Factors affecting completion of the SF-36 in older people

STUART G. PARKER¹, DOROTHEA BECHINGER-ENGLISH², CAROL JAGGER³, NICOLA SPIERS³, JAMES LINDESAY⁴

¹Sheffield Institute for Studies on Ageing, University of Sheffield, Barnsley Hospital NHS Foundation Trust, Gawber Road, Barnsley S75 2EP, UK
²European Institute of Health and Medical Sciences (EIHMS), University of Surrey, Guildford GU2 7XH, UK
³Division of Epidemiology and Public Health,
⁴Division of Psychiatry, University of Leicester, 22-28 Princes Road West, Leicester LE1 6TP, UK

Address correspondence to: S. G. Parker. Fax: +44 (0) 114 2715771. Email: s.g.parker@sheffield.ac.uk

Abstract

Objective: to examine the influence of specific clinical impairments and disabilities on the completion of the SF-36 health status measure among older people.

Design: Prospective observational study.

Setting/Participants: An SF-36 was administered to 245 subjects aged 65 years and older. Subjects were chosen by sampling from a variety of inpatient, outpatient and community sources to ensure a range of relevant disabilities.

Measurements: response rates, overall rates of completion, completion of individual questions and time taken to complete.

Results: severe functional impairment (Barthel index ≤12) was found in 22.4% (51/228), cognitive impairment in 54.1% (132/244), depressed mood in 77.0% (151/196) and visuospatial dysfunction in 71.3% (134/188). The median number of impairments was three (interquartile range 1–4). Specific physical impairments were visual in 13.2% (31/235), hearing in 30.2% (74/245), impaired manual dexterity in 18.0% (44/245) and dysphasia in 23% (55/239). In multivariate analyses, global functional impairment (P = 0.006), cognitive impairment (P = 0.0001) and impaired manual dexterity (P = 0.005) were significantly associated with more dimensions uncompleted, whilst cognitive impairment (P = 0.001), age (P = 0.006) and visuospatial dysfunction (P = 0.0003) were significantly associated with longer completion times.

Conclusion: the most striking finding of the study was that global rather than specific physical and mental dysfunction was associated with the inability to complete the SF-36 questionnaire. The difficulty appears to lie in the performance of a complex task, rather than with specific aspects of the task which could be overcome by adaptation or aids. Our experience is that this relatively complex questionnaire does not adequately measure functional health status in disabled older people because of non-completion and may therefore overestimate the health of populations.

Keywords: health status, SF-36, disability, old age, elderly

Introduction

Older people with chronic disease and disability constitute the majority of healthcare consumers in the developed world. If evidence-based clinical practice is to be a reality for this group, it is important to conduct healthcare evaluations in such a way that we can be certain that they do what is claimed [1]. Measurement of outcomes of healthcare evaluations must therefore be conducted in a manner that is appropriate for the target population. The assessment of health status using standardised questionnaires or interview schedules is an important approach to measuring outcome in healthcare evaluations in old age. These tools are also used to describe the health of populations, and in this context, it is important that response rates are not biased by the presence of the very factors they are trying to measure. A basic requirement of a good patient outcome measure is that it can be collected for all patients with a minimum of missing data.

The SF-36 has been widely recommended for use in measuring health status, particularly in younger patients [2, 3], and older patients with simple pathology [4]. It performs well in community-based surveys of older people [5] and
has been proposed as a potential measure of patient outcome within the NHS [6]. It is a 36-item multiple-response questionnaire with eight dimensions. Dimension scores are calculated using a standard procedure [7], and if at least half of the questions on each scale record a response, inputted values are calculated for the items. Higher scores are associated with positive views of health status.

Response rates of 70–80% have been obtained using the SF-36 but are dependent on patient-related factors [8] such as age and dependency. Published experience of its use specifically in older people was until recently largely confined to community settings [9], ambulatory patients [10] or hospital patients with low levels of co-morbidity [11]. However, older patients with both physical and cognitive dysfunction may be less likely to complete and return a valid SF-36 schedule [12]. The objective of the present study was to identify specific, potentially ameliorable impairments and/or disabilities associated with poor completion of the questionnaire schedule of the SF-36.

Methods

The SF-36 was administered to elderly subjects with a high prevalence of one or more of six impairments, as outlined below. At the same time, the presence or absence of individual impairments was recorded and the impairment quantified.

Subjects were randomly allocated to receive either the standard SF-36 [13] or a version modified to improve the relevance of some of the questions to older people [10].

Local research ethics approval was obtained.

Subjects

Study subjects were drawn from five groups as follows:

- High prevalence of cognitive dysfunction: consecutive patients attending the Leicester Memory Clinic and a psychogeriatric day hospital.
- High prevalence of dysphasia: a random sample of stroke survivors who were inpatients at the Leicester General Hospital stroke assessment and rehabilitation unit.
- High prevalence of locomotor, visual and hearing impairment (inpatient group): a random sample of elderly patients on the rehabilitation unit for the elderly at Leicester General Hospital.
- High prevalence of locomotor, visual and hearing impairment (ambulatory group): consecutive new patients attending the day hospital for the Elderly at Leicester General Hospital.
- High prevalence of impaired cognition, vision, hearing, locomotor disability or dysphasia: a random sample drawn from a general practice list.

The aim was to ensure that subjects with a wide range of common disabilities were included. To avoid selection bias, random sampling was used in sources with an excess of potential subjects.

Data collection

Patients were assessed by one of two observers trained in the appropriate assessment techniques.

Every respondent was asked to self-administer the SF-36, if they felt able to do so.

The interviewer remained with the respondents while they completed the questionnaire. Her presence allowed respondents to ask for help or clarification. When the questionnaire was administered by interview, each question was repeated once, if necessary. Where the respondent could not give or decide on an answer, no answer was recorded.

Mode of completion of the SF-36 (self/interviewer/both) and time taken to complete were recorded. The individual SF-36 dimension scores were calculated according to the manual [7], together with the total number of dimensions completed and whether or not all dimensions had been completed.

Assessment of disabilities

Cognitive impairment: Mini Mental State Examination (MMSE) [14] (score <24), supplemented by a clock drawing test [15] (score >1).

Depressed mood: 30-item Geriatric Depression Scale (GDS) [16] (score >10).

Visual acuity: Snellen test [17] (score <15 in at least one eye).

Impaired hearing: Whisper test, validated against pure tone audiometry [18] (score <8).

Dysphasia: Boston aphasia severity rating scale (simplified version) from Boston Diagnostic Aphasia Examination [19] (score <5).

Impaired manual dexterity: The 9-hole peg test, scored as number of pegs placed after 50 s [20].


Statistical methods

The time taken to complete the SF-36 and the number of dimensions not completed were compared between those with and without each disability by the Mann–Whitney U test. The proportion failing to complete individual SF-36 dimensions was compared between those with and without each impairment by means of chi-squared tests. Given the number of tests performed, a level of 0.005 was taken to denote significance. Spearman’s rank correlation coefficient was calculated to investigate the association between the number of impairments and the time taken to complete the SF-36 and the number of dimensions not completed.

To investigate the joint contribution of impairments and demographic characteristics with failure to self-complete the SF-36, logistic regression models were fitted with failure to self-complete, time taken to complete and number of dimensions unsuccessfully completed as the dependent variables. A forward stepwise approach to model building was taken with variables entering or leaving the model with a significance level of 0.05. The same procedure was used with failure to have a valid SF-36 and with failure to complete individual dimensions of the SF-36. Multiple regression models were used similarly to the above for time taken to complete (after a logarithmic transformation) and number of dimensions not successfully completed.

Information on definitions of SF-36 completion, missing data and inter-observer agreement is provided in Appendix 1 on the journal website (http://ageing.oxfordjournals.org/).
Results

Characteristics of study subjects

The study population comprised 245 subjects (150 females), median age 78 years (IQR 72–86) who were administered an SF-36. Half (n = 119) the subjects lived alone and 2.5% (n = 6) lived in residential care.

Fifteen per cent of subjects (n = 36) were recruited from general practice, 20% (n = 50) from psychogeriatric services, 20% (n = 50) from a day hospital, 22% (n = 53) were stroke survivors and 24% (n = 60) rehabilitation ward inpatients.

Twenty-two per cent of subjects (n = 51) had a Barthel index score <13, 54% (n = 132) scored <24 on the MMSE and 23% (n = 45) scored <11 on the GDS. Thirteen per cent (n = 31) had visual impairment, 30% (n = 74) hearing impairment, 18% (n = 44) impaired manual dexterity, 71% (n = 134) visuospatial dysfunction and 23% (n = 55) dysphasia.

Missing data

Missing data resulting in non-completion of SF-36 questions/items are summarised in Table 1.

Modified questionnaire

No significant differences in completion rates were seen between the modified and unmodified questionnaire.

Time to complete

The SF-36 took subjects a median time of 16.5 min to complete (IQR 11–23).

Completion

Twenty-three per cent (n = 52) of the subjects self-completed the SF-36, 36% (n = 84) required the interviewer to complete and in 40% (n = 92) of subjects part was self-completed and part interviewer completed. The proportion of subjects not self-completing, the number of dimensions not complete and the proportion not completing a valid SF-36, all increased with the number of impairments (Table 1).

Twenty six per cent (n = 63) of subjects completed a valid SF-36. A median of six out of eight dimensions were successfully completed. Non-completion rates of individual dimensions of the SF-36 varied from 27% (mental health) to 46% (physical mental health). Non-completion was highest in the earlier questions (Physical Function and Role-Physical in particular).

Factors affecting completion

Timing

Longer time to complete the SF-36 in bivariate analyses was associated with cognitive impairment (P = 0.001), depressed mood (P = 0.02), visuospatial dysfunction (P < 0.001) and an increased number of impairments (P < 0.001). The final multiple regression model for time to complete the SF-36 included age (P = 0.006) and visuospatial dysfunction (P = 0.0003).

Self-completion

Inability to self-complete in bivariate analyses was associated with cognitive impairment (P < 0.001), visual impairment

| Table 1. Percentage (n) not self-completing and without a valid SF-36 and median (IQR) number of dimensions not completed and time taken to complete by impairment |
|-------------------------------------------------|-------------------------------------------------|-------------------------------------------------|-------------------------------------------------|
| Overall                                         | 77.2 (176)                                      | 2 (0–5)                                         | 16.5 (11–23)                                    | 74.3 (182)                                      |
| No impairments                                  | 50.0 (9/18)                                     | 0 (0–1)                                         | 14 (10–19)                                      | 36.8 (7/19)                                     |
| One impairment                                  | 61.4 (27/44)                                    | 1 (0–2)                                         | 13.5 (9–20)                                     | 54.5 (24/44)                                    |
| Two impairments                                 | 74.5 (38/51)                                    | 2 (1–4)                                         | 15 (11–21)                                      | 77.4 (41/53)                                    |
| Three impairments                               | 89.8 (44/49)                                    | 2 (1–5)                                         | 15 (12–23)                                      | 81.8 (43/53)                                    |
| Four or more impairments                       | 87.1 (61/70)                                    | 4 (2–6)                                         | 20 (15–25)                                      | 88.8 (72/81)                                    |
| Functional impairment                           | Yes                                            | 86.4 (38)                                       | 4 (2–7)*                                        | 92.2 (47)*                                      |
| No                                              | 74.9 (128)                                      | 2 (0–4)                                         | 16 (11–22)                                      | 68.4 (121)                                      |
| Cognitive impairment                            | Yes                                            | 87.2 (102)*                                     | 3.5 (2–6)*                                      | 87.9 (116)*                                     |
| No                                              | 66.4 (73)                                       | 1 (0–3)                                         | 15 (10–20)                                      | 58.0 (65)                                       |
| Depressed mood                                  | Yes                                            | 76.4 (107)                                      | 2 (1–5)                                         | 17 (12–23)*                                     | 77.5 (117)                                      |
| No                                              | 65.9 (29)                                       | 2 (0–6)                                         | 15 (10–20)                                      | 60.0 (27)                                       |
| Visual impairment                               | Yes                                            | 96.4 (27)                                       | 3 (1–5)                                         | 20 (16.5–24.5)                                  | 83.9 (26)                                       |
| No                                              | 74.4 (142)                                      | 2 (0–4)                                         | 15 (11–22)                                      | 71.6 (146)                                      |
| Hearing impairment                              | Yes                                            | 75.8 (50)                                       | 3 (1–5)                                         | 16.5 (11–23)                                    | 79.7 (59)                                       |
| No                                              | 77.8 (126)                                      | 2 (0–5)                                         | 16.5 (12–22)                                    | 71.9 (123)                                      |
| Impaired manual dexterity                       | Yes                                            | 92.1 (35)                                       | 6 (1.5–8)*                                      | 17 (12–23)                                     | 88.6 (39)                                       |
| No                                              | 74.2 (141)                                      | 2 (0–4)                                         | 16 (11–23)                                      | 71.1 (143)                                      |
| Visuospatial dysfunction                        | Yes                                            | 85.6 (107)*                                     | 3 (1–5)*                                        | 17 (13–24)*                                     | 80.6 (108)*                                     |
| No                                              | 56.6 (30)                                       | 0 (0–1)                                         | 13 (9–20)                                       | 44.4 (24)                                       |
| Aphasia                                         | Yes                                            | 91.7 (44)                                       | 3 (1–6)                                         | 17 (12–25)                                      |
| No                                              | 73.3 (129)                                      | 2 (0–4)                                         | 15.5 (11–22)                                    | 44.4 (24)                                       |

*P < 0.0001.
SF-36 in older people

(P = 0.009) and visuospatial dysfunction (P < 0.001), along with the total number of impairments (P < 0.001).

Logistic regression modelling included both cognitive impairment (P = 0.001) and visuospatial dysfunction (P = 0.005) in the final model, with the odds of failure to self-complete with cognitive impairment compared to no or missing cognitive impairment being 3.06 (95% CI 1.54–6.06), and with visuospatial dysfunction (compared to no such dysfunction) 2.59 (95% CI 1.34–5.04).

Completion of a valid SF-36

Failure to complete a valid SF-36 was associated with functional impairment (P < 0.001), cognitive impairment (P < 0.001), visuospatial dysfunction (P < 0.001) and an increased number of impairments (P < 0.001). In the final logistic regression model, functional impairment (P = 0.03) and cognitive impairment (P = 0.0001) were associated with lack of a valid SF-36. The odds of not having a valid SF-36 for subjects with functional impairment compared to those with no or missing functional impairment was 3.33 (95% CI 1.11–10.04) and with cognitive impairment 4.32 (95% CI 2.24–8.34).

Completion of individual dimensions

A significantly greater number of dimensions not being completed (P = 0.001) was associated in bivariate analyses with functional impairment, cognitive impairment, visuospatial dysfunction, impaired manual dexterity and an increased number of impairments. The multiple regression model for number of dimensions completed included functional impairment (P = 0.006), cognitive impairment (P = 0.0001) and impaired manual dexterity (P = 0.005).

Results of the modelling of the joint effects of impairments and age and sex on non-completion rates of each dimension are summarised in Table 2. Cognitive impairment was a significant factor in non-completion of all dimensions.

Discussion

In this study, we took a group of older people from a diverse range of primary and secondary healthcare settings, with a high prevalence of physical and mental disabilities and specific common impairments. We administered the SF-36 questionnaire and measured the nature and extent of their physical and mental capabilities. The presence of the interviewer and the availability of clarification, help or interview administration during completion of the SF-36 may have affected the absolute rate of non-response and completion and hence direct comparison of response and completion rates with for example postal survey results may not be reliable. However, the picture that emerged was one of a difficult-to-complete and time-consuming research measure which produced relatively few completed questionnaires overall. We used a modified version of the SF-36 in half of the subjects, but the modifications were not designed to overcome completion difficulties and were not shown to affect completion (within the limits afforded by the sample size); so the two forms were analysed together.

| Table 2. Final models of probability of non-completion of individual dimensions of SF-36 |
|--------------------------------------|-----------------|-----------------|
| Physical Function                   | Cognitive impairment | 2.33 (1.16–4.68) |
|                                     | Cognit             | 4.07 (2.32–7.16) |
| Role-Physical                       | Functional impair   | 2.89 (1.48–5.66) |
|                                     | Cognitive impairment | 3.47 (1.75–7.06) |
| Pain                                | Cognitive impairment | 3.43 (1.82–6.44) |
|                                     | Impaired manual dexterity | 2.58 (1.23–5.37) |
| General Health                      | Cognitive impairment | 3.06 (1.59–5.89) |
|                                     | Impaired manual dexterity | 2.43 (1.18–5.03) |
| Vitaliy                             | Functional impairment | 2.26 (1.14–4.49) |
|                                     | Cognitive impairment | 3.70 (1.89–7.23) |
|                                     | Depressed mood      | 0.40 (0.21–0.75) |
|                                     | Visuospatial dysfunction | 2.24 (1.19–4.24) |
| Social Functioning                  | Cognitive impairment | 3.56 (1.92–6.61) |
|                                     | Impaired manual dexterity | 3.06 (1.40–6.70) |
|                                     | Visuospatial dysfuncti | 2.10 (1.17–3.79) |
| Role-Emotional                      | Cognitive impairment | 2.66 (1.56–4.51) |
| Mental Health                       | Cognitive impairment | 4.07 (2.00–8.28) |
|                                     | Visual impairment    | 0.35 (0.13–0.94) |
|                                     | Impaired manual dexterity | 2.69 (1.25–5.80) |

*Odds of achieving floor/ceiling with impairment present compared to absent or missing and adjusted for other impairments shown.

A striking finding is that the difficulty the subjects experienced completing the SF-36 was reflected in global measures of physical and mental dysfunction, rather than the measures of specific disabilities or impairments. This implies that strategies designed to overcome specific common disabilities are unlikely to hold the key to improved completion of this relatively complex questionnaire.

We recruited subjects with a range of specific, measurable disabilities from multiple sources; so our study population is unlikely to be typical of any particular group of patients. Rather, they have been used to illustrate a general point about disabled populations of older people. While there may be limitations associated with some of the specific assessments used in this study, taken together, there is a consistent pattern of results between the global measures of physical and cognitive function and the specific measures of impairment in multiple domains, which provides an additional level of confidence in the conclusions drawn. Joint interviews were performed to attempt to secure consistency between interviewers, and levels of agreement were generally moderate to good.

Cut-off points were used to dichotomise measures of impairment for analysis (usually the score representing clinical importance, e.g. GDS [21], MMSE [22] and clock drawing [23]). For global impairment, a score of 13 or more on the Barthel index indicates that a disability was present in three or more of 10 activities of daily living [24], thus classifying subjects with mild or moderate disability (≥12 and <20) as ‘no impairment’. This may have some bearing on the relatively high levels of non-completion among these participants.

The principal conclusion of this study is that the complexity of the SF-36 completion task is too great for people with global physical or cognitive impairments. Consequently, relatively minor adjustments to questions or enhancements to the administration process such as altering layout, using larger print and revising the wording of existing
questions [25] are unlikely to produce major improvements to the completion of the questionnaire. This is likely to be particularly true among the target population for this study—disabled or frail older people in whom there are already questions about difficulty in obtaining responses [12] and validity in relation to clinical care [26]—our data suggest that completion rates for this health status measure may be directly affected by the health status (physical and cognitive function) of such respondents.

Alternative approaches to improving performance on standardised health status measures may be required. Interview administration can achieve high response rates even among severely disabled groups [27] but is not always feasible or affordable. One option may be to simplify the task by using simpler questionnaires. There is some support for this view in the literature on SF-36 completion in stroke patients, in whom the shorter, simpler EuroQol gave more complete data [28] and slightly better test–retest reliability, particularly on the mental health scale [29]. Alternatively, the SF-12 [30] is a relatively new derivation from the SF-36 which collapses the eight dimensions into just two and has been modified [31] and validated [32] for use in older people. The SF-36 physical and mental component summary scales (PCS-36 and MCS-36) capture about 85% of the reliable variance in the eight-scale SF-36. Twelve SF-36 items and improved scoring algorithms reproduced at least 90% of the variance in PCS-36 and MCS-36 in both general and patient populations. This shortened scale of course has the disadvantage that the range of health it describes is reduced to only two measures, but given our experience with the longer and more complex SF-36, this simplification may improve its usefulness for older patients with complex physical and mental disabilities by allowing more complete data capture, as does the EuroQol.

In conclusion, it is not possible on the basis of the evidence provided by this and other studies to recommend the SF-36 for use with populations of older people which may contain large subgroups with complex physical and mental disabilities. Even in general population studies which contain only a small proportion with complex disabilities, the SF-36 is likely to underestimate the prevalence of disability due to associated non-response. Nor is it possible from our data to recommend strategies to overcome specific disabilities interfering with completion. It may be that simpler questionnaires will give more complete and therefore usable data sets, but this hypothesis remains to be tested.

Key points

- The SF-36 is widely used as a measure of health status.
- Response rates tend to be poorest among older, frail subjects.
- This study shows that poor response rates relate to global measures of physical and cognitive function, rather than specific common impairments in old age.

Acknowledgements

The study was funded by the Trent Regional Health Authority.

Conflicts of Interests

None.

References


Received 28 September 2005; accepted in revised form 19 February 2006