (population estimates) and with the respondents to the ONS's general household survey (see later). These indicate that the characteristics of the omnibus sample of responders compare favourably with each of the others. Interviewing was completed within three weeks during August 1996. Recall interviews took place within the interviewing period with a proportion of respondents (about a 10th) as a check that the interviews actually took place and to check consistency of responses. This is reassuring for investigators as a quality control check on interviewers (figures on consistency unavailable from ONS as they are used as a basis for retraining interviewers where necessary and are regarded as confidential data).

#### MEASURES

The HSQ-12<sup>3,4</sup> was used, which asks about the impact of health on social, emotional, and physical functioning over the past four weeks (making it suitable for use with people with more chronic health problems). The scoring involves two steps which can be carried out on the computer (eg, using the *Statistical Package for the Social Sciences - SPSS*); first the recoding of responses, followed by the calculation of scale scores. This involves calculating an average of the recoded response values for the multi-item scales (physical functioning and mental health) to create a summary score. The recoded response value is the summary score for single item scales. The summary scores represent the proportion of total possible values obtainable for the scale (eg a score of 50 on the physical functioning scale represents an achievement of 50% of the possible maximum score). The higher scores represent a positive health attribute (eg no activity limitations or freedom from pain).<sup>4</sup>

The commonly used ONS question was asked: vis "Do you have any longstanding illness, disability or infirmity? By longstanding I mean anything that has troubled you over a period of time or that is likely to affect you over a period of time?" Respondents who said "yes" were asked what the condition was, and whether it "limits your activities in any way?"<sup>12</sup> Standard ONS omnibus sociodemographic items were also included.

This paper focuses on analyses of the HSQ-12 in relation to age, sex, longstanding illness, and type. The data were analysed with the *SPSS*, using univariate, bivariate, and multivariate statistics.

#### Results

Table 1 shows the sociodemographic characteristics of the sample, which are comparable with the characteristics of the national sample for the general household survey (GHS).<sup>12</sup> The GHS was used as a comparison because it is based on a large random sample of adults in Britain and covers a wide range of topics. Thus, while it is based on similar sampling techniques as the omnibus survey (rather than being a census of the population), it does have the unique advantage of permitting comparisons with a wide range of variables. ONS provided the precise GHS figures for respondents in the comparable age groups to the omnibus survey for table 1 (personal communication, ONS). ONS also carried out comparisons with the more recently available 1995 GHS data and with the population estimates for 1995, derived from census data. These 1995 comparisons confirm that the omnibus sample compared favourably with the GHS sample and with the population estimates in age and sex distributions (personal communication, ONS). Checks which the ONS make on non-response bias for the GHS indicate that this is small.<sup>12</sup>

Item response was good, with just between 1 and 7 (0.05%-0.3%) non-responders across these items. Table 1 also shows that the proportions who reported a longstanding illness in both surveys were similar. In the

omnibus survey, 40% of respondents reported this, with 71% of these reporting that the condition limited their activities.

The proportions of those who reported a longstanding illness increased by age for both men and women: 19% of males and 25% of females aged 16<25 reported these, in comparison with 29% of males and 27% of females aged 25<45, 44% of males and 48% of females aged 45<55, 52% of males and 54% of females aged 55<65, 64% of males and 56% of females aged 65<75 and 64% of males and 65% of females aged 75 and over.

The most commonly mentioned longstanding illnesses were those affecting the musculoskeletal system (47%), heart and circulatory system (22%), respiratory system (15%), and digestive system (19%). Other conditions were mentioned by between 1 and 8% of respondents in each case (these included nervous system, endocrine and metabolic, mental disorders, ear complaints, eye complaints, genitourinary, neoplasms and benign tumours, skin complaints and various other conditions).

#### DISCRIMINATIVE POWER 1. VARIATIONS IN RELATION TO AGE AND SEX

Table 2 shows the frequency distributions for the HSQ-12 subscale items in relation to age and sex. The trends with age, controlling for sex, were tested using a  $\chi^2$  trend test (items were dichotomised into "no limitations/problems" versus "any", except for general health perceptions which was dichotomised into "fair/poor" versus "excellent, very good/good" and mental health items which dichotomised the worst two categories versus the rest). The table shows that the items relating to health perception and physical health (physical functioning and role limitations—physical) showed a clear gradient by age, and were statistically significant for both sexes separately at the  $p<0.001$  level. The trends with age were slightly less consistent for role limitations—mental (not significant for males), social functioning (not significant for males), and mental health (not significant for males and two of the three items were not significant for females). The trends with age were not entirely consistent for physical pain and energy/fatigue, although they were generally in the expected direction, and did achieve significance at the  $p<0.001$  level ( $\chi^2$  trend values available from the authors).

While the overall trends with age were in the expected directions, table 2 shows that there were slight improvements in some items with the approach of old age, particularly among men (perhaps reflecting "healthy survivors"), although deteriorations were evident again at age 75+. Table 2 also shows that women tended to report poorer health than men. Item response was again good, with between 1 and 10 non-responders across the items (0.05%–0.5%).

The value of displaying frequencies, rather than following the common practice of simply presenting subscale means and SDs (even with ranges of responses—maximum to minimum scores—displayed), is that the actual number of

people who report problems at each level can be viewed, and hence one can judge the spread of responses, and any floor or ceiling effects which are masked in the presentation of means. Thus, table 2 shows that ceiling effects were often quite high across most age groups. For example, between 77% and 87% of males, and between 69% and 82% of females in each age group reported no limitations on the "social functioning" scale.

The remaining scale items showed a better spread of distributions across categories for purposes of analysis (which is important in longitudinal research and intervention studies). Table 2 shows that the HSQ-12 items detect few positive responses indicating problems among younger age groups. While younger people are expected to be relatively healthy, more sensitive measures showing gradations in health status are needed for the middle age group 45<55. The scale shows a better spread of distributions after age 55. The means for the HSQ-12 subscales reflect these trends with age and sex, although they showed greater consistency in the trends with age. These normative data are shown in table 3. Lower scores reflect poorer health status.

Multiple regression was used to estimate the effects of age and sex on each of the eight HSQ-12 scales. This showed that, in general, age and sex were both significant independently in each model except for general health perceptions (age only was significant), social functioning (age only), mental health (sex only), and there were interactions between age and sex for physical functioning and for pain (figures available from the authors). The trends were in the expected direction (eg females had worse scores than males for mental health). (This was without allowing for longstanding illness—see next section for this).

#### DISCRIMINATIVE POWER 2. VARIATIONS BY LONGSTANDING ILLNESS, DISABILITY OR INFIRMITY AND TYPE

Table 4 shows the HSQ-12 scale means with reported longstanding illness, disability, or infirmity (LSI). This shows that those who reported a LSI had lower (worse) HSQ-12 mean scores than those who did not. For each HSQ-12 scale this difference was statistically significant at the  $p<0.001$  level, using a  $t$  test as follows: general health perceptions  $t$  21.62; physical functioning  $t$  21.73; role limitations—physical  $t$  20.45; role limitations—mental  $t$  10.91; social functioning  $t$  11.30; pain  $t$  20.97; energy/fatigue  $t$  16.49; mental health  $t$  10.54.

These data support the construct validity of the scale. Respondents who were scored as worse on the HSQ-12, but who did not report a longstanding condition are assumed to have acute, self limiting health problems. These findings are mirrored in the frequency results for the HSQ-12 individual items. The "mental health" scale was less markedly discriminatory than the other items.

The frequency distributions of responses to the HSQ-12 scales in relation to longstanding illness are displayed in table 5 in order to assess the sensitivity and specificity of the HSQ-12.

The data in table 5 were tested using the  $\chi^2$  statistic (dichotomised for sensitivity/specificity). All differences were highly significant ( $p < 0.00001$ ) as follows: general health perceptions  $\chi^2$  429.4, 4 df; physical functioning—lifting  $\chi^2$  430.1, 2 df; physical functioning—climbing stairs  $\chi^2$  414.9, 2 df; physical functioning—walking  $\chi^2$  344.7, 2 df; role limitations physical  $\chi^2$  443.6, 4 df; role limitations—mental  $\chi^2$  133.0, 4 df; social functioning  $\chi^2$  157.7, 4 df; pain  $\chi^2$  424.9, 5 df; mental health—calm/peaceful  $\chi^2$  98.74, 5 df; men-

Table 2 Health status questionnaire 12 (HSQ-12) in relation to age and sex

HSQ-12	Men: age (y)						Women: age (y)					
	16<25 % (No)	25<45 % (No)	45<55 % (No)	55<65 % (No)	65<75 % (No)	75+ % (No)	16<25 % (No)	26<45 % (No)	45<55 % (No)	55<65 % (No)	65<75 % (No)	75+ % (No)
Health perception (Q1) self assessed health (generally):												
Excellent	24 (35)	23 (81)	19 (26)	16 (16)	11 (12)	7 (4)*	23 (27)	25 (90)	14 (26)	16 (20)	13 (16)	10 (8)*
Very good	41 (62)	41 (145)	36 (49)	27 (28)	35 (38)	21 (12)	36 (41)	40 (146)	40 (74)	33 (41)	25 (31)	19 (16)
Good	28 (42)	24 (85)	25 (34)	24 (24)	23 (24)	33 (18)	30 (34)	25 (91)	22 (42)	22 (28)	26 (33)	29 (24)
Fair	5 (8)	9 (34)	15 (22)	19 (19)	23 (24)	32 (18)	10 (11)	8 (31)	17 (31)	19 (24)	25 (32)	32 (27)
Poor	2 (3)	3 (11)	5 (7)	14 (14)	8 (9)	7 (4)	1 (1)	2 (9)	7 (13)	8 (10)	11 (14)	10 (8)
Physical functioning (Q2-4) health limits (current):Lifting or carrying groceries:												
A lot	2 (3)	4 (16)	10 (14)	24 (24)	16 (17)	29 (16)*	4 (5)	4 (16)	9 (17)	21 (26)	28 (35)	44 (37)*
A little	2 (3)	4 (13)	8 (11)	14 (14)	23 (24)	21 (12)	13 (15)	15 (53)	26 (48)	19 (23)	35 (44)	30 (25)
Not at all	96 (144)	92 (328)	82 (114)	62 (63)	61 (66)	50 (28)	83 (94)	81 (296)	61 (121)	60 (73)	37 (47)	26 (22)
Climbing several flights of stairs:												
A lot	— (—)	4 (15)	12 (17)	21 (21)	20 (21)	32 (18)*	2 (2)	4 (16)	12 (22)	20 (24)	32 (40)	43 (36)*
A little	9 (14)	6 (21)	8 (11)	20 (20)	25 (27)	27 (15)	19 (22)	10 (36)	21 (39)	25 (31)	28 (35)	37 (31)
Not at all	91 (136)	90 (320)	79 (110)	60 (60)	55 (59)	41 (23)	79 (90)	86 (313)	67 (125)	55 (67)	40 (51)	20 (17)
Walking half a mile:												
A lot	2 (3)	2 (6)	10 (15)	19 (19)	18 (19)	31 (17)*	2 (2)	2 (6)	6 (12)	19 (23)	27 (34)	41 (34)*
A little	5 (8)	4 (15)	4 (6)	10 (10)	16 (18)	21 (11)	9 (10)	6 (23)	11 (21)	11 (14)	20 (25)	19 (16)
Not at all	93 (139)	94 (335)	84 (118)	71 (72)	66 (71)	48 (27)	89 (102)	92 (338)	83 (153)	70 (86)	53 (68)	40 (33)
Role limitations attributable to physical health (role-physical)(Q5): During past 4 weeks, difficulty doing work or other regular daily activities as a result of physical health:												
None at all	87 (130)	85 (306)	73 (102)	62 (63)	67 (72)	44 (24)*	79 (91)	79 (290)	64 (119)	57 (70)	47 (59)	33 (28)*
A little bit	11 (17)	5 (19)	9 (13)	10 (10)	12 (13)	18 (10)	7 (8)	11 (40)	17 (31)	16 (20)	18 (23)	23 (19)
Some	2 (1)	4 (15)	4 (5)	6 (6)	11 (11)	14 (7)	5 (5)	4 (14)	9 (16)	11 (13)	14 (18)	18 (15)
Quite a bit	— (1)	2 (8)	10 (14)	16 (16)	8 (9)	16 (8)	9 (10)	4 (15)	7 (14)	11 (14)	16 (20)	17 (14)
Could not do	— (—)	2 (8)	4 (6)	5 (5)	2 (3)	8 (4)	— (—)	2 (9)	3 (6)	5 (6)	5 (6)	9 (8)
Role limitations attributable to mental health (role-mental)(Q6): During past 4 weeks, accomplished less doing work or other daily activities as a result of emotional problems (such as feeling depressed or anxious):												
Not at all	77 (115)	76 (269)	69 (96)	71 (71)	85 (91)	67 (36)*	74 (84)	73 (265)	64 (120)	66 (80)	62 (78)	56 (47)*
Slightly	18 (27)	12 (43)	18 (25)	15 (15)	7 (8)	11 (6)	14 (16)	15 (56)	17 (31)	9 (11)	20 (26)	18 (15)
Moderately	3 (5)	5 (19)	3 (4)	6 (6)	3 (3)	4 (2)	3 (4)	4 (16)	7 (13)	6 (8)	7 (8)	11 (9)
Quite a bit	2 (4)	6 (21)	5 (7)	7 (7)	3 (3)	13 (7)	8 (9)	6 (21)	9 (17)	16 (19)	9 (11)	11 (9)
Extreme amount	— (—)	1 (3)	5 (7)	1 (1)	2 (3)	5 (3)	1 (1)	2 (7)	3 (6)	3 (4)	2 (3)	4 (4)
Social functioning Q7: During past 4 weeks, physical health or emotional problems interfered with normal social activities with family, friends, neighbours or groups:												
Not at all	84 (126)	84 (296)	77 (106)	80 (81)	87 (94)	81 (44)+	79 (90)	82 (301)	73 (135)	72 (88)	74 (95)	69 (58)*
Slightly	10 (16)	6 (21)	8 (11)	5 (5)	5 (5)	10 (5)	15 (17)	9 (32)	10 (19)	10 (13)	5 (6)	16 (13)
Moderately	3 (4)	3 (11)	5 (6)	7 (7)	2 (2)	1 (1)	— (—)	2 (6)	7 (13)	7 (9)	8 (10)	7 (6)
Quite a bit	1 (2)	5 (18)	5 (7)	7 (7)	3 (3)	9 (5)	5 (6)	6 (21)	9 (16)	8 (10)	9 (11)	7 (6)
Extreme amount	2 (3)	2 (8)	6 (8)	1 (1)	3 (3)	— (—)	1 (1)	1 (5)	2 (3)	3 (3)	4 (5)	1 (1)
Physical pain in past 4 weeks (Q8):												
None	55 (81)	61 (218)	57 (79)	52 (52)	42 (45)	42 (23)*	64 (72)	5 (215)	41 (75)	35 (44)	28 (35)	32 (27)*
Very mild	31 (46)	12 (43)	13 (19)	8 (8)	17 (18)	9 (5)	18 (20)	15 (54)	18 (33)	13 (16)	11 (14)	11 (9)
Mild	7 (11)	12 (43)	9 (12)	9 (9)	14 (15)	14 (8)	9 (10)	12 (45)	12 (23)	13 (16)	18 (23)	15 (13)
Moderate	4 (6)	10 (35)	11 (16)	19 (19)	16 (18)	24 (14)	5 (6)	9 (32)	22 (41)	20 (24)	27 (34)	26 (21)
Severe	3 (5)	3 (2)	5 (7)	7 (7)	8 (8)	3 (2)	2 (3)	4 (16)	4 (8)	14 (17)	12 (15)	17 (14)
Very severe	— (—)	2 (6)	5 (7)	5 (5)	3 (4)	8 (5)	2 (2)	1 (5)	3 (6)	5 (6)	4 (5)	1 (1)
Mental health (Q9,11,12): During past 4 weeks have felt calm and peaceful:												
All of the time	11 (17)	9 (31)	6 (8)	21 (21)	29 (31)	24 (13)+	11 (13)	8 (29)	6 (11)	7 (8)	13 (17)	19 (16)+
Most of the time	57 (85)	49 (176)	50 (69)	35 (35)	52 (56)	49 (27)*	50 (56)	42 (159)	38 (71)	49 (60)	42 (54)	39 (32)
A good bit of the time	16 (23)	15 (54)	20 (28)	19 (19)	7 (8)	5 (3)	21 (24)	17 (60)	18 (34)	8 (10)	12 (15)	10 (8)
Some of the time	12 (18)	16 (57)	14 (20)	16 (16)	9 (9)	13 (7)	14 (16)	21 (76)	25 (47)	29 (35)	23 (29)	21 (17)
Little of the time	2 (3)	9 (30)	8 (11)	5 (5)	2 (3)	7 (4)	3 (4)	9 (32)	10 (18)	3 (3)	6 (8)	7 (6)
None of the time	2 (4)	2 (7)	2 (2)	4 (4)	1 (1)	3 (2)	1 (1)	3 (10)	3 (5)	4 (5)	4 (5)	4 (3)
Felt downhearted and low:												
All of the time	1 (2)	1 (4)	1 (1)	1 (1)	1 (1)	1 (1)+	4 (4)	2 (6)	1 (2)	3 (3)	3 (4)	3 (2)*
Most of the time	2 (3)	4 (15)	4 (5)	3 (3)	2 (3)	— (—)	4 (4)	3 (12)	6 (11)	9 (11)	6 (7)	4 (4)
A good bit of the time	9 (13)	4 (15)	5 (7)	— (—)	1 (2)	5 (3)	3 (4)	8 (28)	7 (13)	4 (5)	12 (15)	10 (8)
Some of the time	8 (12)	16 (55)	20 (28)	19 (19)	11 (11)	19 (10)	16 (19)	17 (64)	26 (49)	20 (24)	20 (25)	29 (23)
Little of the time	42 (64)	40 (141)	38 (52)	31 (31)	25 (26)	34 (19)	47 (54)	39 (140)	33 (61)	29 (35)	23 (29)	19 (16)
None of the time	38 (57)	35 (126)	32 (45)	46 (45)	60 (65)	41 (23)	26 (30)	31 (114)	27 (50)	35 (44)	36 (46)	35 (29)
Been a happy person:												
All of the time	15 (23)	13 (45)	12 (17)	19 (19)	30 (32)	33 (18)+	23 (26)	14 (50)	11 (21)	19 (24)	19 (24)	22 (18)+
Most of the time	59 (89)	57 (207)	48 (67)	56 (55)	55 (58)	38 (21)	45 (52)	49 (180)	46 (85)	47 (58)	44 (56)	39 (32)
A good bit of the time	18 (27)	14 (48)	17 (23)	10 (10)	7 (7)	9 (5)	16 (18)	16 (58)	17 (32)	10 (13)	12 (15)	17 (14)
Some of the time	6 (8)	11 (38)	15 (20)	11 (11)	5 (5)	16 (9)	13 (15)	15 (54)	14 (26)	16 (19)	18 (22)	17 (14)
Little of the time	2 (3)	4 (14)	8 (10)	4 (4)	4 (4)	3 (2)	3 (3)	6 (22)	10 (18)	5 (6)	5 (6)	4 (3)
None of the time	— (—)	1 (5)	— (1)	— (—)	1 (1)	1 (1)	— (1)	— (2)	2 (4)	3 (3)	2 (3)	1 (1)
Energy/fatigue (Q10): During past 4 weeks had a lot of energy:												
All of the time	13 (20)	12 (42)	9 (13)	11 (11)	14 (15)	8 (4)*	7 (6)	8 (30)	4 (7)	6 (8)	7 (9)	6 (5)*
Most of the time	44 (66)	44 (154)	37 (51)	33 (33)	33 (36)	33 (18)	53 (61)	38 (140)	31 (59)	40 (48)	28 (35)	15 (13)
A good bit of the time	25 (38)	17 (60)	25 (34)	15 (15)	18 (20)	8 (5)	20 (22)	16 (60)	17 (32)	12 (15)	10 (13)	15 (12)
Some of the time	14 (21)	21 (74)	13 (18)	21 (21)	21 (21)	30 (17)	9 (10)	25 (91)	27 (49)	19 (24)	24 (31)	31 (25)
Little of the time	3 (4)	5 (19)	13 (18)	10 (10)	8 (8)	10 (6)	11 (12)	10 (35)	17 (31)	14 (17)	21 (27)	18 (15)
None of the time	1 (2)	1 (5)	3 (4)	10 (10)	7 (7)	11 (6)	1 (1)	3 (10)	4 (7)	9 (11)	9 (11)	15 (12)
No of respondents to each item:	150	354-356	138-139	100-101	107-108	55-56	114	364-367	186	122-123	126-127	81-84

$\chi^2$  trend tests (men and women separately; variables dichotomised into no limitations/problems versus any):

\*  $p < 0.001$  (except social functioning -  $p < 0.01$ ).

+ Not statistically significant.

Table 3 Health status questionnaire 12 (HSQ-12) scores in relation to age

HSQ-12	Grouped age (y) - men							Grouped age (y) - women						
	16<25	25<45	45<55	55<65	65<75	75+	Total	16<25	25<45	45<55	55<65	65<75	75+	Total
General health perceptions:														
Mean	76.9	74.1	67.7	58.2	60.4	53.3	68.9	74.6	75.4	65.6	63.0	55.7	51.6	67.5
Median	85.0	85.0	85.0	60.0	60.0	60.0	85.0	85.0	85.0	85.0	60.0	60.0	60.0	85.0
SD	21.9	25.5	29.2	34.2	31.3	29.1	28.5	23.4	24.4	30.0	31.8	32.5	31.1	29.2
Physical functioning:														
Mean	96.1	94.4	85.6	71.6	71.3	57.9	85.8	90.5	91.4	81.2	70.8	57.4	43.1	78.5
Median	100.0	100.0	100.0	100.0	83.3	66.7	100.0	100.0	100.0	100.0	83.3	66.7	33.3	100.0
SD	13.9	17.7	30.8	36.5	34.3	39.2	28.6	19.8	19.3	27.7	36.0	37.6	36.3	31.9
Role limitations - physical:														
Mean	94.5	90.7	81.2	72.2	78.2	61.6	86.4	86.4	87.5	77.8	71.0	63.6	54.0	77.7
Median	100.0	100.0	100.0	100.0	100.0	65.0	100.0	100.0	100.0	100.0	100.0	65.0	65.0	100.0
SD	15.3	24.9	34.3	39.1	34.4	39.8	30.8	29.4	27.4	33.6	37.4	38.8	38.6	34.4
Role limitations - mental:														
Mean	90.1	87.3	82.9	85.0	91.1	78.5	86.8	86.2	85.6	80.0	77.4	79.6	75.1	82.0
Median	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
SD	19.3	24.6	29.3	26.2	23.0	33.8	25.4	25.9	26.2	30.2	33.9	32.3	29.4	29.1
Social functioning:														
Mean	93.1	90.9	86.4	88.8	92.7	90.7	90.5	91.6	91.2	85.9	85.2	84.3	86.2	88.2
Median	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
SD	18.9	23.3	28.7	24.5	22.0	22.1	23.5	20.0	22.1	26.3	27.2	29.7	24.1	24.7
Pain:														
Mean	87.9	84.5	80.3	75.1	74.5	69.8	81.3	87.6	84.1	74.5	67.5	64.4	66.2	76.7
Median	100.0	100.0	100.0	100.0	85.0	85.0	100.0	100.0	100.0	85.0	65.0	65.0	65.0	85.0
SD	18.3	24.1	29.2	30.9	28.8	32.2	26.6	21.9	24.0	27.9	31.5	29.4	28.9	27.9
Energy/fatigue:														
Mean	69.3	66.2	61.6	56.6	61.2	52.5	63.5	66.7	60.5	53.7	55.8	50.0	43.1	56.6
Median	80.0	80.0	60.0	60.0	60.0	40.0	80.0	80.0	60.0	60.0	60.0	40.0	40.0	60.0
SD	21.3	23.4	25.7	30.2	28.5	30.0	22.7	22.7	25.2	25.7	29.4	29.8	28.7	27.2
Mental health:														
Mean	76.0	72.2	70.3	75.3	81.8	76.7	74.3	73.1	69.5	66.1	69.4	68.8	70.0	69.2
Median	80.0	80.0	73.3	80.0	86.7	80.0	80.0	80.0	73.3	73.3	73.3	73.3	73.3	73.3
SD	16.6	19.2	18.7	18.5	17.5	19.3	18.7	16.8	19.8	21.2	22.6	22.0	21.2	20.6
No of respondents	150	354-365	139	100-101	108	54-56	906-909	114	366-367	186	123	127	81-84	996-1001

tal health—downhearted/low  $\chi^2$  77.4, 5 df; mental health—happy  $\chi^2$  98.28, 5 df; energy/fatigue  $\chi^2$  255.2, 5 df.

Modelling for the effects of age, sex, and longstanding illness on the HSQ-12 scale domains, using multiple regression, showed longstanding illness to be a significant factor for all HSQ-12 domains. Sex was a significant factor for all domains except health perceptions and social functioning, but it was not an independent predictor of physical functioning and pain. Age was also a significant factor for all domains except social functioning and role limitations—mental. In addition, all of the six scale domains with which age was significantly associated had interactions with longstanding illness and/or sex. Therefore age was not an independent predictor of HSQ-12 scale scores (details available from the authors).

Table 5 shows that few of those who reported that they did not have a longstanding illness reported any limitations or problems at the HSQ-12. Despite the statistically significant differences obtained between the two LSI groups in their HSQ-12 scale scores, relatively

large proportions of people with a longstanding illness failed to report limitations or problems with the HSQ-12. For example, between 47% and 61% of respondents with a longstanding illness reported that their health did not limit their physical functioning (Q2-4) and 44% said they suffered from no role limitations due to physical health. Fifty eight per cent of those who reported a longstanding illness said that they had experienced no role limitations due to mental health (Q6), and 66% said they had no limitations on their social activities due to their health (Q7).

The sensitivity and specificity of the HSQ-12 in relation to LSI was calculated using dichotomised responses to the HSQ-12 (based on the method used by Ware *et al* for the presentation of dichotomised frequencies for the SF-36<sup>®</sup>); these are given in table 6. The figures show that for all scales high specificity was achieved at the expense of sensitivity.

Table 7 shows the HSQ-12 subdomain mean scores by type of longstanding illness (frequencies available on request to authors). Cross tabulations with LSI were conducted only for those who reported one condition, in order to avoid confounding by comorbidity. Table 7 shows that the variations between mean scores and type of condition do make theoretical sense, supporting the validity of the scale. For example, people who reported disorders of the heart and circulatory system had the worst scores on the physical functioning subscale (mean: 53). People who reported mental disorders scored worst on role limitations—mental (mean: 50), social functioning (mean: 61), and also on mental health (mean: 48). For pain, respondents with disorders of the musculoskel-

Table 4 Health status questionnaire 12 (HSQ-12) subdomains in relation to longstanding illness (LSI)

HSQ-12 subdomains:	Reported LSI Mean (SD)	No LSI reported Mean (SD)
General health perceptions	51.2 (31.5)	79.2 (20.6)*
Physical functioning	63.1 (37.6)	94.4 (15.3)*
Role limitations:		
Physical	61.9 (39.1)	93.4 (20.0)*
Mental	75.3 (33.0)	90.1 (21.3)*
Social functioning	81.0 (30.7)	94.8 (16.5)*
Pain	63.1 (30.7)	89.2 (18.8)*
Energy/fatigue	47.9 (27.3)	67.8 (23.2)*
Mental health	65.6 (22.0)	75.6 (17.3)*

\* p < 0.001 (t test).

etal system scored worst (mean: 57). These associations indicate that the HSQ-12 is able to discriminate between conditions where this was expected theoretically.

One way analysis of variance (ANOVA) was carried out to test for significant differences

between the HSQ-12 means in relation to each type of condition as displayed in table 7, and Bonferroni's *t* pairways multiple comparison procedure was carried out post hoc. This corrects for the effects of multiple testing (it increases the critical F value needed for the

Table 5 Health status questionnaire 12 (HSQ-12) in relation to longstanding illness

	Longstanding illness		
	Yes % (No)	No % (No)	Total % (No)
Health perceptions (Q1) self assessed health (generally):			
Excellent	7 (52)	27 (310)*	19 (363)
Very good	24 (180)	44 (503)	36 (683)
Good	29 (216)	23 (264)	25 (480)
Fair	28 (209)	6 (72)	15 (281)
Poor	13 (98)	0 (5)	5 (103)
Physical functioning (Q2-4) health limits (current):			
Lifting or carrying groceries			
A lot	27 (203)	2 (23)*	12 (226)
A little	25 (189)	8 (95)	15 (284)
Not at all	48 (364)	90 (1033)	73 (1397)
Climbing several flights of stairs			
A lot	23 (208)	2 (24)*	12 (233)
A little	25 (189)	10 (113)	16 (302)
Not at all	47 (357)	88 (1015)	72 (1371)
Walking half a mile			
A lot	23 (177)	1 (14)*	10 (190)
A little	16 (119)	5 (58)	9 (177)
Not at all	61 (459)	94 (1082)	8 (1542)
Role limitations attributable to physical health (role-physical) (Q5) During the past 4 weeks difficulty doing work or other regular daily activities as a result of physical health:			
None at all	44 (335)	88 (1019)*	71 (1354)
A little bit	19 (145)	7 (76)	12 (221)
Some	13 (101)	2 (28)	7 (129)
Quite a bit	16 (120)	2 (23)	7 (143)
Could not do	7 (53)	1 (8)	3 (60)
Role limitations attributable to mental health (role mental) (Q6) During past 4 weeks accomplished less during work or other daily activities as a result of emotional problems (such as feeling depressed or anxious)			
Not at all	58 (437)	79 (916)*	71 (1353)
Slightly	18 (132)	13 (145)	15 (277)
Moderately	7 (52)	4 (45)	5 (97)
Quite a bit	13 (95)	3 (40)	7 (135)
Extreme amount	5 (34)	1 (7)	2 (41)
Social functioning (Q7) During past 4 weeks, physical health or emotional problems interfered with normal social activities with family, friends, neighbours or groups:			
Not at all	66 (499)	88 (1015)*	80 (1514)
Slightly	11 (80)	7 (83)	9 (163)
Moderately	7 (56)	2 (18)	4 (74)
Quite a bit	11 (84)	2 (28)	6 (112)
Extreme amount	4 (32)	1 (9)	2 (41)
Physical pain in past 4 weeks (Q8) Pain rating:			
None	27 (205)	66 (761)*	51 (966)
Very mild	13 (99)	16 (187)	15 (286)
Mild	15 (115)	10 (112)	12 (227)
Moderate	26 (195)	6 (69)	14 (264)
Severe	12 (92)	2 (20)	6 (112)
Very severe	6 (47)	0 (4)	3 (51)
Mental health (Q9,11,12): During past 4 weeks have you:			
Felt calm and peaceful:			
All of the time	7 (54)	14 (161)*	11 (215)
Most of the time	38 (284)	52 (595)	46 (879)
A good bit of the time	16 (199)	14 (167)	15 (285)
Some of the time	26 (191)	14 (156)	18 (347)
Little of the time	10 (73)	5 (54)	7 (127)
None of the time	4 (29)	2 (20)	3 (4)
Felt downhearted and low:			
All of the time	2 (18)	1 (12)*	2 (30)
Most of the time	6 (45)	3 (33)	4 (78)
A good bit of the time	9 (70)	4 (41)	6 (110)
Some of the time	22 (162)	15 (178)	18 (340)
Little of the time	34 (252)	36 (416)	35 (668)
None of the time	27 (203)	41 (471)	35 (674)
Been a happy person:			
All of the time	14 (104)	18 (213)*	17 (317)
Most of the time	42 (313)	56 (645)	50 (958)
A good bit of the time	16 (116)	13 (155)	14 (271)
Some of the time	19 (142)	9 (98)	13 (241)
Little of the time	8 (58)	3 (38)	5 (95)
None of the time	2 (17)	0 (4)	1 (20)
Energy/fatigue (Q10) During past 4 weeks: Had a lot of energy:			
All of the time	4 (29)	12 (143)*	9 (172)
Most of the time	24 (180)	46 (533)	38 (714)
A good bit of the time	16 (122)	18 (203)	17 (325)
Some of the time	28 (214)	16 (190)	21 (403)
Little of the time	18 (139)	5 (63)	11 (202)
None of the time	9 (68)	2 (19)	5 (86)
No of respondents	750-756	1151-1153	1900-1910

\*  $\chi^2$  test  $p < 0.00001$ .

Table 6 Sensitivity and specificity of the health status questionnaire 12 (HSQ-12) in relation to longstanding illness

Scale	Sensitivity (%)	Specificity (%)
General health perceptions	40.7	93.4
Physical functioning	65.0	83.0
Role limitations:		
Physical	55.5	83.3
Mental	41.8	79.4
Social functioning	33.6	88.0
Pain	72.8	66.0
Energy/fatigue	27.5	92.9
Mental health	24.7	90.6

comparison to be declared significant and hence reduces the likelihood of spurious significant results).

The results for the ANOVA and Bonferroni's tests confirmed these observations which made theoretical sense, they were:

- General health perceptions, ANOVA,  $F=4.1$ ;  $p<0.0001$ ; Bonferroni's: the mean for disorders of the heart and circulatory system was significantly different from the means for (in ascending order of significance) musculoskeletal disorders, diseases of the nervous system, "other" conditions, ear complaints;
- Physical functioning,  $F=5.10$ ,  $p<0.0001$ . The mean for disorders of the heart and circulatory system was significantly different from the means for respiratory disorders, "other" conditions, digestive disorders, mental disorders and ear complaints;
- Role limitations—physical,  $F=2.58$ ,  $p<0.01$ . No two groups were significantly different at the 5% level.
- Role limitations—mental,  $F=3.83$ ,  $p<0.001$ . The mean for mental disorders was significantly different from the means for musculoskeletal disorders, disorders of the digestive system, "other" conditions, disorders of the heart and circulatory system, respiratory, nervous system and ear conditions.
- Social functioning,  $F=3.46$ ,  $p<0.001$ . The mean for mental disorders was significantly different from the means for disorders of the heart and circulatory system, disorders of the respiratory system, musculoskeletal system, nervous system, "other" conditions, and ear complaints.
- Pain,  $F=7.51$ ,  $p<0.0001$ . The mean for musculoskeletal disorders was significantly different from the means for disorders of the heart and circulatory system, respiratory conditions, "other" conditions, and ear complaints.
- Energy/fatigue,  $F=4.31$ ,  $p<0.0001$ . The mean for ear complaints was significantly different from the means for neoplasms, mental disorders, disorders of the heart and circulatory system, and musculoskeletal disorders.
- Mental health,  $F=4.23$ ,  $p<0.0001$ . The mean for mental disorders was significantly different from the means for all conditions except neoplasms.

## Discussion

This paper presents the first British normative data for the HSQ-12, with analyses of its discriminative power in relation to age, sex, longstanding illness and type, and sensitivity and specificity in relation to reported longstanding illness. It contributes to the programme of international research into the HSQ-12.

The HSQ-12 was highly acceptable to respondents, and there were exceptionally high rates of item response demonstrating the value of short and simple scales, as well as of scales administered by well trained interviewers. An advantage of the HSQ-12 over the alternative SF-12,<sup>6</sup> is that, unlike the SF-12, the HSQ-12 excludes the *examples* of items which were included the original question wording of the SF-36, and which some people may perceive to be irrelevant, possibly leading them not to respond to items. The inclusion of "irrelevant" item examples might explain the high item non-response reported for the SF-36 in England.<sup>8 9 13</sup> For example, the physical functioning scale in the HSQ-12 simply asks about three limitations due to health: lifting or carrying groceries, climbing several flights of stairs, and walking several blocks (half a mile). In contrast, the SF-36 and SF-12 physical functioning scales both include examples for some items (eg "Does your health now limit you in:—moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf?"). Examples can be leading or even encourage people to interpret the scale as irrelevant to them if they are not part of their everyday lives (eg lifting furniture, specific (often social class related) sports activities).

The main advantages of the HSQ-12, like the SF-12, is its brevity and simplicity of completion and analysis. These considerations are important in busy clinical settings, with frailer people and where a generic core scale is required to supplement disease specific and other batteries of measures. With the increasing popularity of disease specific measurement scales in order to increase the sensitivity and relevance of measures, the demand for shorter core, generic scales as a supplement is also increasing. Inevitably, the psychometric properties of a scale are reduced with short versions,<sup>6</sup> but this has to be balanced against the threat of incomplete data with longer instruments and batteries.

The frequency distributions for HSQ-12 items in relation to age and sex, and by reporting of longstanding illness were presented here in order to demonstrate ceiling effects. For example, most respondents in all age groups achieved high (good) scores on the "social functioning" subscale. Items which have such high ceiling effects are of questionable value in longitudinal research on health outcomes.

In contrast with older age groups, the HSQ-12 items detected relatively few positive responses indicating problems among younger age groups, which limits use as a population health survey instrument, as opposed to its use



Table 7 Health status questionnaire 12 (HSQ-12) subdomains (mean (SD)) by type of longstanding illness (based on those with just one condition)

HSQ-12 subdomain	Neoplasms & benign growths Mean (SD)	Endocrine & metabolic Mean (SD)	Mental disorders Mean (SD)	Nervous system Mean (SD)	Ear complaints Mean (SD)	Heart & circulatory system Mean (SD)	Respiratory system Mean (SD)	Digestive system Mean (SD)	Musculoskeletal system Mean (SD)	Other Mean (SD)	Total Mean (SD)
General health perceptions	61.0 (38.2)	55.6 (29.0)	43.2 (31.8)	62.0 (31.6)	73.3 (23.9)	42.0 (29.7)	54.6 (28.0)	56.6 (31.5)	57.7 (29.5)	68.7 (26.9)	55.7 (30.3)
Physical functioning	62.2 (44.7)	76.9 (31.6)	85.3 (24.3)	71.0 (39.8)	89.0 (16.8)	52.7 (39.2)	72.9 (35.1)	84.2 (26.2)	63.6 (36.2)	79.2 (33.8)	68.6 (36.2)
Role limitations:											
Physical	57.9 (44.4)	74.7 (36.5)	69.5 (34.7)	69.2 (38.3)	87.1 (28.8)	60.5 (38.7)	73.3 (38.5)	71.7 (38.0)	63.1 (37.7)	86.0 (28.5)	68.0 (37.7)
Mental	68.1 (44.2)	74.7 (36.4)	49.9 (32.7)	81.7 (27.2)	93.0 (20.1)	80.9 (31.0)	81.7 (25.8)	79.2 (29.3)	76.8 (32.5)	79.4 (30.5)	77.4 (31.7)
Social functioning	70.2 (40.8)	80.6 (35.0)	60.7 (37.5)	89.9 (22.6)	96.1 (12.5)	82.6 (29.1)	84.6 (24.5)	83.3 (27.5)	84.7 (28.9)	91.8 (22.4)	83.7 (28.9)
Pain	65.2 (35.5)	74.7 (30.0)	73.8 (26.1)	63.2 (28.9)	85.0 (23.4)	70.3 (27.7)	78.8 (26.2)	67.1 (30.1)	56.5 (29.3)	82.5 (20.7)	67.0 (29.5)
Energy/fatigue	35.4 (34.7)	55.9 (28.8)	40.0 (19.3)	60.6 (29.7)	71.6 (19.2)	43.8 (28.3)	45.7 (26.0)	54.8 (20.5)	48.8 (26.1)	54.4 (20.8)	50.7 (26.6)
Mental health	55.6 (26.4)	71.4 (23.3)	47.8 (23.2)	71.2 (18.9)	75.8 (15.7)	68.2 (21.6)	67.1 (18.8)	70.1 (19.3)	65.6 (22.2)	69.8 (15.5)	66.4 (12.6)
No of respondents	14	26	29-31	36	20	74-75	79-80	30	216-218	30-32	557-561

among older and patient populations. Differences in HSQ-12 mean scores and frequencies by construct (sex, age group, longstanding illness) were generally in the expected directions. While age was associated with most of the HSQ-12 domains, this association had interactions with longstanding illness or sex. The slight inconsistency in trends with age reflects results for other scales (eg, those measuring anxiety and depression<sup>14</sup>). The slight improvements in some scores among older age groups could reflect either lower expectations with increasing age which affects perceptions of health or the survival of the fittest.

There were clear differences in mean HSQ-12 scores and reporting, or not, of a longstanding illness, although the specificity of the scales was achieved at the expense of sensitivity. The discriminative ability of the HSQ-12 was evident when analysing the *type* of longstanding illness. The differences in mean HSQ-12 scores with type of longstanding illness made theoretical sense and these support the discriminative power of the scale.

Although item response was good with interviewer administration, it awaits testing by postal administration in order to assess whether item non-response increases. However, item non-response to the SF-36 has been reported to be problematic among older people whether interviewer or postal (self) administered,<sup>8</sup> which suggests that a trade off could be made between a shorter, simpler instrument with a consequent loss of some minor psychometric precision and the longer instrument with high rates of item non-response. The sensitivity of the HSQ-12 to changes over time also requires testing.

The results reported here support the use of the HSQ-12 with older populations, and particularly with specific chronic illnesses, although it will reveal relatively few problems among younger populations. It will require supplementation with more sensitive disease and/or domain specific scales in the areas of interest or intervention (without ceiling or floor effects), and is not recommended for use in isolation. It provides an acceptable core for more generally tapping some of the key domains of health related quality of life.

We would like to thank the omnibus survey staff at the Office for National Statistics for the data reported here, in particular Fiona Dawe, Angie Osborne and Paul Hunter; and Lesley Marriot for typing the tables. Those who carried out the original data collection and analysis hold no responsibility for the further analysis and interpretation of them. Material from the ONS omnibus survey, made available through the ONS, has been used with the permission of the Controller of HM Stationery Office and ONS. The dataset will be held on the ESRC Data Archive at the University of Essex from April 1998.

Funding: internal.  
Conflicts of interest: none.

- 1 Bergner M, Bobbitt, RA, Carter, WB *et al*. The sickness impact profile: development and final revision of a health status measure. *Med Care* 1981; 19:787-805.
- 2 British Geriatrics Society/Royal College of Physicians. *Standardised clinical instruments and measurement scales for elderly patients. Working Party report*. London: Royal College of Physicians, 1992.
- 3 Radosevich D, Pruitt M. *Twelve-item health status questionnaire. HSQ-12 Version 2.0*. Bloomington, MN: Health Outcomes Institute, 1995.

- 4 Health Outcomes Institute. *Twelve-item health status questionnaire (HSQ-12) version 2.0. user guide*. Bloomington, MN:Health Outcomes Institute, 1996.
- 5 Ware JE, Snow KK, Kosinski M, Gandek B. *SF-36 health survey: manual and interpretation guide*. Boston, MA:The Health Institute, New England Medical Centre, 1993.
- 6 Ware JE, Kosinski M, Keller, SD. *SF-12:How to score the SF-12. Physical and mental health summary scales*. Boston, MA:The Health Institute, 1995.
- 7 Brazier JE, Harper R, Jones NMB *et al*. Validating the SF-36 health survey questionnaire:new outcome measure for primary care. *BMJ* 1992; **305**:160-4.
- 8 Brazier JE, Walters SJ, Nicholl JP, Kohler B. Using the SF-36 and Euroqol on an elderly population. *Quality of Life Research* 1996; **5**:195-204.
- 9 Hill S, Harries U. Assessing the outcome of health care for the older person in community settings:should we use the SF-36? *Outcomes Briefing, UK Clearing House for the Assessment of Health Outcomes* 1994; **4**:26-7.
- 10 Ware JE, Sherbourne CD. The MOS 36-item short form health survey (SF-36):I. Conceptual framework and item selection. *Med Care* 1992; **30**:473-83.
- 11 Jenkinson C, Coulter A, Wright, L. Short form 36 (SF 36) health survey questionnaire:normative data for adults of working age. *BMA* 1993; **306**:1437-40.
- 12 Foster K, Jackson B, Thomas M *et al*. General household survey 1993. London:Office of Population Censuses and Surveys, HMSO, 1995.
- 13 Hayes V, Morris J, Wolfe C, Morgan M. The SF-36 health survey questionnaire:is it suitable for use with older adults? *Age Ageing* 1995; **24**:120-5.
- 14 Goldberg DP, Williams P. *A user's guide to the general health questionnaire*. Windsor:NFER-Nelson, 1988.