

# Patient centred assessment of quality of life for patients with four common conditions

Danny A Ruta, Andrew M Garratt, Ian T Russell

## Abstract

**Objectives**—To assess the reliability, validity, and responsiveness of a new quality of life measure, the patient generated index (PGI) of quality of life, in patients with four common clinical conditions.

**Design**—Prospective one year follow up study.

**Setting**—Outpatient departments and four general practices in Grampian, Scotland.

**Subjects**—1746 patients consulting a general practitioner in one of four practices, or referred to outpatients from all Grampian practices over a four month period, with low back pain, menorrhagia, suspected peptic ulcer, and varicose veins.

**Main outcome measures**—Postal questionnaire including the PGI, SF-36 health survey, and clinically derived condition specific measures of disease severity.

**Results**—Test-retest reliability was satisfactory for group comparisons (intraclass correlation coefficient 0.65). Validity was confirmed by the observed association of the PGI with the SF-36, condition specific instruments, and sociodemographic variables. For low back pain, the PGI and the SF-36 pain scale were found to be most responsive to clinical change. For patients with menorrhagia and suspected peptic ulcer, only the condition specific instruments detected larger changes than the PGI.

**Conclusions**—It is possible to develop a patient generated index of quality of life that not only assesses the extent to which patients' expectations are matched by reality but also satisfies criteria of reliability and responsiveness to change. Further work is required to make the PGI more acceptable and meaningful to patients, but it is believed that it offers an exciting new approach to the evaluation of medical care.

(Quality in Health Care 1999;8:22-29)

Keywords: quality of life; health outcomes; patient generated index; SF-36

Improving "quality of life" has become a main focus of interest in medicine, yet most outcome measures assess much narrower concepts such as impairment and disability.<sup>1</sup> A true quality of life measure would focus less on the health factors necessary to sustain life, such as physical and psychological wellbeing, and more on the goals and expectations that give life its meaning and purpose. Calman proposes a definition of quality of life—"the extent to which our hopes

and ambitions are matched by experience".<sup>2</sup> He argues that to improve quality of life, health care should "narrow the gap between a patient's hopes and expectations and what actually happens".<sup>2</sup> This definition provides the conceptual basis for a measure that we have developed called the patient generated index (PGI) of quality of life.<sup>3</sup> The PGI incorporates two existing techniques developed in unrelated fields. The first was developed as part of a condition specific measure for use in chronic lung disease.<sup>4</sup> Patients were asked to choose the five most important areas of their lives affected by their breathing problems and to rate how badly affected they were in each area. The second technique, the priority evaluator method, was developed by Hoinville,<sup>5</sup> and has been used by town planners to take account of the preferences of future residents. In its original form, subjects are presented with a set of residential characteristics—for example garden size, parking facilities, and shopping amenities—and are asked to indicate their preferences for the kind of town they would like to live in, by distributing "points" between the different characteristics.

The PGI is completed in three stages as a self administered or interviewer administered questionnaire (appendix 1). Table 1 illustrates these stages with reference to a 30 year old man experiencing low back pain. In stage one, patients are asked to specify the five most important areas of life affected by their condition, which in this case is low back pain. Patients are told that the area might be small and personal to them, for example they may have difficulty playing with their children or might have a continual worry at the back of their mind. If they are unclear about the question, they are provided with a trigger list of the areas most frequently mentioned by patients with the same condition. If patients feel that fewer than five areas of their life are affected they are instructed to write "none" in the remaining boxes.

Stage two asks patients to rate how badly affected they are in each chosen area on a scale of 0 to 100, where 0 represents the worst they can imagine for themselves and 100 represents exactly as they would like to be. A sixth box is provided for them to rate all other areas of their life. This can include areas of their life affected by their medical condition but not important enough to be included in the top five boxes, as well as areas of their life which might be unrelated to their condition or even to their health. The second column of table 1 shows that the respondent's five areas, particularly his work, are considerably affected by his back pain, however he has few complaints about the rest

Department of  
Epidemiology and  
Public Health,  
Ninewells Hospital and  
Medical School,  
Dundee, UK  
D A Ruta, senior lecturer

Department of Health  
Sciences and Clinical  
Evaluation, University  
of York, UK  
A M Garratt, research  
fellow  
I T Russell, professor of  
health sciences

Correspondence to:  
Dr D Ruta, Department of  
Epidemiology and Public  
Health, Ninewells Hospital  
and Medical School, Dundee  
DD1 9SY, UK.

Accepted 11 December 1998

Table 1 Stages in the completion of the PGI of quality of life

Stage 1: area or activity affected		Stage 2: score out of 100	Stage 3: spend your 60 points	Calculation of final score
1	Work suffers	10	×	10/60 = 1.7
2	Makes me moody	30	×	10/60 = 5.0
3	Always thinking	30	×	5/60 = 2.5
4	Can't play with kids	50	×	20/60 = 16.7
5	My sex life suffers	70	×	10/60 = 11.7
6	All other aspects of your life not mentioned above	90	×	5/60 = 7.5
				45.1

of his life which is quite close to how he would like it to be.

In the final stage, patients are asked to imagine that they can improve some or all of their chosen areas. They are given "points" to "spend" across one or more areas that they would most like to improve. The points they give to each area are taken to represent the relative importance of potential improvements in that area. Patients are asked to imagine that areas receiving no points remain unimproved. In table 1, although the man's ability to play with his children is less affected than his work, he places greater value on an improvement in this area and so gives this area more points. To generate an index, the ratings for each area are multiplied by the proportion of points awarded to that area and summed to give a score between nought and 100. In the example that table 1 shows, a score of 45.1 was obtained. This score is intended to represent the extent to which reality matches expectations—that is, perceived quality of life—in those areas of life in which patients most value an improvement.

The PGI has been validated in patients with low back pain.<sup>3</sup> In this article we assess the PGI for reliability, validity, and responsiveness to change in three other common conditions—menorrhagia, suspected peptic ulcer, and varicose veins—and compare the results with patients with back pain.

## Methods

### DATA COLLECTION

Between March and June 1991 we identified patients in Grampian, Scotland presenting with one of the four conditions as part of a large prospective study called the Grampian Health Outcomes Study. The aim of this study was to examine four different approaches to health related quality of life measurement. Patient centred assessment using the PGI comprised one approach. Patients were identified in one of two ways: from referral letters to hospital outpatient departments within Grampian, and by general practitioners (GPs) from four large training practices. A questionnaire including the PGI was sent to patients in general practice within two weeks of their initial consultation and to referred patients before their first outpatient appointment. Two further reminders were sent to non-responders. Patients not wishing to take part in the study were asked to return their questionnaires blank. Patients taking part in the study were sent a follow up questionnaire at six months and one year after completing the first questionnaire. Only one year follow up data are reported here. At follow

up, patients were asked to complete a blank PGI questionnaire without being reminded of their previously chosen areas.

The questionnaire also included a well validated generic health measure, the SF-36 health survey questionnaire,<sup>6,7</sup> clinically derived condition specific measures of symptom severity, and sociodemographic questions. The four condition specific clinically derived measures were developed through consultation with local clinicians. The resulting items are designed to reflect the areas of health that clinicians consider to be important in the management of a particular condition. After the selection of individual items, the four sets of clinical questions were subjected to rigorous testing. This approach has been shown to produce condition specific measures which are valid, reliable, and responsive to change.<sup>8-12</sup>

Patients completing the PGI correctly were compared with those who omitted or did not correctly complete the PGI using stepwise logistic regression.<sup>13</sup> A PGI response was considered incomplete if all three stages were not completed, or if the allocation of points in stage three did not add up to 60, as this made it impossible to generate a final index score.

### VALIDATING THE PGI

To validate the PGI we assessed its reliability, validity, and responsiveness to changes in quality of life over time. Reliability is the extent to which measurements on the same individual under different circumstances are similar.<sup>14</sup> Test-retest is the most appropriate method for assessing reliability if the instrument in question is intended as an evaluative tool.<sup>14</sup> Validity is the extent to which an instrument measures what is intended,<sup>14</sup> in this case quality of life, as defined by the gap between individuals' expectations and their reality. Validity is usually assessed by correlating a new measure with established measures of the concept under study which is known as criterion validity.<sup>14</sup> However, in the absence of an established or "gold standard" measure of quality of life, we tested the PGI for construct validity—that is, the extent to which a new measure is related to specified variables in accordance with an established theory or "hypothetical construct".<sup>14</sup>

To the extent that health influences quality of life, we would expect a positive relation to exist between the PGI scores and health as measured by the eight scales of the SF-36 and by the clinical condition specific measures. If the PGI were a valid measure of quality of life we would expect significant correlations between PGI scores and the scores on the health measures. We would not expect the scores to be

Table 2 Response analysis: logistic regression analysis of differences between those completing the PGI correctly and those failing to complete or not attempting the PGI

Variable	Regression coefficient	Standard error	Wald statistic	Significance	Odds ratio
Age on leaving full time education	0.123	0.035	13.969	0.000	1.152
Physical functioning	0.014	0.003	17.052	0.000	1.014
Role: physical	-0.009	0.002	17.763	0.000	0.991
Energy and fatigue	-0.007	0.004	3.957	0.047	1.993
Patient divorced	-0.544	0.234	5.378	0.020	0.581
Home rented from the council	-0.502	0.158	10.023	0.002	0.606
Patient retired	-0.517	0.247	4.393	0.036	0.597
(Constant)	-1.328	0.605	4.816	0.028	—

too high, however, as this would imply that the PGI was merely measuring health.

If the PGI is to have potential as an outcome measure it must be shown to be responsive or sensitive to clinically important changes in quality of life. Several different methods have been put forward for assessing responsiveness.<sup>15-17</sup> If meaningful comparisons are to be made between instruments, then a standardised measure of responsiveness is required. The standardised response mean, which represents the mean change in score over two points in time divided by the standard deviation (SD) of the score differences, allows such comparisons to be made.<sup>17</sup>

#### RELIABILITY

Test-retest reliability was assessed by sending patients an additional questionnaire two weeks after they returned the one year follow up questionnaire. Reliability was estimated for those patients reporting no change in health using the intraclass correlation coefficient and the distribution of differences between the two sets of scores.<sup>18</sup>

#### VALIDITY

In the absence of a “gold standard” measure of quality of life, we tested the PGI for construct validity. We hypothesised a positive relation between PGI scores and scores on the measures of health shown by the Pearson correlation coefficient. Stepwise multiple regression was used to model the relation between the PGI and the health measures.<sup>13</sup> Sociodemographic variables and patient source were also included as potential explanatory variables. Further construct validation was done by testing hypothesised relations between PGI scores and various clinical and socio-demographic variables. The following hypotheses were tested about PGI scores: they would be higher in non-referred than in referred patients; related to symptom severity as perceived by the GP; lower in women; higher in patients owning their own homes; higher in married than single patients; and lower in unemployed patients. The questionnaires given to patients with low back pain and suspected peptic ulcer contained questions relating to medication and family history of ulcer disease, respectively. We hypothesised that the use of analgesics by patients with low back pain and the presence of a family history of ulcer disease in patients with a suspected peptic ulcer would be reflected in lower PGI scores.

#### RESPONSIVENESS

The standardised response mean was used to compare the relative responsiveness of the various measures.<sup>17-19</sup> Higher standardised response means indicate greater sensitivity to clinical change, with standardised response means of 0.2, 0.5, and 0.8 or above representing small, moderate, and large changes, respectively.<sup>17</sup> The sampling distribution (mean, SD) of the standardised response means were estimated using a jackknife procedure.<sup>20</sup> This allowed testing for statistically significant differences in standardised response means between instruments.

## Results

#### DATA COLLECTION

A total of 1317 of 1746 correctly identified patients took part, giving a response rate of 75.4%. Their ages ranged from 16 to 86 (mean 42.7 years) and 870 (66.1%) were women. The patients who did not respond were significantly younger (mean age 39.9 years;  $p < 0.01$ ). However, they did not differ significantly from respondents in sex, clinical condition, source, or symptom severity as reported by their general practitioners. Of the 1317 patients taking part in the study, 672 (51.0%) completed the PGI correctly and 106 (8.0%) indicated that no aspects of their life were affected by their condition. Their ages ranged from 16 to 86 (mean 41.1 years), and 460 were women (68.5%). After adjusting for confounding variables using logistic regression, seven significant differences were found between respondents completing the PGI correctly and those not attempting or not completing it (table 2). As the patient's age on leaving full time education increased so did the likelihood of the patient completing the PGI. Patients with higher scores on the SF-36 scale of physical functioning were more likely to return a questionnaire with a completed PGI, whereas patients with higher scores on the SF-36 scales of role limitations due to physical problems and energy/fatigue were less likely to return a questionnaire with a completed PGI. Divorced or separated patients, patients living in homes rented from the council, and retired patients were less likely to return a questionnaire with the PGI completed.

Of the original 1317 patients agreeing to take part in the main Grampian Health Outcomes Study, 1148 remained in the study after one year. These patients were all sent a one year follow up questionnaire—including the PGI—regardless of whether or not they had com-

Table 3 Validity: correlation between PGI, the condition specific scores, and eight scales of the SF-36 for condition specific groups

Measure	PGI Scores			
	Low back pain (n=359)	Menorrhagia (n=200)	Suspected peptic ulcer (n=120)	Varicose veins (n=158)
Condition specific scores <sup>a</sup>	-0.42**	-0.25**	-0.39**	0.30**
SF-36 health survey scores:				
Physical functioning	0.26**	0.13	0.14	0.24*
Social functioning	0.38**	0.29**	0.24*	0.24*
Role limitations due to physical problems	0.27**	0.24**	0.18	0.15
Role limitations due to emotional problems	0.18**	0.20*	0.12	0.32**
Mental health	0.23**	0.28**	0.24*	0.29**
Energy/fatigue	0.27**	0.23**	0.33**	0.28**
Pain	0.47**	0.29**	0.30**	0.29**
General health perception	0.13*	0.27**	0.39**	0.06

<sup>a</sup> The condition specific instruments are scored in the opposite direction to the SF-36; zero is reserved for patients ticking the least severe response categories to each question and 100 for patients ticking the most severe response to each question.

\*\*Significant at the 0.1% level.

\*Significant at the 1% level.

pleted the PGI at baseline. Of these, 775 returned a questionnaire and 239 completed the PGI at baseline and one year follow up.

#### RELIABILITY

To test the reliability of the instrument, 239 patients who had returned a one year follow up questionnaire with a correctly completed PGI were sent a retest questionnaire. Of the 217 patients returning the questionnaire, 148 patients stated that their health had stayed the same, making them eligible for the test-retest analysis. Correlating the two sets of PGI scores achieved a reliability coefficient of 0.65 ( $p < 0.001$ ). The differences between the first and second set of scores was approximately normally distributed with a mean of  $-0.6$ , a SE of 1.4, and a SD of 16.6. This gives 95% confidence intervals of  $-3.3$  to 2.1 for the mean difference, and 95% limits of agreement of  $-33.1$  to 31.9 for individual comparisons.

#### VALIDITY

PGI scores for the four conditions achieved highly significant negative correlations with the respective condition specific scores (table 3).

Three of these correlations were above 0.3. For all conditions, PGI scores showed significant small to moderate correlations with the majority of the SF-36 scales (table 3).

Table 4 shows the findings of stepwise regression. Between 15% (menorrhagia) and 26% (suspected peptic ulcer) of the variation in PGI scores could be explained by health and sociodemographic variables. Table 5 shows further tests of construct validity. All but three failed to achieve statistical significance, but in the majority of cases (19 out of 25) the mean score differences favour the "construct". For example, for all four patient groups non-referred patients have a higher PGI score than referred patients. For patients with low back pain and varicose veins, PGI scores appear to be related to the general practitioner severity ratings for the none, mild, and moderate categories. Women had lower PGI scores than men, whereas patients owning their own homes had higher scores than those living in rented accommodation. For low back pain and varicose veins, married patients had higher PGI scores than single or separated patients. For all four conditions, unemployed patients had lower PGI scores. Finally, for low back pain patients, analgesic use and strength appear to be associated with lower PGI scores, as does the presence of family history with suspected peptic ulcer, although these differences were not significant.

Table 4 Validity: stepwise regression of PGI scores on condition specific scores, eight SF-36 scales, and sociodemographic variables by condition specific groups

Condition variable	Regression coefficient	Standard error	t	Significance level
Low back pain				
Pain	0.310	0.059	5.21	.0000
Condition specific instrument	-0.236	0.078	-3.00	.0029
Retired	7.75	3.3	2.37	.0184
(Constant)	30.0	4.7	6.43	.0000
$R^2 = 0.25^*$				
Menorrhagia				
Social	0.162	0.062	2.62	.0094
Pain	0.139	0.053	2.61	.0099
Age	0.323	0.16	2.06	.0404
(Constant)	-0.0179	7.0	-0.00	.9980
$R^2 = 0.15$				
Suspected peptic ulcer				
General health perception	0.259	0.077	3.36	.0011
Condition specific instrument	-0.393	0.12	-3.34	.0011
(Constant)	35.7	6.7	5.29	.0000
$R^2 = 0.26$				
Varicose veins				
Role: emotional	0.112	0.038	2.98	.0034
Pain	0.214	0.069	3.12	.0022
Female sex	9.38	3.5	2.65	.0090
(Constant)	16.5	4.9	3.39	.0009
$R^2 = 0.19$				

\*Amount of variation these variables are able to account for in PGI scores (that is, 25%).

#### RESPONSIVENESS

Of the 672 patients completing the PGI at baseline, 239 completed all measures allowing a comparison of standardised response means. In table 6, significant improvements in mean PGI scores are seen for three conditions. No significant improvements were found on any of the measures for patients with varicose veins. For low back pain, the largest standardised response mean was seen for the PGI and the SF-36 pain scale, with five of the eight SF-36 scales recording significantly lower standardised response means relative to the PGI. For patients with menorrhagia and suspected peptic ulcer, only the condition specific instruments registered larger standardised response means than the PGI.



Table 5 Validity: PGI scores by referral, severity, sex, housing, marital status, employment, and condition specific variables

Variable	Low back pain		Menorrhagia		Peptic ulcer		Varicose veins	
	n	PGI Score (SD)	n	PGI Score (SD)	n	PGI Score (SD)	n	PGI Score (SD)
Patient referred:								
No	206	34.7** (18.8)	47	33.6 (20.0)	62	43.8 (21.3)	29	41.5 (17.4)
Yes	153	29.2 (17.6)	153	30.4 (16.0)	58	38.3 (17.5)	129	39.9 (20.4)
GP severity rating:								
None	6	44.7 (28.8)	0	—	4	45.2 (34.6)	1	50.0
Mild	81	37.0 (16.8)	14	33.3 (17.8)	23	43.6 (18.9)	13	46.1 (14.1)
Moderate	103	32.2 (19.1)	30	33.2 (20.6)	29	45.3 (22.7)	21	35.8 (20.9)
Severe	12	36.0 (22.9)	6	33.8 (22.0)	2	43.3 (21.2)	2	50.4 (11.2)
Sex:								
Male	191	32.6 (17.4)	—	—	75	43.8 (21.4)*	36	48.8 (20.2)**
Female	167	31.9 (19.6)	—	—	45	36.7 (15.6)	121	37.8 (19.0)
Housing tenure:								
Self ownership	255	33.4 (17.9)	145	31.9 (16.9)	82	40.9 (20.1)	107	41.3 (20.6)
Rented	104	29.7 (19.7)	55	29.2 (17.4)	34	39.8 (17.6)	48	38.1 (18.1)
Marital status:								
Married	278	32.6 (18.8)	169	31.1 (31.1)	95	40.8 (19.7)	126	40.6 (19.2)
Single	78	31.4 (17.4)	31	31.9 (31.9)	23	40.8 (19.7)	29	37.6 (22.5)
Patient unemployed:								
No	350	32.4 (18.6)	194	31.3 (17.1)	118	40.8 (19.6)	152	40.5 (19.7)
Yes	8	28.0 (16.4)	6	27.6 (17.7)	0	—	4	28.0 (21.7)
Analgesic medication being taken:								
No	64	35.9 (21.2)	—	—	—	—	—	—
Yes	295	31.6 (17.8)	—	—	—	—	—	—
Analgesic strength*:								
Mild-moderate	58	33.1 (18.4)	—	—	—	—	—	—
Moderate-severe	172	30.2 (16.1)	—	—	—	—	—	—
Family history of ulcer disease:								
No	—	—	—	—	62	43.8 (20.1)	—	—
Yes	—	—	—	—	55	38.8 (18.6)	—	—

Asterisks denote significant differences between baseline and follow up scores: \*p<0.05; \*\*p<0.01.

Table 6 Mean changes in PGI, condition specific variables, and SF-36 scores over one year and standardised response means (SRMs) for condition specific groups

Measure	Low back pain (n=124)		Menorrhagia (n=61)		Suspected peptic ulcer (n=27)		Varicose veins (n=27)	
	Mean (SD)	SRM	Mean (SD)	SRM	Mean (SD)	SRM	Mean (SD)	SRM
PGI	11.68** (22.53)	0.52	8.67** (20.64)	0.42	9.89* (21.35)	0.47	-4.02 (16.40)	-0.17
Condition-specific instruments	8.13** (17.43)	0.46	6.01** (13.12)	0.45	11.86** (17.56)	0.58	0.96 (8.61)	0.15
SF-36								
Physical function	2.11 (21.39)	0.09	-0.14 (13.78)	-0.01	1.99 (12.13)	0.08	2.96 (14.63)	0.19
Social functioning	4.55 (29.98)	0.12	3.90* (23.94)	0.17	10.33 (21.13)	0.37	-2.74 (21.66)	-0.09
Role: physical	17.67** (41.08)	0.43	14.75** (46.85)	0.31	15.24 (43.60)	0.26	11.11 (41.79)	0.28
Role: emotional	8.74* (48.53)	0.17	15.85 (44.97)	0.35	7.91 (41.47)	0.22	-1.23 (52.69)	-0.07
Mental health	1.19 (17.65)	0.07	2.16 (14.62)	0.16	5.21 (17.76)	0.27	-1.04 (12.43)	-0.00
Energy/fatigue	4.34* (19.34)	0.23	2.51 (18.5)	0.15	9.02 (21.33)	0.16	0.56 (14.37)	0.10
Pain	12.81** (24.75)	0.52	2.37 (21.86)	0.12	18.53 (28.29)	0.39	3.70 (23.47)	0.16
General health perception	1.59 (17.00)	0.10	-1.40 (13.07)	-0.11	-2.72 (17.35)	0.26	-0.96 (12.17)	-0.09

Underlined values represent significantly lower standardised response means relative to PGI: double underline p<0.05; single underline p<0.01.

## Discussion

We validated the PGI by comparing it with generic and condition specific measures of health. Health is not usually valued for its own sake, but for the extent to which it influences our ability to enjoy life. Hence we expected to find a moderate correlation between the PGI and measures of health. This was the case for the four conditions studied. Using multiple regression we found that the variables entering the regression equation were able to explain a substantial proportion of the variation in PGI scores. Furthermore, for each condition the variables that one would expect to have the greatest impact on quality of life—for example, pain in low back pain and social function in menorrhagia—entered the regression equation. Low back pain impacts on quality of life primarily through the physical impairment and disability that results. Because these factors are captured by the clinical back pain questionnaire, it is perhaps not surprising that this was the only condition specific instrument that entered the regression model for the four conditions.

In further tests of validity we constructed several hypothetical constructs. None of our results conflicted with these hypotheses, but in most cases the differences in PGI scores were small, and few reached statistical significance. These equivocal findings may simply reflect small subsample sizes, alternatively they may indicate the weaknesses of the underlying hypotheses. If perceived quality of life depends on the gap between a person's reality and his or her expectations, then individuals may "stabilise" their perceived quality of life by adapting their expectations to suit their particular social or economic circumstances. For example, the tendency for quality of life to "return" to a stable baseline after a sudden dramatic deterioration or improvement in life circumstances has been observed in individuals with quadriplegia and in lottery winners.<sup>21</sup> This process of adaptation could explain the small observed differences in PGI scores between patients with different sociodemographic characteristics.

Our findings suggest that the PGI shows an acceptable level of reliability when used to make group comparisons, for example in a

clinical trial. In its present form, however, it is not suitable for routine use in the assessment of individual patients. Although, on average, PGI scores remain fairly stable, for an individual a score can change by up to 33 points when no change in quality of life has actually occurred. The PGI was found to be quite responsive to changes in perceived quality of life over time. Using the standardised response mean as a measure of responsiveness, the PGI was able to detect small to moderate change in three of the four conditions over a one year period. In the absence of a gold standard, the extent of any true change is unknown. However, in those conditions showing an improvement after one year, standardised response means are larger for the PGI than for all of the eight SF-36 scales, with the exception of the SF-36 pain scale in patients with low back pain. For the same conditions the PGI has standardised response means of a similar magnitude as the condition specific instruments. Recent research in Oxford provides further evidence for the responsiveness of the PGI to changes in quality of life over time. In a study of 89 patients with obstructive sleep apnoea treated for three months with positive airways pressure therapy, the PGI was able to detect larger clinical changes (measured by the effect size) than the SF-36, the EUROQoL, and even a condition specific measure of sleepiness.<sup>22</sup>

When given as a self completed questionnaire, certain patients encounter difficulties in understanding and completing the PGI. Patients who successfully completed and returned the SF-36 but not the PGI were found to be less well educated, to live in council accommodation, and to be retired. Although non-response is rarely a problem when the PGI is done by interview, the low response for the postal version has implications for the practical use of the PGI in routine health service settings. We have tried to address these problems by developing a more user friendly version (appendix 2). Experience with the use of the concise version in patients with atopic dermatitis suggests that interviewer administration is still necessary at initial assessment,<sup>23</sup> although once patients have been introduced to the PGI, they are then able to self complete it at subsequent follow up. Another study, which used the revised PGI in a postal survey of patients with limiting long term illness without an initial interview administration, found that the response rate was no better than that reported for the original PGI.<sup>24</sup> We believe that these findings do not preclude the use of the PGI in routine health service settings, however. Anecdotal reports from medical and nursing staff who have used the PGI suggest that when incorporated into the initial patient consultation, rather than adding to staff workload, the PGI acts as a powerful care management tool. By basing treatment strategies on the goals of care identified by the patient in stage one of the PGI, and by focusing care activity on those goals given greatest weight in stage three, the PGI may improve quality of patient care independently of its effects as an outcome measure. This potential for a patient centred quality of life assessment

tool such as the PGI to improve the process of care needs to be evaluated in further studies.

The PGI is still under development. It is unclear whether the current version goes far enough in capturing the impact of disease on quality of life in a meaningful way, or whether it is possible to push it further without compromising reliability and responsiveness to change. It is also unclear whether the way in which the questions are framed, and the scores derived, can be modified to improve these measurement properties.

## Conclusion

We have shown that it is possible to develop a PGI of quality of life that not only assesses the extent to which patients' expectations are matched by reality but also at the same time satisfies criteria of reliability and responsiveness to change. The PGI is still at an early stage in its development, and further work is required to make it more acceptable and meaningful to patients, but we believe it offers an exciting new approach to the evaluation of medical care.

Those interested in the PGI may visit the website at <http://www.dundee.ac.uk/epidemiology/PGI>

We thank the staff at Inverurie, Portlethen, Rubislaw Place, and Westhill practices for recruiting patients; Jeremy Grimshaw, Jenny Duncan, and Alison De Ville for help with data collection; John Ware and his colleagues at the Health Institute of the New England Medical Center for permission to use the SF-36; Mona Abdalla for statistical advice; and Elizabeth Russell for comments.

This research and the Health Services Research Unit are funded by the Chief Scientist Office of the Scottish Office Health Department; however, the opinions are those of the authors and not necessarily those of the Scottish Office Health Department

- 1 Ruta DA, Garratt AM. From health status to quality of life measurement. In Jenkinson C, editor. *Measuring health and medical outcomes*. London: UCL Press, 1995.
- 2 Calman KG. Quality of life in cancer patients—a hypothesis. *J Med Ethics* 1984;10:124–7.
- 3 Ruta DA, Garratt AM, Leng M, et al. A new approach to the measurement of quality of life: the patient generated index (PGI). *Med Care* 1994;32:1109–26.
- 4 Guyatt GH, Berman LB, Townsend M, et al. A measure of quality of life for clinical trials in chronic lung disease. *Thorax* 1987;42:773–8.
- 5 Hoinville G. *The priority evaluator method. Methodological working paper 3*. Department of Social and Community Planning and Research. London: London University, 1977.
- 6 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.
- 7 Garratt AM, Ruta DA, Abdalla MI, et al. The SF 36 health survey questionnaire: an outcome measure suitable for routine use within the NHS? *BMJ* 1993;306:1440–4.
- 8 Garratt AM, Macdonald LM, Ruta DA, et al. Towards the measurement of outcome for patients with varicose veins. *Quality in Health Care* 1993;2:5–10. 9 Ruta DA, Garratt AM, Wardlaw D, et al. Developing a valid and reliable measure of health outcome for patients with low back pain. *Spine* 1994;17:1887–96.
- 10 Ruta DA, Garratt AM, Chadha YC, et al. Assessment of patients with menorrhagia: how valid is a structured clinical history as a measure of health status? *Quality of Life Research* 1995;4:33–40.
- 11 Garratt AM, Ruta DA, Russell I, et al. Developing a condition-specific measure of health for patients with dyspepsia and ulcer-related symptoms. *J Clin Epidemiol* 1996;49:565–71.
- 12 Garratt AM, Ruta DA, Abdalla MI, et al. Responsiveness of the SF-36 and a conditionspecific measure of health for patients with varicose veins. *Quality of Life Research* 1996;5:223–34.
- 13 Armitage P, Berry G. *Statistical methods in medical research*. Second edition. Oxford: Blackwell Scientific Publications, 1987.
- 14 Streiner DL, Norman GR. *Health measurement scales: a practical guide to their development and use*. Oxford: Oxford University Press, 1990.
- 15 Guyatt GH, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chron Dis* 1987;40:171–8.

16 Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care* 1989; 27(suppl):178-89.

17 Katz JN, Larson MG, Phillips GB, et al. Comparative measurement sensitivity of short and longer form health status instruments. *Med Care* 1992;30:917-25.

18 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;i:307-10.

19 Garratt AM, Ruta DA, Abdalla MI, et al. SF-36 health survey questionnaire: ii. responsiveness to changes in health status in four common clinical conditions. *Quality in Health Care* 1994;3:186-92.

20 Liang MH, Larson MG, Cullen KE, et al. Comparative measurement efficiency and sensitivity of five health status instruments for arthritic research. *Arthritis Rheum* 1985; 28:542.

21 Brickman P, Coates DD, Janoff-Bulman. Lottery winners and accident victims: is happiness relative? *Journal of Personality and Social Psychology* 1978;36:917-27.

22 Jenkinson C, Stradling J, Petersen S. How should we evaluate health status? A comparison of three methods in patients presenting with obstructive sleep apnoea. *Quality of Life Research* 1998;7:95-100.

23 Herd RM, Tidman MJ, Ruta DA, et al. Measurement of quality of life in atopic dermatitis: correlation and validation of two different methods. *Br J Dermatol* 1997;136:502-7.

24 Macduff C, Russell E. The problem of measuring change in a disabled population. *Quality of Life Research* 1998; in press. Appendix

**Appendix 1**  
**The patient generated index**

This section of the questionnaire is presented in three stages. Your answers should be entered on the answer sheet, which is marked "stage 1", "stage 2", and "stage 3". The instructions on how to fill it in also are divided into three stages. You might find it easiest to read through the instructions for stage one, then go to the answer sheet and fill in stage one, and then do the same for stages two and three.

When you are filling in each stage, we would like you to think about when you were at your

worst with your back pain in the last month. If you are at your worst now, then think about how you feel now.

**Stage 1**

At this stage, we would like you to think of the different areas in your life, or activities in your life, that have been affected by your back pain in the past month. When you think of areas or activities, you might think of some small area or activity that may be quite personal and special to you such as "I can't play with my kids".

We want you to write the five most important areas or activities of your life that are affected by your back pain in the boxes provided in stage one on the answer sheet. Put one area or activity in each box.

You may be able to think of more than five areas, but you can only write down the five most important ones. If you can't think of five areas or activities that are affected, there is a list of areas and activities that have been mentioned by other people suffering with back pain. You might want to use these areas or activities to fill in the boxes if you feel that they apply to you.

You don't have to write down five areas of your life if you don't feel that five areas of your life have been affected. If you have less than five, you can write "none" in the empty boxes and move on to stage two.

If you feel that your life is not affected by your back pain at all, then just send back the questionnaire with "none" in each box. In this case you don't need to go on to stages two and three.

Once you have written down the most important areas or activities that have been affected by your back pain, you can move on to stage two.

**Stage 2**

Looking at the next page, you will see a scale from 0 to 100 in multiples of 10. This scale is supposed to show you how badly affected you are for each of the areas or activities you have mentioned. A score of 0 would mean that when you were at your worst in the last month, you felt that this was really the worst you could imagine for yourself. The score of 100 is meant to represent exactly how you would like to be, in that area or activity of your life (even if it is impossible for you to reach).

For each area or activity that you have mentioned, write down a score out of 100 that you would give to reflect how you were affected when you were at your worst in the last month.

You will notice that we have filled in "all other aspects of your life" as the final "area or activity." This is meant to include all the other areas of your life affected by your back pain, but that are not important enough to go in the top five boxes. It will also include areas in your life that might be totally unaffected, such as the size of your house.

You might suffer from another illness as well as your back pain, and any other areas that are affected by this illness would be included in this box.

Please give a score out of 100 to the "all other areas of your life" box in the same way that you scored the other "areas and activities". Even if you leave other boxes empty, you must fill in this box.

**Stage 3**

For the final stage, we would like you to imagine that we can grant you a wish to improve any area of your life, including the areas that have nothing to do with your back pain.

STAGE 1 area/activity (eg sport)	STAGE 2 score each area/ activity out of 100	STAGE 3 spend your 60 points between the different areas	
<input type="text"/>	<input type="text"/>	<input type="text"/>	Total number of points should add up to 60
<input type="text"/>	<input type="text"/>	<input type="text"/>	
<input type="text"/>	<input type="text"/>	<input type="text"/>	
<input type="text"/>	<input type="text"/>	<input type="text"/>	
<input type="text"/>	<input type="text"/>	<input type="text"/>	
All other aspects of your life not mentioned above	You must fill <input type="text"/> in this box	<input type="text"/>	

- 100 Exactly as you would like to be
- 90 Close to how you would like to be
- 80 Very good but not how you would like to be
- 70 Good but not how you would like to be
- 50 Fair
- 40 Between poor and fair
- 30 Poor but not the worst you could imagine
- 20 Very poor but not the worst you could imagine
- 10 Close to the worst you could imagine
- 0 The worst you could imagine

**Appendix 2**

**A PATIENT GENERATED INDEX<sup>®</sup>  
OF QUALITY OF LIFE**

Your answers to the following steps will tell us how your life is affected by your BACK PAIN. It will also tell us how you would like to see your life improved

STEP 1: Identifying Areas	STEP 2: Scoring each area	STEP 3: Spending points
<p>We would like you to think of the most important areas of your life that are affected by your BACK PAIN. Please write up to FIVE areas in the the boxes below.</p>	<p>In this part we would like you to score the areas you mentioned in step 1. This score should show how badly affected you were over the past MONTH. Please score each area out of 10 using this scale:</p>	<p>We want you to imagine that any or all the areas of your life could be improved. You have 14 imaginary points to spend to show which areas you would most like to see improve. Spend more points on areas you would most like to see improve and less on areas that are not so important. You don't have to spend points in every area. You can't spend more than 14 points in total.</p>
<p>Here are some areas other patients have mentioned that might help you to think how your life has been affected:</p>	<p>10 = Exactly as you would like to be                      9 = Close to how you would like to be                      8 = Very good but not how you would like to be                      7 = Good, but not how you would like to be                      6 = Between good and fair                      5 = Fair                      4 = Between poor and fair                      3 = Poor but not the worst you could imagine                      2 = Very poor but not the worst you could imagine                      1 = Close to the worst you could imagine                      0 = The worst you could imagine</p>	<p><input type="text"/>  <input type="text"/>  <input type="text"/>  <input type="text"/>  <input type="text"/>  <input type="text"/>  <input type="text"/>  <input type="text"/>  <input type="text"/>  <input type="text"/></p>
<p>Social life, work, housework, hobbies, interests, makes me moody, loss of independence, can't go shopping, walking, climbing stairs, sitting, sleep, tiredness</p>	<p><input type="text"/>  <input type="text"/>  <input type="text"/>  <input type="text"/>  <input type="text"/>  <input type="text"/>  <input type="text"/>  <input type="text"/>  <input type="text"/>  <input type="text"/></p>	<p>Remember total must add up to 14</p>
<p>AREAS AFFECTED BY OTHER HEALTH PROBLEMS</p>	<p>Please use the last two boxes to score all areas affected by other health problems and all other non-health areas of life</p>	<p><input type="text"/>  <input type="text"/></p>
<p>ALL OTHER NON-HEALTH AREAS OF LIFE</p>	<p><input type="text"/>  <input type="text"/></p>	<p><input type="text"/>  <input type="text"/></p>

Imagine that you are given 60 points to improve your score in any of the areas you have mentioned. You cannot have more than 60 points in total, but you can spend them anyway you like. For example, you could give 10 points to each area, or you might give 60 points to one area. The choice is your to split the points up any way you like, but you cannot have more than 60 points in total.

If you don't give any points to an area of your life, you must try to imagine that this area will stay exactly as it is.

Go through the boxes in stage three and distribute your points to those areas or activities in which you would most like to improve. You can keep changing your mind until you feel that you have reached the best distribution of points. Remember that the total across all areas must add up to 60.

You have finished this section. It will tell us how your back pain has affected your life and also which aspects of your life you would most like to see improved.