Age and Ageing 1998; **27**: 13–18

Measuring health status in older patients. The SF-36 in practice

STUART G. PARKER, SUSAN M. PEET, CAROL JAGGER[1], MAHDI FARHAN[2], C. MARK CASTLEDEN

Division of Medicine for the Elderly, [1]Department of Epidemiology and Public Health and [2] Department of General Practice, University of Leicester, Leicester General Hospital, Gwendolen Road, Leicester LE5 4PW, UK

Address correspondence to: S. G. Parker. Fax: (+44) 116 2854666. email: sgp5@leicester.ac.uk

Abstract

Objective: to examine the use of the Medical Outcomes Study short form 36-item (SF-36) health status measure in older patients receiving health care; to explore the influence of age and physical and cognitive status on response to and completion of the SF-36 questionnaire.
Design: prospective observational study.
Settings: hospital and ambulatory patients aged 65 years and over in the medical and surgical wards and outpatient department of a teaching hospital and a local general practitioner's surgery.
Subjects: 1014 hospital inpatients, 80 hospital outpatients and 40 patients attending their general practitioner's surgery.
Main outcome measures: response rates, overall rates of completion (sufficient to calculate a valid SF-36 score) and completion of individual questions.
Results: 37 out of 40 ambulatory patients in general practice (93%) and 71 out of 80 outpatients (89%) returned a self-completed questionnaire. In hospital inpatients the overall response rate was only 46% (369 of 802). This was improved by interview administration to 77.4% (164 of 212). Logistic regression analysis revealed that self-completion, cognitive dysfunction, disability and age were all independently associated with poor overall response rates. Among those patients who returned a completed questionnaire, completion of individual questions sufficient to calculate a valid score was variable. Only 62.5% of inpatients who self-completed a questionnaire gave sufficient response to calculate a score on the mental health subscale, compared with 93.7% of general practice patients.
Conclusion: the self-administered questionnaire is unacceptable for older hospital inpatients. Use of an interviewer improves response, but factors which influence health status, such as physical and cognitive dysfunction, have a significant effect on response rates. Therefore the utility of the SF-36 in its present form as a routine health status measure for use in older hospital inpatients is questioned.

Keywords: *disability, elderly, health status, inpatients, quality of life, SF-36*

Introduction

Improving quality of life and reducing disability are identified as major priorities for the measured performance of geriatric medical services [1, 2]. Standardized clinical instruments and measuring scales may be used in routine practice to measure outcomes in the care of elderly patients [3]. In this context, measurement of health status using standard instruments is especially relevant. If such instruments are to be used to evaluate the impact of health care interventions, it is important that they can be used among typical patients. Dependency levels in residential care and in the community are rising as the population ages [4, 5], and multiple pathology and its attendant complex disability are commoner in older people [6], who are much more likely to be disabled in some way than younger patient groups [7]. These factors may not only worsen health status but also interfere with its measurement. For example, measurement of health status on a rehabilitation unit for older people has been said to be problematic—perhaps because of impaired communication and/or cognitive function [3].

The Medical Outcomes Study short form 36-item questionnaire (SF-36) was developed as part of a large series of outcomes studies in the USA [8]. It has recently been adapted for use in the UK and has been subjected to extensive validation in people aged <65 years [9, 10]. It is widely recommended as the instrument of choice for health status assessment in

health care settings [11]. Recently an anglicized version has become available that is reported to perform well in the under 65s [12-14]. Its use in older people has been evaluated in a community setting when administered by interview, with an 82% response rate and good (98%) data completeness [15]. It has been shown to be suitable for use with older surgical inpatients, who had similar global health perceptions to younger patients, although response rates were not stated [16]. However, experience of its use in ambulatory care has suggested that older people (>75 years) and those with poor physical or mental health have difficulty in returning self-completed questionnaires and that some of the specific questions are problematic [17].

This study examined the value of using the SF-36 in subjects aged 65 years and over who were medical or surgical inpatients and in smaller groups of outpatients and visitors to general practice surgeries. The aims of the study were to determine the proportion of older hospital inpatients for whom the SF-36 can provide health status data, the proportion of these data that can be used to derive valid scores and which questions pose particular problems for older patients. The study examined the extent to which physical and cognitive dysfunction or the method of administration influence rates of response and completion of the SF-36 in this group.

Methods

The hospital inpatients were participating in a survey of discharge processes and outcomes. A random, stratified sample was drawn of 10 wards: five medical, two surgical, one gynaecological and two orthopaedic. Consecutive patients aged 65 years and over admitted to each ward over an 8-month period were recruited by trained fieldworkers. Written, informed consent was obtained for all patients prior to participation in the study. All eligible consenting patients were interviewed whilst on the ward to provide baseline data on cognitive function and physical disability (see below). The SF-36 questionnaire was introduced to participants as a survey of their views about their health. If administered as an interview, the participant was advised that the response categories would be read out or shown and they should choose the category that most closely fitted how they had been feeling. Written instructions to this effect were provided on the questionnaire for those conducting the self-completed version. The SF-36 was usually completed towards the end of their illness or after their surgical procedure had been carried out. No time limit was stipulated for completion of either the self-administered or interviewer-completed questionnaire. In the case of the self-completed version, interviewers returned on the same day and/or the following and subsequent days until completed.

Responses to the self-completion version of the SF-36 (but not physical or cognitive function measures) were also obtained from patients attending in a consecutive series at a general practitioner's surgery or as medical outpatients (ambulatory care).

The SF-36 contains 36 items in nine separate domains. Eight of these comprise two or more items and the final one has only one item. Each domain is computed to give a scaled score of 0 to 100%. Incomplete domains can still be calculated providing at least 50% of items in the scale are present; in these circumstances an arithmetic mean is calculated from the completed items and used as an imputed value for the missing items [18].

Response and completion rates for the SF-36 were measured in three ways:

1. Overall response rate: the percentage of patients returning the SF-36 form.
2. Scale completion rate: the proportion of those patients returning the SF-36 who completed 50% or more items of a scale (allowing the calculation of a valid score [16]; this was calculated separately for each individual scale).
3. Item completion rate: the proportion of patients completing each individual question.

The following information was also collected for each of the hospital inpatients: age, gender, type of patient—categorized as medical or surgical (including gynaecology and orthopaedics), cognitive function score (the information/orientation subscale of the Clifton assessment procedures for the elderly) and a physical disability score, based on activities of daily living assessed by interview with the patient.

This physical disability score has been described previously [19]. Briefly, the patient's ability to dress, transfer from bed and chair, move around the house and use the toilet were rated as independent (score 0), needing an aid or help (score 1), needing an aid and help (score 2) or not able (score 3). The final disability score was derived by calculating [(1 × dressing) + (1 × bed transfer) + (1 × chair transfer) + (0.5 × mobility) + (0.25 × use of toilet)] and the scores were categorized as 0 = no disability, 0.25-1.00 mild/moderate disability and >1.00 = severe disability.

Step-wise logistic regression analysis was used to determine which (if any) of these additional variables significantly influenced overall response rate. At each step, every variable was assessed and the one which most significantly improved the discrimination between the two groups was included in the model. The process was terminated when no variable had a further significant effect on the discrimination ($P < 0.05$). SPSS PC for windows [20] was used and significant improvement in discrimination was determined by a change in deviance, measured by the Wald χ^2 statistic.

Results

Subjects

The SF-36 was administered to 1016 inpatients (554 women), median age 76 years [interquartile range (IQR) 71-82], 80 patients in medicine for the elderly outpatient clinics (48 women), median age 79 years (IQR 72-83) and 40 patients attending a general practitioner's surgery (23 female), median age 77 years (IQR 68-81). The first 212 inpatients were assessed using the interview schedule, the remainder received the version for self-completion.

Overall response rates

Overall response rates were significantly lower in the inpatient group (533/1016, 53%) compared with outpatients (57/80, 71%) and the general practitioner group (37/40, 93%; $P < 0.001$). In inpatients, administration of the SF-36 by interview resulted in significantly higher overall response rates (164/212, 77% *vs* 369/804, 46%; $P < 0.001$).

Factors influencing overall response rates

Subjectively, a number of barriers to completion of the SF-36 were observed. These included the patients feeling unwell or tired, lack of understanding of the concepts of multiple choice questionnaires, loss of concentration, interruptions on the ward and mechanical difficulties such as hearing, visual deficit and inability to hold a pen.

Objectively, patients were less likely to return an SF-36 with increasing age ($P < 0.001$), cognitive impairment ($P < 0.001$) or physical disability ($P < 0.001$; Mann-Whitney). Table 1 shows the effect of physical disability and cognitive dysfunction on overall response rates among inpatients, the lowest rates being seen in the most severely physically and cognitively impaired patients. Logistic regression analysis was performed with completion as the dependent variable and age, gender, cognitive function, physical function, type of patient (medical or surgical) and method of administration (interview or self) as independent variables. The result of this analysis for variables that showed an independent association with overall response rate is shown in Table 2. The method of administration had the largest influence, but even mild degrees of cognitive or physical dysfunction significantly affected response (Table 2). The effect of age appeared to be independent of physical and cognitive function and was most marked in the very old (85 years or older).

Table 2. Results of multiple logistic regression analysis: odds ratios (ORs) for response compared to non-response and 95% confidence intervals after adjustment for method of administration, age, cognitive function score and physical disability

	OR[a]	95% CI
Method of administration		
Interview	1.00	
Self-completion	0.13	0.08-0.20
Age (years)		
65-69	1.00	
70-74	0.79	0.52-1.21
75-79	0.79	0.51-1.22
80-84	0.72	0.46-1.14
85+	0.38	0.22-0.64
Cognitive function (CAPE-IOS score)		
12	1.00	
10-11	0.67	0.49-0.91
8-9	0.23	0.13-0.41
0-7	0.06	0.02-0.15
Missing	0.10	0.05-0.21
Physical disability score		
None	1.00	
Mild/moderate	0.60	0.40-0.87
Severe	0.56	0.41-0.77
Missing	0.45	0.15-1.37

[a] Odds ratio for response *vs* non-response adjusted for other variables shown.
CAPE-IOS, information/orientation subscale of the Clifton assessment procedures for the elderly.

Table 1. Overall response rate (%) in inpatients by level of physical disability and cognitive function

	No. of subjects, by physical disability			
Cognitive function	All	None ($n = 334$)	Mild/moderate ($n = 176$)	Severe ($n = 412$)
Cape-IOS score				
12	120	74	54	58
11	114	65	49	54
10	200	40	52	41
<10	498	35	30	20

CAPE-IOS, information/orientation subscale of the Clifton assessment procedures for the elderly.

Table 3. Scale completion rate: percentage of patients who returned a SF-36 satisfactorily completing 50% or more items for individual scales, allowing calculation of a valid score for that dimension

	% responding, by group					
			Inpatients			
SF-36 dimension	GP group (n = 37)	Outpatients (n = 57)	Questionnaire (n = 369)	Interview (n = 164)	χ^2	P
Physical functioning	91	91	73	95	58.90	<0.001
Physical role limitation	97	84	77	96	48.16	<0.001
Emotional role limitation	97	86	73	95	57.06	<0.001
Social functioning	100	89	92	94	5.65	<0.25
Mental health	94	86	62	75	38.65	<0.001
Vitality	97	87	85	95	17.30	<0.001
Pain	100	86	87	85	7.08	<0.1
General health perceptions	94	79	80	96	30.75	<0.001

GP, general practice.

Table 4. Non-response to individual questions, ordered by response rate

Question	No. of non-responders	Response rate (%)	Non-responses (%)
1	10	99	0.46
2	12	98.8	0.55
7	20	98	0.92
6	34	96.7	1.57
3b	39	96.2	1.8
3j	39	96.2	1.8
3c	42	95.9	1.94
3f	46	95.5	2.12
3g	53	94.8	2.44
3a	55	94.6	2.54
3e	56	94.5	2.58
9i	57	94.4	2.63
9f	60	94.1	2.77
3h	62	93.9	2.86
3i	62	93.9	2.86
8	62	93.9	2.86
3d	63	93.8	2.9
9c	63	93.8	2.9
9d	63	93.8	2.9
9h	64	93.7	2.95
9g	67	93.4	3.09
9e	68	93.3	3.14
9a	69	93.2	3.18
11b	69	93.2	3.18
4b	70	93.1	3.23
4d	71	93	3.27
9b	71	93	3.27
11a	75	92.6	3.46
11c	77	92.4	3.55
4a	85	91.6	3.92
4c	86	91.5	3.96
5b	91	91	4.2
10	94	90.7	4.33
5a	104	89.8	4.79
5c	110	89.2	5.07

Scale completion rates

Rates of completion for the eight scales of the SF-36 are shown in Table 3. In general, self-administration in inpatients was associated with poorer scale completion rates than interview administration in the same patient group or self-completion in ambulatory patients.

Item completion rates

Three questions, representing 11 scale items (30%) accounted for 41% of missing item responses (Table 4). Question 4 (14% of total non-responses) deals with problems with work or other daily activities as a result of physical health. Question 5 (14%) deals with problems with work or other daily activities as a result of emotional problems. Question 11 (14%) seeks perceptions of health status.

Discussion

The SF-36 has been widely recommended for use in measuring health status and has undergone evaluation and validation in a variety of populations. It has been shown to be of use in evaluating the impact of a wide range of clinical interventions, particularly in younger patients [21, 22] and older patients with simple pathology [23]. Response rates of 70–80% have been recorded for self-completion in older patients and are said to be dependant on patient-related factors such as age and dependency [24–30]. An anglicized version has recently become available which has been recommended as a potential measure of patient outcome within the UK's National Health Service [8–11]. Published experience of its use specifically in older people is largely confined to community settings [8], ambulatory patients [15] or hospital patients with low levels of co-morbidity [14]. Low response rates and problems with missing data have been noted in ambulatory care, particularly in the very elderly [15].

In the present study the self-completion version was shown not to be a useful instrument for routine use in hospital inpatients because of difficulty in obtaining adequate response rates and the influence of physical and cognitive co-morbidity on response. The overall response rate and scale completion rates for inpatients were improved by interview administration, but completion was still unacceptable in one-quarter of the patients. Higher response and completion rates were seen in ambulatory care and the overall response rates in these settings in our study are similar to those reported elsewhere [15], indicating that the SF-36 may be a useful tool in these settings where lower rates of physical and cognitive impairment exist.

The combination of low overall response rate and low rates of scale completion amongst responders makes interpretation of SF-36 scores difficult in this population and severely limits its potential clinical utility. Health status in older patients is influenced by a variety of factors, including the presence of specific impairments or disabilities [31]. The influence of physical disability and cognitive dysfunction on the response rate is important because it makes it more difficult to measure health-related quality of life in those patients with health-related disability and impaired cognition. In older people, concurrent use of physical disability and cognitive function scales, as here, helped to contribute to the assessment and revealed some aspects of health which themselves interfere with measurement of health status using the SF-36.

Other reservations about the use of the SF-36 in older patients concern the relevance of some of the individual questions to the experience of life of older people [14] and a possible lack of sensitivity to significant changes in health [32]. Experience from the present study with responses to individual questions is similar to that of Hayes *et al.* [15]. Questions 4 and 5, concerning problems with work or other daily activity as a result of physical health or emotional problems, appear to be particularly problematic. Question 11 also accounted for a significant proportion of non-completion; this question contains four items probing the patients' perceptions of their current health status.

The results of the present study are in accord with those from previous studies in which satisfactory response rates were obtained in relatively healthy elderly subjects. However, use in older people with physical and complex disabilities is problematic as aspects of their health may interfere with completion and interpretation of the results.

Key points

- The SF-36 is a health status instrument that has undergone extensive validation.
- Older people at home are able to complete and return the questionnaires.
- Disabled older patients, such as hospital inpatients, respond less reliably.
- Response rate is related to age, disability and cognitive impairment and is improved by interview administration.
- The self-administered questionnaire is unacceptable for older hospital inpatients.

References

1. Roberts H, Khee TS, Philp I. Setting priorities for measures of performance for geriatric medical services. Age Ageing 1994; 23: 154–7.

2. Roberts H, Philp I. Prioritizing performance-measures for geriatric medical-services—what do the purchasers and providers think? Age Ageing 1996; 25: 326–8.

3. Parker SG, Du X, Bardsley MJ *et al.* Measuring outcomes in care of the elderly. J Roy Coll Phys Lond 1994; 28: 428-33.

4. Lindesay J. The Guy's/Age Concern survey: physical health and psychiatric disorder in an urban elderly community. Int J Geriatr Psychiat 1990; 5: 171-8.

5. Campbell Stern M, Jagger C, Clarke M *et al.* Residential care for elderly people: a decade of change. Br Med J 1993; 306: 827-30.

6. Gosney M, Tallis R. The burden of chronic illness in local authority residential homes for the elderly. Health Trends 1990/91: 22: 153-7.

7. Martin J, Metzer H, Elliot D. The Prevalence of Disability among Adults. London. HMSO, 1988.

8. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992; 30: 473-83.

9. Jenkinson C, Coulter A, Wright L. Short form 36 (SF-36) health survey questionnaire: normative data for adults of working age. Br Med J 1993; 306: 1437-40.

10. Lyons RA, Fielder H, Littlepage BN. Measuring health status with the SF-36: the need for regional norms. J Public Health Med 1995; 17: 46-50.

11. Garrat AM, Ruta DA, Abdalla MI, Buckingham JK, Russell IT. The SF-36 health survey questionnaire: an outcome measure suitable for routine use within the NHS? Br Med J 1993; 306: 1440-4.

12. Ruta DA, Abdall MI, Garratt AM, Coutts A, Russell IT. SF-36 health survey questionnaire: I. Reliability in two patient based studies. Quality Health Care 1994; 3: 180-5.

13. Brazier JE, Harper R, Jones NM *et al.* Validating the SF-36 health survey questionnaire: new outcome measure for primary care. Br Med J 1992; 305: 160-4.

14. Garratt AM, Ruta DA, Abdulla MI, Russell IT. SF-36 health survey questionnaire: II. Responsiveness to changes in health status in four common clinical conditions. Quality Health Care 1994; 3: 186-92.

15. Lyons RA, Perry HW, Littlepage BNC. Evidence for the validity of the short form 36 questionnaire (SF-36) in an elderly population. Age Ageing 1994; 23: 182-4.

16. Mangione CM, Marcantonio ER, Goldman L *et al.* Influence of age on measurement of health status in patients undergoing elective surgery. J Am Geriatr Soc 1993; 41: 377-83.

17. 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.

18. Ware JE, Snow KK, Kosinski M, Gandek B. How to Score the SF-36 Health Survey. Boston MA: The Medical Outcomes Trust 1993.

19. Jagger C, Clarke M, Davies R. The elderly at home: indices of disability. J Epidemiol Comm Health 1986; 40: 139-42.

20. Norussis M. SPSS for Windows: base system user's guide, release 6.0. Chicago: SPSS, 1993.

21. Mahler DA, Mackowiak JI. Evaluation of the short-form 36-item questionnaire to measure health-related quality of life in patients with COPD. Chest 1995; 107: 1585-9.

22. Jenkinson C, Lawrence K, McWhinnie D, Gordon J. Sensitivity to change of health status measures in a randomized controlled trial: a comparison of the COOP charts and the SF-36. Quality Life Res 1995; 4: 47-52.

23. Bombardier C, Melfi CA, Paul J *et al.* Comparison of a generic and a disease-specific measure of pain and physical function after knee replacement surgery. Med Care 1995; 33 (suppl. 4): AS 131-44.

24. Boline PD, Kassak K, Bronfort G, Nelson C, Anderson AV. Spinal manipulation *vs* amitriptyline for the treatment of chronic tension-type headaches: a randomized clinical trial. J Manipulative Physiol Ther 1995; 18: 148-54.

25. Wagner AK, Keller SD, Kosinski M *et al.* Advances in methods for assessing the impact of epilepsy and antiepileptic drug therapy on patients' health-related quality of life. Quality Life Res 1995; 4: 115-34.

26. Lacey KO, Meier GH, Krumholz HM, Gusberg RJ. Outcomes after major vascular surgery: the patients' perspective. J Vasc Nurs 1995; 13: 8-13.

27. Gliklich RE, Hilinski JM. Longitudinal sensitivity of generic and specific health measures in chronic sinusitis. Quality Life Res 1995; 4: 27-32.

28. Johnson PA, Goldman L, Orav EJ, Garcia T, Pearson SD, Lee TH. Comparison of the Medical Outcomes Study Short-Form 36-Item Health Survey in black patients and white patients with acute chest pain. Med Care 1995; 33: 145-60.

29. Meyer KB, Espindle DM, DeGiacomo JM, Jenuleson CS, Kurtin PS, Davies AR. Monitoring dialysis patients' health status. Am J Kidney Dis 1994; 24: 267-79.

30. McHorney CA, Ware JE Jr, Raczek AE. The MOS 36-item short-form health survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993; 31: 247-63.

31. Carabellese C, Appolonio I, Rozzini R *et al.* Sensory impairment and quality of life in a community elderly population. J Am Geriatr Soc 1993; 41: 401-7.

32. Hill S, Harries U. The outcomes process: some reflections from research with people in their 60s and 70s. Crit Rev Public Health 1993; 4: 21-8.

Received 29 January 1997