Short form 36 (SF 36) health survey questionnaire: normative data for adults of working age

Crispin Jenkinson, Angela Coulter, Lucie Wright

Abstract

Objectives—**To gain population norms for the short form 36 health survey questionnaire (SF 36) in a large community sample and to explore the questionnaire's internal consistency and validity.**

Design—**Postal survey by using a booklet containing the SF 36 and several other items concerned with lifestyles and illness.**

Setting—**The sample was drawn from computerised registers of the family health services authorities for Berkshire, Buckinghamshire, Northamptonshire, and Oxfordshire.**

Sample—**13 042 randomly selected subjects aged 18-64 years.**

Main outcome measures—**Scores for the eight health dimensions of the SF 36.**

Results—**The survey achieved a response rate of 72% (n=9332). Internal consistency of the different dimensions of the questionnaire was high. Normative data broken down by age, sex, and social class were consistent with those from previous studies.**

Conclusions—**The SF 36 is a potentially valuable tool in medical research. The normative data provided here may further facilitate its validation and use.**

Introduction

One of the most dramatic developments in health care in the past 10 years has been the increasing consensus about the importance of subjective accounts of health in monitoring medical outcomes.[1 2] Traditional measures of morbidity and mortality are generally agreed to be too narrow to measure the potential benefits of health care interventions that can influence a wide number of variables such as physical mobility, emotional wellbeing, social life, and overall wellbeing.[3] Researchers have developed many questionnaires which are specific for various illnesses and are intended to tap the various domains ill health can affect,[4-6] but the search for a generic questionnaire that is easy to administer, acceptable to patients, and short as well as being fully validated has been a venture with few successes. One of the more widely used questionnaires has been the Nottingham health profile.[7 8] This questionnaire contains 38 items, is acceptable to patients, and requires only a few minutes to complete. It has, however, been criticised because of its inability to detect low levels of disability that are important not only clinically but also to respondents.[9]

A new measure has recently been introduced into this country that was developed from the Rand Corporation's health insurance experiment in the United States of America. The purposes and methods of the Rand study have been fully summarised elsewhere.[10 11] The new questionnaire, developed from an original lengthy battery of questions, is the short form 36 (SF 36), a questionnaire containing 36 items covering nine variables.

The purpose of the present paper is to provide population norms for the SF 36 from a large scale community sample. Community norms are important as they provide a base level of results on the questionnaire. As the designers of the SF 36 have themselves stated, the problem with scale scores is that they do not lend themselves to easy interpretation.[12] This is especially true in questionnaires such as the SF 36, in which certain items are weighted. The designers suggest, however, that when normative data sets exist these can be used for comparison purposes with other populations and samples.[12] To date only one study has provided any normative data for the SF 36 in Britain. The purpose of that study, however, was not primarily to gain normative data and hence the sample size was smaller than the one we report here.[3] As well as providing normative data we also studied the internal consistency and validity of the questionnaire and outline the possible limitations of which potential users must be aware.

Methods

The SF 36 is a short questionnaire with 36 items which measure eight multi-item variables: physical functioning (10 items), social functioning (two items), role limitations due to physical problems (four items), role limitations due to emotional problems (three items), mental health (five items), energy and vitality (four items), pain (two items), and general perception of health (five items). There is a further unscaled single item on changes in respondents' health over the past year. For each variable item scores are coded, summed, and transformed on to a scale from 0 (worst possible health state measured by the questionnaire) to 100 (best possible health state). Minor modifications to the wording of six items on the SF 36 were made to make it acceptable for British subjects. The changes are similar to those reported in the Sheffield study.[3]

Our results are based on data gained from the Oxford healthy life survey. This was a postal survey in which the SF 36 and questions on lifestyle and demographics were incorporated into a booklet. A covering letter, explaining the purpose of the study, was sent with the questionnaire. To those who did not respond to the initial questionnaire a reminder on a postcard was sent about four to six weeks later. If this elicited no response then another questionnaire and covering letter were sent.

In addition to the SF 36 the questionnaire booklet contained questions on whether or not the respondents had any long standing illness and also whether they had consulted a medical practitioner in the past two weeks because of problems with their health.

The questionnaire booklet was sent to 13 042 randomly selected subjects, unstratified by age or sex,

Health Services Research Unit, Department of Public Health and Primary Care, University of Oxford, Radcliffe Infirmary, Oxford OX2 6HE
Crispin Jenkinson, *research officer in health outcomes*
Angela Coulter, *deputy director*
Lucie Wright, *research officer*

Correspondence to: Dr Jenkinson.

BMJ 1993;306:1437-40

aged 18-64 years from the computerised registers of the family health services authorities for Berkshire, Buckinghamshire, Northamptonshire, and Oxfordshire.

STATISTICAL ANALYSIS

Normative data are provided in the form of means and standard deviations broken down by age, sex, and social class. Not all of the questionnaires returned had every item on the SF 36 answered. The developers of the SF 36 suggest a method of gaining scores for missing values but owing to the sample size and as the purpose of this paper is to provide normative data we decided not to use this protocol: hence missing data are not substituted. The results are based on the eight multi-item scales of the SF 36. Internal consistency—that is, the extent to which there is correlation between items on a scale—was assessed by Cronbach's α, an interitem correlation statistic with a range of 0-1.[13] Higher values indicate items on a dimension are correlated and therefore that the scale is tapping an underlying single variable on the questionnaire. Construct validity—that is, the extent to which the questionnaire supports predefined hypotheses— was assessed by determining the extent to which scores on different variables reflected the expected distribution of health status for and between certain groups. As has been suggested elsewhere,[3] scores might be expected to be lower, reflecting poorer perceived health, for women than for men, for those in lower social classes, for those who report a longstanding illness, and for those who have had a medical consultation in the recent past. We used *t* tests to analyse whether SF 36 scores differed significantly between these groups. Furthermore, the construct validity can be assessed by a correlational matrix of items against scales. Items would be expected to correlate most strongly with the variables to which they contribute and less strongly to others in the questionnaire.

Results

Completed questionnaires were obtained from 9332 of 13 042 people originally contacted, giving a response rate of 72%.

Sample characteristics—Table I shows the response rates to the survey broken down by age and sex. The proportion of returned questionnaires was high for a postal survey. To ensure the sample was representative, however, its composition was compared with population data; the sociodemographic characteristics of the sample were shown to mirror closely the characteristics of the general population when compared with 1991 population estimates and the social class distribution in the 1981 census.[13] Data on social class, determined by the present occupation of the respondent or if retired by their last full time occupation, were obtained from 8242 (88·3%) of the respondents and coded by using the registrar general's standard occupational classification.[14] Table II provides data on the social class composition of the sample.

The internal consistency of the variables in the SF 36 was assessed with the Cronbach's α statistic.[15] Table III shows the results. Internal consistency was good; an α of 0·5 or above is usually deemed acceptable,[16] although Nunnally recommends values of 0·7 and above.[17] In practice, however, well used scales should ideally gain α values in excess of 0·8.[18] Such high α values were gained for all dimensions of the SF 36 except for the social functioning scale (α=0·76); considering the small number of items in this dimension the result is acceptable. Similar results were found in the Sheffield study.[3] Breaking the data down into social classes revealed similar α values, as did an analysis of the internal consistency of variables of the SF 36 broken down by those who reported long term illnesses and those who did not.

TABLE III—*Chronbach's α statistic for variables of SF 36 questionnaire*

Variables	No of subjects	α
Physical functioning	8883	0·90
Social functioning	9219	0·76
Role limitations:		
Physical	9151	0·88
Emotional	9159	0·80
Mental health	9014	0·83
Energy/vitality	9089	0·85
Pain	9214	0·82
General health perceptions	9085	0·80

Normative data—Lower scores on the SF 36 reflect poorer health. Tables IV and V provide normative data in the form of means and standard deviations broken down by age and sex. Overall, women reported poorer health on all variables of the SF 36 than did men ($p<0·001$) except for general health perception. Table VI provides normative data for both sexes broken down by social class; a significant difference ($p<0·001$) was found on scores on all variables of the SF 36 between manual (classes III manual, IV, and V) and non-manual respondents (classes I, II, and III non-manual). Significant differences for the sample as a whole ($p<0·001$) were found between scores for women and men on all variables of the questionnaire. The data were broken down by those reporting long standing illness and those who reported a medical consultation in the two weeks before completing the questionnaire (table VII). In both cases significant differences were found between the groups, with those with long standing illness gaining significantly lower scores than those who did not report long standing illness ($p<0·001$ on all dimensions), and those who had consulted a doctor in the preceding two weeks gaining significantly lower scores than those who had not ($p<0·001$). These data, which support the findings of the Sheffield study,[3] provide evidence for the construct validity of the questionnaire. This was further supported by correlations for items against scale scores being highest for items against the variable to which they contribute (data not reported but available on request).

TABLE I—*Rate of response to SF 36 questionnaire broken down by age and sex*

	Age (years)					
	18-24	25-34	35-54	45-54	55-64	Total
No (%) of men	587 (60)	992 (60)	1039 (68)	879 (70)	732 (74)	4229 (66)
No (%) of women	832 (74)	1274 (75)	1232 (79)	979 (77)	786 (80)	5103 (77)
Total (%)	1419 (67)	2266 (68)	2271 (74)	1858 (74)	1518 (77)	9332 (72)

TABLE II—*Distribution of social class determined by Registrar General's standard occupational classification among subjects who responded to SF 36 questionnaire*

	Occupational class						
	I	II	III non-manual	III manual	IV	V	Unclassified*
No (%) of men (n=3991)	298 (3·3)	1416 (15·8)	414 (4·6)	1232 (13·8)	406 (4·5)	76 (0·9)	149 (1·7)
No (%) of women (n=4949)	92 (1·3)	1222 (13·7)	1894 (21·2)	397 (4·4)	604 (6·8)	191 (2·1)	549 (6·1)

*Could not be coded by using classification—for example, those who cared for home and children.

TABLE IV—*Mean (SD) scores and sample sizes for eight variables of SF 36 for men by age*

Variable	Age (years) 18-24	25-34	35-54	45-54	55-64
Physical functioning	92·8 (16·8) n=545	93·9 (14·2) n=929	91·9 (14·5) n=993	87·9 (17·4) n=815	80·0 (22·1) n=681
Social functioning	90·2 (16·4) n=556	91·3 (16·3) n=935	90·5 (17·0) n=1009	89·8 (18·7) n=844	86·9 (22·6) n=729
Role limitations:					
Physical	91·8 (22·6) n=557	92·0 (23·2) n=928	89·5 (25·5) n=1008	87·6 (28·3) n=841	78·8 (36·1) n=717
Emotional	82·9 (31·1) n=560	87·1 (27·9) n=931	86·0 (28·6) n=1008	85·7 (29·5) n=843	85·8 (29·9) n=714
Mental health	74·8 (15·4) n=548	75·8 (15·2) n=927	75·0 (16·1) n=990	76·0 (16·7) n=822	78·0 (17·3) n=697
Energy/vitality	66·4 (17·1) n=553	64·5 (17·3) n=927	63·5 (18·6) n=1001	62·9 (19·9) n=837	62·9 (20·3) n=707
Pain	86·6 (17·9) n=555	87·5 (17·7) n=934	85·6 (19·7) n=1008	81·8 (22·2) n=843	78·8 (23·6) n=724
General health perceptions	72·0 (20·1) n=838	76·7 (17·7) n=929	74·1 (18·5) n=998	72·0 (20·1) n=838	68·1 (22·9) n=709

TABLE V—*Mean (SD) scores and sample sizes for eight variables of SF 36 for women by age*

Variable	Age (years) 18-24	25-34	35-54	45-54	55-64
Physical functioning	90·1 (16·4) n=780	92·9 (13·3) n=1274	89·4 (16·1) n=1183	84·8 (18·3) n=917	74·8 (23·5) n=684
Social functioning	85·7 (19·7) n=791	87·1 (18·9) n=1294	86·7 (20·5) n=1210	87·0 (20·8) n=973	85·9 (22·6) n=783
Role limitations:					
Physical	88·6 (25·5) n=786	86·9 (29·2) n=1294	84·0 (32·0) n=1210	82·4 (32·0) n=960	76·6 (36·9) n=757
Emotional	78·8 (33·0) n=792	80·6 (34·0) n=1291	80·3 (33·6) n=1207	80·8 (33·6) n=965	83·3 (32·5) n=756
Mental health	70·2 (17·4) n=787	71·6 (15·2) n=1280	71·6 (17·8) n=1187	73·2 (18·2) n=950	74·4 (18·5) n=742
Energy/vitality	59·8 (19·4) n=784	58·3 (19·5) n=1269	58·2 (19·9) n=1200	59·4 (20·3) n=957	59·0 (21·4) n=763
Pain	81·7 (20·8) n=790	82·1 (21·1) n=1299	79·4 (22·0) n=1211	77·4 (22·3) n=965	75·0 (25·1) n=779
General health perceptions	72·1 (20·3) n=787	77·3 (18·5) n=1285	74·1 (20·3) n=1190	73·1 (19·9) n=950	68·0 (22·0) n=747

Discussion

The results of our study provide evidence for the internal consistency of the domains of the SF 36 and provide normative data, which will, we hope be useful to researchers who intend to use this new measure of health assessment. Brazier *et al* have suggested that the SF 36 is acceptable to patients, has high levels of internal validity, and good test-retest properties.[3] The questionnaire has advantages over, for example, the sickness impact profile,[19] in that it is considerably shorter, and the Nottingham health profile,[7 8] which has been found to be insensitive to lower levels of dysfunction and disability.[3 9] Results from the SF 36 have thus far been encouraging, but more research is required in Britain to determine its validity, reliability, and sensitivity to change—or "responsiveness" in patients with particular illnesses. At present there is little work on the last of these properties in health assessment questionnaires in general and the SF 36 in particular. This aspect, however, is increasingly seen as one of the important features of measures of health status, especially in assessments of clinical interventions.[20 21]

Clinicians and other users of questionnaires such as the SF 36 should also consider that "off the shelf" indices that have high levels of reliability, validity, and responsiveness may not necessarily be appropriate for what they want to measure. The SF 36, for example, contains no variable on sleep, and a treatment aimed at reducing disturbance of sleep may not be appropriately judged by using the SF 36 alone. In most cases it will be necessary to use not only a generic questionnaire but also one designed to tap variables specific to the particular case under study. Furthermore, not all questionnaires are appropriate for all age groups. Brazier *et al* reported a lower response rate for the SF 36 among those over 65 years, and suggested that different questionnaires may be required for this group.[3] There are certainly several questionnaires designed specifically to assess the health of this age group, and they may be more appropriate.[5] Gaining a summary score or single index figure from this questionnaire is not possible, yet while questionnaires that can be summed to provide a single figure permit easy comparisons between different groups of patients, they may not always tap dimensions appropriate for particular groups. Existing questionnaires that result in a single index have yet to be validated to ensure their appropriateness for this task.[22]

If used with consideration to the points raised above the SF 36 is a potentially useful tool for measuring health status. We would advise users to retain the format used in the Sheffield study[3] and the one reported here. There are presently three versions available: the original questionnaire made available

TABLE VI—*Mean (SD) scores and sample sizes for eight variables of SF 36 by standard occupational classification*

Variable	Occupational classification I	II	III non-manual	III manual	IV	V
Physical functioning	93·4 (11·7) n=387	90·4 (15·4) n=2558	88·9 (16·5) n=2274	87·6 (19·5) n=1566	87·6 (18·8) n=988	84·3 (21.3) n=236
Social functioning	91·0 (16·7) n=391	89·0 (18·2) n=2620	87·8 (19·6) n=2299	88·6 (19·3) n=1616	87·6 (19·9) n=1007	85·7 (21·3) n=267
Role limitations:						
Physical	89·9 (25·2) n=390	86·7 (28·8) n=2626	86·1 (29·3) n=2284	87·1 (29·1) n=1608	84·6 (31·8) n=991	82·8 (33·0) n=258
Emotional	87·3 (26·5) n=391	83·1 (31·6) n=2625	82·5 (32·0) n=2283	85·9 (29·3) n=1616	82·1 (32·7) n=990	79·7 (34·5) n=260
Mental health	76·6 (14·7) n=384	75·2 (16·2) n=2585	72·9 (17·2) n=2252	74·6 (17·3) n=1573	72·9 (18·0) n=978	70·8 (20·0) n=258
Energy/vitality	63·7 (18·8) n=389	61·9 (19·0) n=2595	60·1 (19·4) n=2267	62·0 (20·1) n=1599	60·8 (20·0) n=987	58·7 (20·3) n=258
Pain	88·2 (16·2) n=387	83·0 (20·2) n=2620	81·2 (21·7) n=2296	81·6 (22·4) n=1617	80·0 (23·0) n=1008	78·6 (23·2) n=266
General health perceptions	75·1 (17·8) n=387	75·2 (18·6) n=2587	74·3 (19·5) n=2274	72·8 (20·1) n=1604	72·1 (20·8) n=988	70·3 (21·2) n=255

TABLE VII—*Mean (SD) scores and sample sizes for eight variables of SF 36 by self reported long standing illness and consultation with doctor in two weeks before completion of questionnaire*

Variable	Respondents not reporting long standing illness	Respondents reporting long standing illness	Respondents who did not report medical consultation	Respondents who did report medical consultation
Physical functioning	92·5 (13·4) n=6301	78·3 (23·2) n=2489	89·9 (16·2) n=7245	81·6 (23·0) n=1596
Social functioning	91·3 (15·8) n=6506	80·2 (24·8) n=2621	90·5 (16·9) n=7524	76·9 (26·0) n=1671
Role limitations:				
Physical	91·4 (23·2) n=6468	71·9 (38·9) n=2509	90·0 (24·9) n=7461	66·9 (41·3) n=1642
Emotional	85·6 (29·3) n=6485	76·3 (36·4) n=2506	85·2 (38·6) n=7472	72·7 (29·6) n=1643
Mental health	75·4 (16·3) n=6367	69·9 (18·7) n=2553	75·1 (16·3) n=7357	68·0 (19·9) n=1633
Energy/vitality	64·0 (18·2) n=6403	54·0 (21·1) n=2590	63·0 (18·7) n=7415	52·9 (21·9) n=1650
Pain	86·3 (17·9) n=6497	69·8 (25·4) n=2615	84·6 (18·8) n=7529	67·7 (27·6) n=1662
General health perceptions	78·8 (15·7) n=6402	60·8 (23·0) n=2584	75·7 (18·2) n=7417	63·7 (23·7) n=1644

from InterStudy (now called the Health Outcomes Institute) in America, on which the anglicised versions used in the Sheffield study and the Oxford healthy life study were based, and two more recent versions. One of these available from John Ware at the New England Medical Center Hospital[2] and the other from the Rand Corporation.[23] The latter versions have thus far not been validated for use in England and, despite the relatively minor changes between these and the original questionnaire, we would suggest that users continue to use the original InterStudy version, adapted for use in England.

We thank Sue Bradshaw, Peter Brooks, and Diana Harwood for help with the administration of the survey, Sue Ziebland for comments on an earlier draft of the paper, Berkshire, Buckinghamshire, Northamptonshire, and Oxfordshire Family Health Services Authorities for help in drawing the sample, and Oxford Regional Health Authority for funding the study.

1 Geigle R, Jones S. Outcomes measurement: a report from the front. *Inquiry* 1990;27:7-13.
2 Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36) 1: conceptual framework and item selection. *Med Care* 1992;30: 473-83.
3 Brazier JE, Harper R, Jones NMB, O'Cathain A, Thomas KJ, Usherwood T, *et al.* Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ* 1992;305:160-4.
4 Wilkin D, Hallam L, Doggett M. *Measures of need and outcome for primary care.* Oxford: Oxford University Press, 1992.
5 McDowell I, Newell C. *Measuring health: a guide to rating scales and questionnaires.* New York: Oxford University Press, 1987.
6 Bowling A. *Measuring health: a review of quality of life measurement scales.* Milton Keynes: Open University Press, 1991.
7 Hunt S, McEwen J, McKenna SP. Measuring health status: a new tool for clinicians and epidemiologists. *J R Coll Gen Pract* 1985;35:185-8.
8 Hunt S, McEwen J, McKenna SP. *Measuring health status.* London: Croom Helm, 1986.
9 Kind P, Carr-Hill R. The Nottingham health profile: a useful tool for epidemiologists? *Soc Sci Med* 1987;25:905-10.
10 Ware JE, Brook RH, Williams KN, Stewart AL, Davies-Avery A. *Conceptualisation and measurement of health for adults in the health insurance study.* Vol 1. Model of health and methodology. Santa Monica, California: Rand Corp, 1980.
11 Stewart AL, Ware JE, eds. *Measuring functioning and well being: the medical outcomes study approach.* London: Duke University Press, 1992.
12 Stewart AL, Hays RD, Ware JE. Methods of validating the MOS health measures. In: Stewart AL, Ware JE, eds. *Measuring functioning and well being: the medical outcomes study approach.* London: Duke University Press, 1992.
13 Wright L, Harwood D, Coulter A. *Health and lifestyles in the Oxford region.* Oxford: Health Services Research Unit, 1992.
14 Office of Population Censuses and Surveys. *Standard occupational classification. Vol 2.* London: HMSO, 1990.
15 Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297-334.
16 Helmstadter GC. *Principles of psychological measurement.* New York: Appleton-Century-Crofts, 1964.
17 Nunnally JC. *Psychometric theory.* 2nd ed. New York: McGraw Hill, 1978.
18 Carmines E, Zeller R. *Reliability and validity assessment. Quantitative applications in the social science.* Beverley Hills: Sage, 1979.
19 Bergner M, Bobbitt RA, Carter WB, Gilson BS. The sickness impact profile: development and final revision of a health status measure. *Med Care* 1981;19:787-805.
20 Fitzpatrick R, Ziebland S, Jenkinson C, Mowat A, Mowat A. The importance of sensitivity to change as a criterion for selecting health status measures. *Quality in Health Care* 1992;1:89-93.
21 Ziebland S, Fitzpatrick R, Jenkinson C. Assessing short term outcome. *Quality in Health Care* 1992;1(2):141-2. (Letter.)
22 Carr-Hill R, Morris J. Current practice in obtaining the "Q" in QALYs: a cautionary note. *BMJ* 1991;303:699-701.
23 Rand Health Sciences Program. *Rand 36-item health survey 1.0.* Santa Monica: Rand, 1992.

(Accepted 16 March 1993)

The SF 36 health survey questionnaire: an outcome measure suitable for routine use within the NHS?

Andrew M Garratt, Danny A Ruta, Mona I Abdalla, J Kenneth Buckingham, Ian T Russell

Abstract

***Objective*—To assess the validity, reliability, and acceptability of the short form 36 (SF36) health survey questionnaire (a shortened version of a battery of 149 health status questions) as a measure of patient outcome in a broad sample of patients suffering from four common clinical conditions.**

***Design*—Postal questionnaire, followed up by two reminders at two week intervals.**

***Setting*—Clinics and four training practices in north east Scotland.**

***Subjects*—Over 1700 patients aged 16-86 with one of four conditions—low back pain, menorrhagia, suspected peptic ulcer, or varicose veins—and a comparison sample of 900 members of the general population.**

***Main outcome measures*—The eight scales within the SF36 health profile.**

***Results*—The response rate exceeded 75% in the patient population (1310 respondents). The SF36 satisfied rigorous psychometric criteria for validity and internal consistency. Clinical validity was shown by the distinctive profiles generated for each condition, each of which differed from that in the general population in a predictable manner. Furthermore, SF36 scores were lower in referred patients than in patients not referred and were closely related to general practitioners' perceptions of severity.**

***Conclusions*—These results provide support for the SF36 as a potential measure of patient outcome within the NHS. The SF36 seems acceptable to patients, internally consistent, and a valid measure of the health status of a wide range of patients. Before it can be used in the new health service, however, its sensitivity to changes in health status over time must also be tested.**

Department of Public Health, University of Aberdeen, Foresterhill, Aberdeen AB9 2ZD
Andrew M Garratt, *research assistant*
Danny A Ruta, *lecturer*

Health Services Research Unit
Mona I Abdalla, *research fellow*
J Kenneth Buckingham, *research fellow*
Ian T Russell, *director*

Correspondence to: Mr Garratt.

BMJ 1993;306:1440-4

Introduction

The government has responded to the need for efficient provision of health care by introducing managed competition into the National Health Service.[1] In the internal market health authorities and fundholding practices may purchase care from competing provider units. If the reforms are to be successful, then purchasers will require valid, reliable, and sensitive measures of outcome to allow them to allocate scarce resources in the most cost effective manner. Similarly, service providers who can demonstrate the effectiveness of the care they provide in improving patient outcome will be better placed to compete for purchaser funds. Without such information health care will be purchased on the basis of cost alone, with serious consequences for its quality.

Few outcome measures currently available for routine use satisfy the criteria of validity, reliability, and sensitivity to changes in health status. For example, the Nottingham health profile has been criticised for failing to detect low levels of morbidity.[2] A measure that deserves careful consideration is the short form 36 (SF36) health survey questionnaire.[3,4] The SF36 is a shortened version of a battery of 149 health status questions developed and tested on a population of over 22 000 patients as part of the medical outcome study,[5,6] designed to help understand how specific components of the American health care system affect the outcomes of care. A key objective of the study was to develop more practical tools for monitoring patient outcomes in