Patient Centered Assessment of Ankylosing Spondylitis-Specific Health Related Quality of Life: Evaluation of the Patient Generated Index

KIRSTIE L. HAYWOOD, ANDREW M. GARRATT, KRYSIA DZIEDZIC, and PETER T. DAWES

ABSTRACT. ***Objective.*** To evaluate the Patient Generated Index (PGI) for the measurement of individualized health related quality of life in ankylosing spondylitis (AS).

Methods. The PGI asks patients to nominate areas of their lives affected by their disease that they consider the most important. Semistructured interviews with AS patients produced a trigger list of areas of life affected by AS. The PGI was self-completed by UK patients taking part in a multicenter postal survey. The instrument was assessed for data quality, reliability, validity, and responsiveness.

Results. The PGI had acceptable completion rates. Scores covered the available range and approximated normality. Test-retest reliability estimates support the use of the PGI in group evaluation (intraclass correlation coefficient > 0.80). Comparisons with scores for other health status instruments provided evidence for the validity of the PGI. The largest levels of correlation were found for the AS Quality of Life Questionnaire (ASQoL) and the EuroQol. The informed and open format of the PGI had the strongest linear relationship with responses to both specific and general health transition questions ($p < 0.01$), and was the most responsive format of the PGI.

Conclusion. The PGI is the first AS-specific individualized measure of health related quality of life. There is good support for the content validity of the instrument and patient acceptability is high. Adequate levels of data quality and reliability support the use of the PGI in group evaluation. Moderate levels of responsiveness to changes in health were produced by the informed and open format of the PGI. (J Rheumatol 2003;30:764–73)

Key Indexing Terms:
ANKYLOSING SPONDYLITIS PATIENT GENERATED INDEX
MEASUREMENT PROPERTIES

Ankylosing spondylitis (AS) has a substantial effect on an individual's health related quality of life (HRQL)[1,2], encompassing broad issues relating to role and physical functioning, psychological well being, and social interactions. AS is a chronic and often progressive inflammatory disorder, and although primarily affecting the axial skeleton[3], peripheral joints, entheses, and extraarticular sites may also be involved[4]. Although the HRQL of an individual with a disease such as AS is considered by many an important indicator of disease impact[5,6], the Assessment in Ankylosing Spondylitis Group (ASAS) indicated that quality of life could not currently be included as a core domain in AS evaluation due to uncertainty over the best measurement approach[7]. There is an obvious need for a reliable, valid, and responsive AS-specific and patient assessed measure of HRQL.

HRQL is specific to an individual, their priorities, expectations, and experience of life and ill health. To ensure the representation of general patient concerns within patient assessed health instruments, many developers involve patients in the generation of items. However, developers commonly adopt a summated rating scale format whereby patients respond to a predetermined set of items. The majority of AS-specific patient assessed health instruments follow this format[2]. The AS-Quality of Life Questionnaire (ASQoL)[8,9] is a new measure of disease-specific HRQL, comprising 18 items. There is no published evaluation of the ASQoL. Items were generated following patient interviews, and items are summated to provide the final score. Such highly standardized instruments often have good psychometric properties, but may omit issues of importance to individual patients[10,11] while containing items of little relevance to others, thus introducing noise into the evaluation[12] and compromising instrument accuracy and validity[13]. Items

From the Interdisciplinary Research Centre in Health, Physiotherapy and Dietetics Subject Group, School of Health and Social Sciences, Coventry University, Coventry; the Unit of Health-Care Epidemiology, Department of Public Health, University of Oxford, Oxford; Department of Physiotherapy Studies and Primary Care Sciences Research Centre, Keele University, Keele; and Staffordshire Rheumatology Centre, Stoke-on-Trent, England.

Supported by a grant from the Arthritis Research Campaign and the Staffordshire Rheumatology Centre.

K.L. Haywood, PhD, DPhil, Research Fellow, Senior Lecturer in Physiotherapy (Research Methods), University of York; A.M. Garratt, PhD, Research Officer, University of Aberdeen, Scotland; K. Dziedzic, PhD, Arthritis Research Campaign Senior Lecturer in Physiotherapy, Keele University; P.T. Dawes, MB, ChB, FRCP(UK), Consultant Rheumatologist.

Address reprint requests to Dr. A.M. Garratt, Unit of Health-Care Epidemiology, Institute of Health Sciences, Old Road, Oxford OX3 7LF, UK. E-mail: andrew.garratt@uhce.ox.ac.uk

Submitted January 28, 2002; revision accepted October 28, 2002.

may appear detached from the contextual setting, thus losing the social or personal significance that may be afforded by a more individually tailored evaluation[14,15]. The ultimate measure of HRQL will capture a patient's individuality and uniqueness[13], while recognizing and facilitating the dynamic nature of health and disease, both within and between patients.

HRQL has also been evaluated using individualized, patient centered measures of health, which aim to be more sensitive to individual needs, demands, and change in status. Several instruments attempt to provide a more patient centered approach to evaluation, for example, the Disease Repercussion Profile[16] and the Schedule for the Evaluation of Individual Quality of Life[15], but the individualized approach has yet to be applied in the evaluation of AS-specific HRQL.

The Patient Generated Index (PGI)[17,18] is an individualized measure of HRQL, for which acceptable completion rates and measurement properties following both self- (postal) and interview-based completion have been reported[19,20]. The individualized format of the instrument produces good content validity and it has been found to be responsive to change. The PGI has been validated for use in patients with varied medical conditions, including low back pain and menorrhagia[17,19], dermatitis[21], and chronic sleep apnea[20]. To facilitate completion, a disease-specific trigger list of issues considered important by patients is included.

We evaluated the PGI in a large and representative population of patients with AS in the United Kingdom. The reliability, validity, and responsiveness of the PGI are evaluated and the instrument is compared to patient assessed health instruments that are based on summated rating scales.

MATERIALS AND METHODS

Completion of the PGI. The 3 stages of instrument completion are shown in Table 1. The first stage asks the patient to list up to 5 of the most important areas of life affected by AS in blank boxes. The trigger list of important areas commonly mentioned by AS patients and a completed example of the instrument aid completion. Patients may choose not to use the trigger list and identify areas that relate to their own individual experience of disease. Recognizing the role that comorbidity may play in the evaluation of HRQL, patients are given the opportunity to consider the effect of other health problems in a sixth box. The seventh box asks patients to consider non-health related areas of life. In stage 2, patients rate how badly affected they are in each of the areas on a scale of 0 to 10, where 0 represents the worst they can imagine and 10 exactly as they would like to be. Finally, patients "spend points" in stage 3 to reflect their priorities for improvement. Fourteen points can be given to any combination of chosen areas. Multiplying each of the 7 possible ratings (stage 2) by the proportion of points given to that area (stage 3) and summing produces a score from 0 to 10. The score is designed to represent the extent to which reality falls short of expectations in those areas of life in which patients would most value an improvement[17].

Evaluation of the PGI in AS. All patients had a confirmed diagnosis of AS (Modified New York Criteria[22]), were registered with one of a group of specialist centers of rheumatology in England and Scotland, and were aged between 18 and 75 years. Pregnancy, learning difficulties, or an inability to comprehend written English were exclusion criteria. The study was approved by the Northern and Yorkshire Multi-centre Research Ethics Committee and relevant local research ethics committees. Informed and written consent was given by all participants. Different samples of patients were used at each stage of the evaluation.

There were 3 stages in the evaluation of the PGI in AS patients: the development of an AS-specific trigger list; instrument pretesting; and evaluation of measurement properties in a postal survey.

1. Developing an AS-specific trigger list. Following the developers, a trigger list of the most important areas of life affected by AS was developed for the PGI[17,19]. A random sample of 40 patients [n = 33 men; mean age 46.2 yrs (SD 10.1)] registered with the Staffordshire Rheumatology Centre (SRC) were invited to attend the SRC to discuss the effect of AS on their day-to-day life with the lead investigator (KLH). Two further letters were sent to nonresponders. Patients not wishing to participate in the interviews were asked to return the precoded consent form. Twenty-nine (72.5%) patients representing a broad spectrum of disease duration [n = 24 men; mean age 48.4 yrs (SD 10.1), range 31–69 yrs; mean duration of AS diagnosis 11.0 yrs (SD 10.7), range 2–41 yrs) agreed to an interview.

Semistructured interviews were conducted to elicit a patient's free responses about the influence of AS on their everyday life and the importance of areas affected by AS. Interviews took place in a private room and lasted between 30 and 60 min. With the consent of participants, interviews

Table 1. Stages of completion of the PGI.

Stage 1 Area or Activity	Stage 2 Score Out of 10	×	Stage3 Spend Your Points	Total
1. Impact on ability to work	6		0/14	0
2. Worry about the future	4		3/14	0.86
3. Relationship with my partner	3		6/14	1.30
4. Unable to plan ahead	2		2/14	0.28
5. Feelings of low self-esteem	3		2/14	0.43
6. Areas affected by health problems other than AS	10		0/14	0
7. All other non-health areas of your life	7		1/14	0.50
Total score			14	4.23

Stage 1: Area or activity: Identify up to 5 of the "most important areas of your life affected by AS". Stage 2: Score out of 10: scale of 0 to 10, where 0 represents the worst they can imagine for themselves and 10 that they are exactly as they would like to be. Stage 3: Spend your points: points "spent" to reflect patient priorities for improvement. Fourteen points can be given to any combination of areas listed in stage 1. Total score: calculated by multiplying each of the 7 possible ratings (stage 2) by the proportion of points given to that area (stage 3) and summing produces an index score from 0–10, where higher scores indicate better health related quality of life.

were audio-recorded and later transcribed[23]. Following a content analysis of the transcripts, verbatim statements reflecting important and common themes related to the HRQL of patients with AS were listed[14]. Related themes were highlighted, grouped together, and organized by conceptual categories[24]. The content analysis and category identification was discussed between members of the development team and assessed for repetition and ambiguity. Area selection for the trigger list was determined by conceptual categories, frequency of endorsement, and available space in the formatted instrument.

2. Pretesting the PGI. The instrument was pretested for acceptability and feasibility as a self-completed postal questionnaire in a random sample of patients from the SRC AS database (n = 10) (9 men; mean age 47.7 yrs, SD 12.9). Questionnaire completion was followed by semistructured interviews at the SRC to identify any difficulties with instrument completion. With the consent of participants, verbatim statements in relation to the acceptability and feasibility of the PGI were noted (by hand).

3. Postal evaluation of PGI. The reliability, validity, and responsiveness of the PGI were assessed following self-completion in a multicenter postal survey with participants from the North, Midlands, and South of the UK. Following evaluations of patient assessed health instruments, a population of over 400 patients was deemed acceptable for the postal survey[25,26]. Patients not wishing to participate were asked to return the questionnaires using a reply-paid envelope. Nonresponders were sent reminders at 2 and 4 weeks. Patients taking part in the study were sent a second retest questionnaire 2 weeks after completing the first, and a third questionnaire at 6 months. The postal questionnaire, which included the PGI, disease-specific and generic health instruments, 2 health transition items, and questions relating to age, disease duration, marital status, post-school education, occupational status and housing tenure, was sent to 451 patients.

The disease-specific instruments reflect domains considered important in the evaluation of patients with AS[7]: HRQL, disease activity, functional disability, and pain. The 18 item ASQoL uses "yes/no" scaling. Items are summed to produce a score from 0 to 18, where a lower score indicates a better level of AS-specific HRQL[8]. The Bath AS Disease Activity Index (BASDAI) is a 6 item measure of disease activity that uses visual analog scales[27]. Scores are summed and transformed to a 0–10 scale; lower scores indicate less disease activity. The Revised Leeds Disability Questionnaire (RLDQ) is a 16 item measure of AS-specific functional disability[28]. Items use 4 point adjectival scales and sum to produce a score from 0 to 48; higher scores indicate greater functional disability. The Body Chart allows patients to indicate the areas and severity of global pain[29]. Areas of current pain are sketched onto a body manikin with anterior and posterior views. Each area is scored on a 4 point scale. Area scores are summed; lower scores indicate less bodily pain. The instrument was developed and tested following interview administration in a clinic environment, and there is evidence for satisfactory measurement properties[4,29].

The EuroQol[30] and the Short Form 12 item Health Survey (SF-12)[31] were identified as 2 short and comprehensive generic approaches to assessing overall health. There has been no published evaluation of the EuroQol and SF-12 in AS, but both instruments have good evidence for their measurement properties when applied in the evaluation of patients with disease similar to AS[32,33]. The EuroQol has 2 sections: the first (EQ-5D) has 5 items covering the domains of mobility, self-care, usual activity, pain/discomfort, and anxiety/depression. Items use 3 point adjectival response scales. Scores are transformed using utility weights derived from the general population to produce a single index. Scores range from –0.59 to 1.00; 1.00 is perfect health and a score below 0 is considered worse than death. The second section (EQ-VAS) includes a visual analog scale on which the patient rates their overall health today from 0 (worst imaginable) to 100 (best imaginable). The SF-12 comprises 12 items derived from the Medical Outcome Study Short Form 36 item Health Survey (SF-36)[23] and uses adjectival scales. It produces mental and physical health summary scales based on scores for the general population that range from 0 to 100; higher scores indicate a better HRQL.

Patient reported health transition questions have been widely used as external criteria in the evaluation of instrument test-retest reliability and instrument responsiveness, providing a valid reflection of the extent and direction of change in specific or general health over time[34,35]. Two health transition questions addressing self-reported change in AS-specific and general health at 2 weeks and 6 months were included in the respective followup questionnaires (Compared to 2 weeks ago/6 months ago how would you rate your AS/general health now — much better, somewhat better, about the same, somewhat worse, much worse?).

Evaluation of the different formats is required to facilitate interpretation of the PGI[20] and to identify the most appropriate format for use in clinical trials and routine practice. Three formats of the PGI that varied according to the provision or absence of areas identified in stage 1 at baseline were used for followup completion. First, the "blind" format does not give patients their baseline areas. Second, the "closed" format gives patients their baseline areas, but they cannot change them. Third, the "informed and open" format gives patients their baseline areas and they can maintain or change them in any way they want. Participants were randomly assigned to complete the blind or closed format at 2 weeks, and the blind or informed and open format at 6 months.

Reliability. Patients were mailed a second questionnaire at 2 weeks and randomly allocated to receive the blind or closed format of the PGI. Two weeks is considered an appropriate period for test-retest reliability in patients with a stable condition[36]. Reliability was assessed for patients whose general and AS-specific health had remained the same according to their responses to the health transition questions. The intraclass correlation coefficient (ICC) (2, 1)[37] was used to assess agreement between test and retest[26]. For group comparisons, levels of reliability over 0.70 are required[36,38], and for evaluation of individuals levels above 0.90 have been recommended[36,38].

Following the instrument developers, the effect of area change (stage 1) on PGI reliability was also assessed[17]. Area substitution is given 1 point. Area addition and removal are given half a point. Points are summed to reflect area change (a scale of 0–5, where 0 indicates no area change and 5 change in all areas). This calculation is only possible for the blind format of the PGI.

Validity. Construct validity was assessed by correlating (Pearson r) the PGI index scores with those for disease-specific and generic instruments. Hypothesized theoretical relationships between instrument scores were considered *a priori*[38]. The PGI (individualized disease-specific), ASQoL (disease-specific), BASDAI (disease activity), Body Chart (bodily pain), RLDQ (functional disability), EuroQoL (general health), and the SF-12 (general health) measure related aspects of HRQL. The majority of the content of these instruments overlaps with the PGI trigger list. However, rather than simply producing a summed score for the different areas, the PGI also includes the ability to fulfil one's expectations. It is hypothesized that this, together with the non-health component of the PGI, will limit the correlation with scores for the disease-specific instruments to a moderate level. Five of the 7 areas in stage 1 of the PGI are disease-specific, the sixth relates to general health, and the seventh relates to non-health. Therefore, a small to moderate level of correlation is hypothesized with scores for the generic instruments. To test this hypothesis further, a score for the PGI was calculated following the omission of area 7, and following the omission of both areas 6 and 7, and the same relationships between instruments was assessed.

Responsiveness. Patients who responded to the baseline questionnaire were mailed a questionnaire at 6 months. Patients were randomly allocated to complete the blind or informed and open format of the PGI. The PGI was assessed for responsiveness to change by calculating the modified standardized response mean (MSRM), which is equal to the mean change in scores (6 month minus baseline) divided by the standard deviation of change scores in patients defined as stable[38]. Guidance for interpretation suggests that a score greater than 0.8 represents a large level of responsiveness, a score of 0.5 moderate responsiveness, and a score of 0.2 a small

level[38]. MSRM were calculated for patients reporting an improvement or deterioration on the health transition questions (general or AS-specific).

Changes in instrument scores and patient response to the transition questions at 6 months were also assessed for a linear trend[26,38]. To the extent that the patient assessed instruments are valid measures of health capable of measuring change, a strong association with a patient reported health transition item is expected[26,38].

RESULTS

Developing an AS-specific trigger list. Twenty-nine (72.5%) patients agreed to participate in the interviews [n = 24 men (82.7%); mean age 48.4 yrs (SD 10.1); mean duration AS diagnosis 11.0 yrs (SD 10.7)]. Nonparticipants were significantly younger than responders (mean age 40.3 yrs, SD 7.68, range 28–52) (t test, p = 0.02), but there was no significant difference by sex (Fisher's exact test, p = 0.64). Following the content analysis, 16 conceptual categories were identified. The trigger list contains 37 areas reflecting these categories.

Pretesting the PGI. Six patients were interviewed following postal self-completion (n = 5 men, mean age 46.0 yrs, SD 15.7). The interviews did not reveal any problems with completion of the PGI.

Postal evaluation of PGI. Of 451 patients who were sent a postal questionnaire, 349 (77.4%) returned them completed; 303 (69.4%) and 289 (64.1%) patients returned the 2 week questionnaire and 6 month questionnaire, respectively.

The majority of patients were male (n = 259; 74.2%) with a mean age of 46.1 years (SD 12.6, range 18–75). The mean symptom duration of participants was 19.8 years (SD 11.8, range 1–56), suggesting a broad spectrum of disease presentation.

Instrument evaluation. Correct completion of the PGI requires all 3 stages to be completed, including the correct allocation of points in stage 3[19]. PGI scores were computable for 303 (87.5%) patients returning baseline questionnaires. Score distributions approximated normality, with a mean baseline score of 4.05 (SD 1.65) on a scale 0–10, where 10 is the best possible HRQL. Baseline data for all instruments completed in the postal survey are summarized in Table 2.

All patients completing the PGI entered at least one area in stage 1 (areas 1–5) (n = 303). Five areas were entered by 68.9% (n = 209) of patients, 87.7% (n = 266) entered 4 areas, 95.0% (n = 288) entered 3, and 96.7% (n = 293) entered 2 areas. In total, 68 areas were entered including all 37 trigger list areas (Table 3). The most frequently listed areas affected by AS were "work" and "sleep." The PGI trigger list addresses all but one of the items within the ASQoL ("I often get frustrated"), which is evidence for the content validity of the list. The item not included in the trigger list was entered by 4 (1.3%) patients; 213 (70.5%) patients reported health problems additional to AS (area 6), and 294 patients (97.3%) provided a score (stage 2) for area 7 (non-health related areas of life).

Patients with higher index scores tended to include a lower number of areas in stage 1 (F 4.38; p = 0.002).

Reliability. Of the 303 patients who returned the questionnaire at 2 weeks, 269 completed both transition questions (88.8%). For both formats of the PGI, test-retest reliability was assessed for patients reporting no change in both AS and general health: 173 patients (57.0%) reported no change in health at 2 weeks (AS and general health had improved, n = 40; AS and general health had deteriorated, n = 56). Of those reporting no change in health, 144 patients (83.2%) correctly completed the PGI to allow calculation of an index score at baseline and at 2 weeks (blind format, n = 75; closed format, n = 69). Reliability was also assessed against areas changes in stage 1 for patients completing the blind format.

Table 4 shows that PGI scores had higher levels of reliability for the closed format (ICC = 0.88) than for the blind format (ICC = 0.82), although both formats are suitable for use in groups of patients. High levels of test reliability were found for patients not changing any areas or scoring up to 1 point on the number of area changes following blind completion. Increasing area changes are associated with reduced reliability, and reliability falls below acceptable levels when 3 to 5 area changes are made.

Validity. All the correlations between the PGI and other instrument scores were in the hypothesized direction, and moderate correlations between the PGI and the disease-specific and generic instruments were found (Table 5). As hypothesized, a hierarchy of association was found, the largest correlations being between the PGI and disease-specific instruments, followed by the generic instruments. The largest correlation was with the ASQoL and the smallest correlation with a disease-specific instrument was with the RLDQ. Moderate levels of correlation were found with the EQ-5D and SF-12 Mental Component Scale (MCS). Slightly larger correlations were found with the EQ-VAS and SF-12 Physical Component Scale (PCS). After removal of area 7 and the omission of both areas 6 and 7 from the index score, levels of correlation were in broad agreement with those for the total score (Table 5).

Responsiveness. The results of responsiveness testing for the blind, and informed and open formats of the PGI are shown in Table 6a and 6b, respectively. For clarification, whereas the blind format does not give patients their baseline areas, the informed and open format gives patients their baseline areas, which may be maintained or changed as desired.

The change scores for the informed and open format of the PGI reflect the categories of both AS-specific and general health transition (Table 6b). The largest changes were found for the informed and open PGI, BASDAI, EQ-5D, EQ-VAS, and SF-12 MCS on both AS-specific and general health transition. For example, patients who say that their AS is better have an average improvement in PGI score of 0.75 over the 6 months (on a scale of 0–10), and where

Table 2. Scale properties of study instruments at baseline (n = 349).

Instrument	Correct Completion, n	Missing, %	Mean SD	Response Options*	
				Floor, %	Ceiling, %
PGI	303	13.2	4.06 (1.6)	0.7	0.3
ASQoL	339	2.9	8.77 (5.5)	7.1	5.3
BASDAI	318	8.9	4.43 (2.2)	3.8	0.9
Body Chart	310	11.2	16.16 (16.2)	2.3	0.3
RLDQ	342	2.0	14.24 (10.2)	4.7	0.0
EQ-5D	340	2.6	0.53 (0.34)	0.3	7.1
EQ-VAS	343	1.7	57.93 (21.29)	0.3	0.9
SF-12 Mental component scale	331	5.2	45.49 (11.61)	0.0	0.0
SF-12 Physical component scale	331	5.2	36.63 (11.03)	0.0	0.0

PGI: scored 0–10, where higher scores indicate better health related quality of life. ASQoL: scored 0–18, where lower scores indicate better health related quality of life. BASDAI: scored 0–10, where higher scores indicate greater disease activity. Body Chart: scored from 0 upwards, with no maximum score limit. Higher scores indicate greater levels of perceived body pain. RLDQ: scored 0–48, where higher scores indicate increased limitation in functional ability. EuroQol EQ-5D: scored –0.59 to 1.0, where –0.59 is the worst and 1.0 the best possible health. EuroQol-VAS: scored 0–100, where higher scores indicate better health states. SF-12 uses norm-based scoring from the general population. Scales are transformed: mean of 50 (SD 10), range 0–100. * Floor of the scale is the lowest possible score; ceiling is the highest possible score.

Table 3. Frequency of areas entered in stage 1 of PGI*.

Trigger List	Frequency (%)	Additional Areas Cited	Frequency (%)
Impact on / unable to work	121 (39.9)	Morning stiffness / stiffness	9 (2.9)
Disturbed sleep	92 (30.4)	Gardening	9 (2.9)
Worry about the future	80 (26.6)	Ability to complete tasks / do simple tasks	7 (2.4)
Sporting activities / exercise	70 (23.2)	Slow to do things	7 (2.4)
Feeling tired	64 (21.2)	Lack of spontaneous thought / mental concentration	6 (1.9)
Difficulty with housework / DIY / lifting	47 (15.6)	Named body part (other than back/knee/hands)	6 (2.0)
Walking	50 (16.5)	Travelling / travel distances	6 (2.0)
Ability to remain physically active / general mobility	51 (16.8)	Dressing and bathing / personal hygiene	5 (1.6)
Poor self body image / posture / embarrassment / self-conscious	44 (14.6)	Shopping	5 (1.7)
Fatigue / loss of energy / lethargy / stamina	43 (14.2)	Normal activities	5 (1.6)
Pain / discomfort	41 (13.5)	Back	4 (1.3)
Feelings of depression	40 (13.2)	Impact of medication / side-effects / efficacy	4 (1.3)
Driving / getting into and out of car	40 (13.6)	General fitness (physical)	4 (1.3)
Social life / holidays / relationships with friends	39 (12.9)	Ability to relax / relaxation	4 (1.2)
Relationship with wife / husband / partner	39 (12.9)	Frustration / anxiety	4 (1.2)
Specific limitation to joint / spinal mobility	38 (12.6)	Vision / iritis	2 (0.6)
Feeling of low self-esteem / confidence	37 (12.2)	Financial impact	3 (1.0)
Ability to play with / look after children / grandchildren	37 (12.2)	Concern over weight gain	2 (0.6)
Pursuing hobbies / past-times / leisure activities	29 (9.6)	Fear of being knocked / standing in crowds	2 (0.6)
Getting going in the morning	28 (9.2)	Quality of my life	2 (0.6)
Sex life	27 (8.8)	Health	2 (0.6)
Family life / relationship with family / children	23 (7.6)	Hands	2 (0.6)
Ability to plan ahead	23 (7.7)	Knees	2 (0.6)
Difficulty standing / standing for long periods	23 (7.6)	Difficulty with transfers — crouch to standing / out of chairs / out of bath	2 (0.6)
Level of independence / dependency on others	21 (6.3)	Breathing	1 (0.3)
Enjoyment of life	20 (6.6)	Crossing the road	1 (0.3)
Loss of motivation	18 (6.0)	Reaching above head	—
Difficulty sitting	18 (5.9)	Getting out of bed / turning over in bed	1 (0.3)
Difficulty lying down	16 (5.4)	Drinking	1 (0.3)
Feeling moody / miserable / irritable	14 (4.6)	Inability to physically defend oneself / partner	1 (0.3)
Difficulty sitting / standing / lying down	11 (3.7)	Sneezing	1 (0.3)
Control over life / life in general / daily living	10 (3.3)	Concern over childbirth / future childbirth	1 (0.3)
Fear of falling	7 (2.3)	Impact on choice of foot wear	1 (0.3)
Letting people down / meeting commitments	7 (2.3)		

*Baseline postal survey (n = 303). DIY: “do-it-yourself” handiwork.

Table 4. PGI test-retest reliability* for index score and by area changes.

PGI	n	PGI index score ICC (95% CI)
Closed format**	69	0.88 (0.81–0.92)
Blind format***	75	0.82 (0.72–0.88)
Blind format by number of area changes		
0 to 1 area changes	25	0.91 (0.80–0.96)
1.5 to 2.5 area changes	24	0.80 (0.57–0.91)
3 to 5 area changes	26	0.56 (0.23–0.78)

* AS and general health the same at 2 weeks by health transition. ** Closed: informed of baseline areas, but not allowed to change or add to list. *** Blind: blind to baseline areas. ICC: intraclass correlation coefficient.

general health is better have an average score improvement of 1.34. Those whose AS is worse have an average score deterioration of –0.80. All instruments have a stronger relationship with general health transition than with AS-specific transition. Of the 2 PGI formats, the informed and open format has a stronger relationship with both transition questions. The strongest linear relationships with both AS and general health transition were observed for the BASDAI ($p < 0.01$).

The informed and open format of the PGI produced moderate to large levels of responsiveness for groups of patients whose general or AS-specific health had improved or deteriorated according to transition question (Table 6b). In patients reporting an improvement or deterioration in health, responsiveness statistics over 0.5 were found for the informed and open format, representing a level of change that is at least one-half a standard deviation (SD) of the change scores for stable patients. Small to moderate levels of responsiveness were found for patients, indicating an improvement in general or AS-specific health, when completing the blind format of the PGI. Very low levels of responsiveness were found for patients, indicating deterioration in general or AS-specific health, when completing the blind format.

The BASDAI produced large levels of responsiveness for groups of patients whose AS-specific or general health had improved or deteriorated according to transition question responses. In general, the EQ-VAS produced moderate to high levels of responsiveness for groups of patients reporting improvement or deterioration in general or AS-specific health, and the SF-12 produced large levels of responsiveness for patients reporting improvement in general health only. Lower levels of responsiveness were found for the remaining instruments.

DISCUSSION

The measurement of an individual's subjective perception of HRQL is now considered a core component in the evaluation of health outcome[38]. Although many developers involve patients in item generation to ensure the representation of patient concerns, patient-assessed instruments typically use summated rating scales that use standardized items. By providing the opportunity for identification of areas of life that a patient deems to be of greatest importance, the PGI allows an individual's perspective to be considered within the evaluative process. This study undertook an evaluation of the PGI in a large and representative population of outpatients with AS. Instrument performance was assessed following self-completion within a postal survey.

The first step in the adoption of the PGI required the development of an AS-specific trigger list. The list addresses a wide diversity of areas such as relationships with family, fear of falling, ability to plan ahead, and the level of social embarrassment associated with poor posture and reduced mobility. It also captures patients' concern about the future direct and indirect consequences of the disease. For example, the effect of disease on the ability to work and the resulting financial impact. Many areas are distinctively associated with AS although many may differ between patients and over time, a feature common with other patient centered measures of HRQL[16,25]. While the majority of PGI trigger list areas were entered more frequently than supplementary areas introduced by individual patients, the additional areas highlight the diversity and individuality of HRQL as a concept, a diversity that is not captured by traditional summated rating scales. Further, end effects, where the majority of item scores accrue at the ceiling or floor[36], may be less likely in individualized measures, where patients representing the extremes of the disease severity spectrum can enter areas of personal importance.

Limited resources prevented stratified or purposive sampling to ensure representation of all possible disease scenarios in the development of the trigger list. Although the aim of the interviews was to identify areas of life affected by

Table 5. Correlation between the PGI and other patient assessed instruments. Postal survey (n = 343).

	ASQoL	BASDAI	Body Chart	EuroQol EQ-5D	EuroQol VAS	RLDQ	SF-12 MCS	SF-12 PCS
PGI index*	–0.58	–0.56	–0.42	0.55	0.60	–0.47	0.50	0.52
PGI-6 areas**	–0.57	–0.57	–0.42	0.55	0.60	–0.47	0.49	0.51
PGI-5 areas***	–0.56	–0.56	–0.40	0.55	0.58	–0.47	0.48	0.52

* All correlations statistically significant ($p < 0.01$). ** PGI score calculated following omission of area 7. *** PGI score calculated following omission of areas 6 and 7.

Table 6A. PGI blind format—mean score change (SD) and modified standardized response mean (MSRM) by 6 month AS-specific and general health transition.

	AS Health Transition					General Health Transition				
	Better (n = 16)		Same (n = 48)	Worse (n = 23)		Better (n = 14)		Same (n = 52)	Worse (n = 18)	
Instrument	Mean (SD)	MSRM	Mean (SD)	Mean (SD)	MSRM	Mean (SD)	MSRM	Mean (SD)	Mean (SD)	MSRM
PGI Blind	0.51 (1.6)	0.33	0.32 (1.5)	0.09 (1.5)	0.05	0.75 (1.4)	0.50	0.39 (1.5)	–0.24 (1.5)	–0.16
ASQoL	–0.66 (3.5)	–0.25	–0.02 (2.7)	1.46 (3.1)	0.54	–1.32 (3.4)	–0.49	0.19 (2.6)	1.57 (3.5)*	0.58
BASDAI	–0.78 (1.3)	–0.62	0.37 (1.2)	0.61 (1.4)**	0.48	–1.03 (1.3)	–0.84	0.36 (1.2)	0.87 (1.4)**	0.70
RLDQ	0.00 (2.9)	0.00	–1.07 (3.9)	1.60 (4.2)**	0.40	0.07 (2.9)	–0.02	–1.03 (3.7)	1.90 (4.9)*	0.50
Body Chart	–0.06 (24.8)	–0.002	8.97 (29.7)	20.65 (34.8)	0.69	1.07 (26.6)	0.03	12.65 (34.1)	10.66 (25.6)	0.31
EQ-5D	0.14 (0.25)	0.07	0.39 (1.96)	–0.09 (0.33)*	–0.04	0.22 (0.27)	0.11	–0.35 (1.8)	–0.12 (0.36)	–0.06
EQ-VAS	9.12 (22.7)	0.50	–1.58 (18.1)	–3.73 (22.0)	–0.21	12.35 (22.8)	0.78	–0.48 (15.8)	–7.20 (26.0)*	–0.45
SF-12 MCS	5.71 (11.2)	0.67	–1.60 (8.5)	0.62 (7.0)*	–0.07	7.75 (10.4)	0.88	–1.39 (8.8)	0.57 (5.8)**	0.06
SF-12 PCS	4.80 (6.1)	0.65	1.05 (7.3)	–1.19 (7.8)*	–0.16	4.72 (6.7)	0.66	1.37 (7.1)	–2.45 (7.6)*	–0.34

F test for linearity * $p < 0.05$; ** $p < 0.01$. MSRM: mean change in scores (6 months minus baseline) divided by SD of change scores in patients defined as stable. PGI: scored 0–10, where higher scores indicate better health related quality of life. Blind: blind to baseline areas; Informed & Open — informed of baseline areas and allowed to retain or change as required. ASQoL: scored 0–18, where lower scores indicate better health related quality of life. BASDAI: scored 0–10, where higher scores indicate greater disease activity. Body Chart: scored from 0 upwards, with no maximum score limit. Higher scores indicate greater levels of perceived body pain. RLDQ: scored 0–48, where higher scores indicate increased limitation in functional ability. EuroQol EQ-5D: scored –0.59 to 1.0, where –0.59 is the worst and 1.0 the best possible health. EuroQol-VAS: scored 0–100, where higher scores indicate better health states. SF-12 uses norm-based scoring from the general population. Scales are transformed: mean of 50 (SD 10), range 0–100.

Table 6B. PGI informed and open format — mean score change (SD) and modified standardized response mean (MSRM) by 6 month AS-specific and general health transition.

	AS Health Transition					General Health Transition				
	Better (n = 19)		Same (n = 36)	Worse (n = 20)		Better (n = 14)		Same (n = 42)	Worse (n = 19)	
Instrument[†]	Mean (SD)	MSRM[a]	Mean (SD)	Mean (SD)	MSRM	Mean (SD)	MSRM	Mean (SD)	Mean (SD)	MSRM
PGI Informed & Open[c]	0.75 (1.6)	0.50	–0.14 (1.5)	–0.80 (1.2)**	–0.53	1.34 (1.8)	1.00	–0.17 (1.3)	–1.01 (1.1)**	–0.78
ASQoL	–1.53 (4.0)	–0.43	–1.00 (3.5)	2.25 (2.3)**	0.64	–3.2 (4.2)	–1.43	0.31 (2.3)	0.84 (4.7)*	0.36
BASDAI	–1.32 (1.5)	–0.76	0.30 (1.7)	1.30 (1.9)**	0.76	–1.78 (1.3)	–1.43	–0.07 (1.2)	1.05 (2.7)**	0.84
RLDQ	–2.12 (4.1)	–0.48	–0.39 (4.4)	2.90 (5.8)*	0.65	–3.21 (4.1)	0.90	0.19 (3.5)	2.14 (7.3)*	0.61
Body Chart	–1.78 (19.8)	–0.09	3.48 (19.1)	5.71 (11.9)	0.30	4.50 (35.2)	0.74	–0.21 (6.1)	8.05 (16.2)	0.49
EQ-5D	0.13 (0.22)	0.69	–0.03 (0.22)	–0.26 (0.29)**	–1.23	0.14 (0.24)	0.77	–0.006 (0.18)	–0.21 (0.4)**	–1.16
EQ-VAS	13.89 (17.2)	0.87	–0.51 (15.9)	–10.50 (23.1)	–0.66	18.85 (16.2)	1.26	–3.81 (14.9)	–12.31 (23.85)	–0.82
SF-12 MCS	3.83 (8.09)	0.46	–1.06 (8.3)	–8.49 (9.5)**	–1.02	6.79 (10.6)	1.01	0.004 (6.7)	–8.01 (10.3)**	–1.19
SF-12 PCS	4.44 (7.2)	0.55	1.58 (8.0)	–3.06 (4.4)*	–0.12	6.01 (5.6)	1.02	0.30 (5.8)	–1.02 (10.0)	–0.10

F test for linearity * $p < 0.05$; ** $p < 0.01$.
MSRM: mean change in scores (6 months minus baseline) divided by SD of change scores in patients defined as stable. [†] Instruments as defined in Table 6A.

AS and considered important by patients, it is suggested that disease severity is not synonymous with a patient's level of quality of life[6,23]. Further, the trigger list acts only as a prompt for patients completing the PGI. Area selection and subsequent weighting is specifically individualized. Patients are not required nor expected to respond to each area in the list in the manner of a summated rating scale. The sex, age range, and duration of diagnosis of the interview participants suggests that the study population was broadly representative of patients with AS[3,4]. In addition, the sample size was supported by instrument development work described by other investigators[14,25]. No new significant themes emerged during the last few interviews, supporting the concept of sampling to redundancy[24].

The response rate for the postal survey was good at recruitment and followup, and compared favorably to other studies[17,25]. There was a statistically significant age difference between responders and nonresponders at baseline (46.09 vs 41.36 yrs, respectively). There was no significant difference by sex. The peak incidence of AS disease onset is between 25 and 34 years of age[39], and the bias of responders toward the older age group may reduce generalizability. For example, the results may have a greater relevance to patients with more severe disease.

Limited resources prevented further contact with nonresponders to investigate reasons for nonresponse. The saliency of questionnaire content is an important factor influencing response rates in mailed surveys, and the level

of importance attributed to the questionnaire by the respondent may have a greater influence over response than actual questionnaire length[40]. Completion and acceptability of questionnaires may be influenced by many factors — for example, time to complete, legibility and understanding of items, appearance and complexity of the questionnaire, and the possibility of distress when completing sensitive items[38,40]. Further, self-completion by the targeted individual cannot be guaranteed, but is an important consideration when assessing measurement properties of patient assessed and individualized instruments.

The other disease-specific measure of HRQL included in the study, the ASQoL, adopts a summated rating scale and includes many items frequently mentioned by patients completing the PGI. However, several areas considered important to the HRQL of patients with chronic disease[41,42] that are included in the PGI trigger list are omitted by the ASQoL. In particular, the influence of AS on work is not assessed. The effect of AS on work is an important issue for many people[1,43], and was the most frequently mentioned area for the PGI in our study. Omission from the ASQoL reflects the need for applicability to the wider AS population, recognizing that many patients may not be in employment. This example illustrates the benefit, in terms of content validity, afforded by individualized measurement. As with all standardized instruments[13], the content validity of the ASQoL is compromised because it does not reflect what constitutes AS-specific HRQL for all patients. All areas within the PGI are uniquely individual in both selection and weighting, fulfilling the requirements of an individualized and patient centered measure of HRQL[13].

Completion rates for the PGI were improved compared to levels reported previously[17,19], which could be due to changes in the instrument structure. The trigger list is more extensive than those used previously, which may help improve completion. During pretesting, several patients commented on the relevance of areas to their own life. However, 13.2% of patients failed to complete the baseline PGI sufficiently to allow calculation of a final score. Acceptability of the self-completed format may be improved by simplifying the "spending of points" in stage 3. A revised version of the PGI is currently being tested for patient acceptability. The developers of the PGI evaluated the performance of the blind format of the instrument only. All 3 completion formats of the PGI were assessed in our study: blind, informed and open, and closed formats. Evaluation of these different formats is required to ensure that patients, clinicians, and researchers are able to make meaningful interpretations when completing the instrument[44].

The test-retest reliability evaluation of the blind and the closed formats, and the influence of area change in stage 1 of the blind format PGI, increases our understanding of the contribution of each stage to overall reliability. The reliability of both formats was greater than 0.80 for the index score, and almost 0.90 for the closed format, supporting the use of both formats in group evaluation[36]. This is an improvement on all previous reliability estimates[17,19,45].

High levels of reliability were found for patients not changing any areas in stage 1, or scoring up to 1 point on the number of area changes (0.91). Although we used only a small sample size, this supports application of the blind format in individual evaluation (> 0.90). However, increased area changes reduced reliability, and when more than 3 area changes were made, reliability was not acceptable for group evaluation. In choosing the most appropriate followup format of the PGI, the tradeoff between reliability and content validity must be considered. The informed and open format of the PGI should be evaluated for test-retest reliability to determine the influence of area provision and freedom to change or retain areas as necessary.

Evidence for the construct validity of the PGI was provided by the moderate correlations with widely used disease-specific and generic instruments, which met our *a priori* hypotheses. These relationships may be a function of the alternative approach to measuring HRQL presented by the PGI, the role of explicit weighting, and the influence of areas relating to other health and non-health issues on the score. Users of future versions of the PGI should consider removing both areas 6 and 7, which did not appear to be making a large contribution to the index scores as determined by the levels of correlation. The sixth box should allow consideration of other aspects of the targeted disease not listed in the 5 available boxes. The levels of correlation between the PGI and generic measures of HRQL that we found were greater than those reported by others comparing the original format of the PGI to the SF-36[17,19]. This could be due to improved reliability and validity or the greater effect of AS on health than disorders assessed in earlier studies, for example, low back pain, menorrhagia, peptic ulcer, and varicose veins.

The informed and open format of the PGI showed a moderate level of responsiveness for self-perceived improvement and deterioration in both AS and general health. The blind format was not responsive to change and did not correlate significantly with AS or general health transition. More information on the role of the different completion formats for the PGI in patients indicating change in health is required. An indication of the number of area changes in stage 1 of the PGI (identifying areas), as calculated for test-retest reliability, may reveal information about change in areas over the 6 month period in patients indicating improvement or deterioration in health, and the relative responsiveness of the PGI associated with this change. However, the results of investigation of instrument responsiveness should be interpreted with caution due to the small sample size when the 2 formats of the PGI are considered separately. The RLDQ and EQ-5D produced moderate and

small levels of responsiveness, respectively, in a larger population[46,47]. The measurement properties of the PGI should be assessed in larger sample sizes to improve confidence in the results including specific formats — blind, informed and open, or closed.

In conclusion, the PGI offers a unique patient centered and individualized approach to the evaluation of disease-specific health related quality of life. The open and dynamic nature of the PGI places the patient at the center of the evaluative process, and was developed to provide a sufficiently short and simple instrument that would be feasible for application following self-completion in postal surveys. Evidence suggests that the instrument is acceptable to patients with AS in a self-completed format and provides broader item coverage than other AS-specific measures of health related quality of life that adopt a summated rating scale format. The lack of comparative evidence for the different formats of the PGI makes recommendation difficult. The closed format has a high level of reliability, but evidence is lacking for the informed and open format. Compared to the closed format the informed and open format more closely follows the individuality of measurement proposed by the PGI developers. The enhanced content validity, together with evidence for instrument responsiveness, supports application of the informed and open format of the PGI for evaluative purposes at a group level.

ACKNOWLEDGMENT

We are grateful to all patients who completed questionnaires. We also thank the following Consultant Rheumatologists for allowing access to patient databases and local physiotherapists for their support: Prof. Roger Sturrock and Fiona Gough; Prof. Ian Haslock, Dr. Mike Plant and Kay West; Dr. Tom Price, Carol David, and Louise Preston; Prof. Hill Gaston and Julie Isaacson; Dr. Paul Creemer and Rachel Lewis; all consultant rheumatologists, nursing and clinic staff from the Staffordshire Rheumatology Centre, with particular thanks to Jackie Waterfield for assistance with data collection. Thanks to Dr. Kelvin Jordan for statistical advice.

REFERENCES

1. Ward MM. Quality of life in patients with ankylosing spondylitis. Rheum Dis Clin North Am 1998;24:815-27.
2. Haywood KL. Health outcomes in ankylosing spondylitis: an evaluation of patient-based and anthropometric measures. [DPhil thesis]. York: University of York; 2000, 353 p.
3. Russell AS. Ankylosing spondylitis – history. In: Klippel JH, Dieppe PA, editors. Rheumatology. 2nd ed. London: Mosby; 1998:14.1-2.
4. Dziedzic K. Ankylosing spondylitis. In: David C, Lloyd J, editors. Rheumatological physiotherapy. London: Mosby; 1998:97-114.
5. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. Ann Intern Med 1993;118:622-9.
6. Aronson KJ. Quality of life among persons with multiple sclerosis and their caregivers. Neurology 1997;48:74-80.
7. van der Heijde D, Bellamy N, Calin A, Dougadas M, Khan MA, van der Linden S. Preliminary core sets for endpoints in ankylosing spondylitis. J Rheumatol 1997;24:2225-9.
8. Reynolds S, Doward LC, Spoorenberg A, et al. The development of the Ankylosing Spondylitis Quality of Life Questionnaire (ASQoL) [abstract]. Qual Life Res 1999;8:651.
9. Chamberlain A, Jones P, Paul E, Garnham R, Garrod R, Bestall J. The development of the ASQoL: A quality of life instrument specific to ankylosing spondylitis. 1998. Website, Department of Health, UK. PCD Programme — Commissioned Research. Ref. A3215es. [Cited November 18, 2002]; available from http://www.doh.gov.uk/research/swro/rd/national/pcd/funded/completed/
10. Stratford P, Gill C, Westaway M, Binkley J. Assessing disability and change on individual patients: a report of a patient specific measure. Physiotherapy Canada 1995;47:258-63.
11. Carr AJ, Thompson PW, Kirwan JR. Quality of life measures. Br J Rheumatol 1996;35:275-81.
12. Tugwell P, Bombardier C, Buchanan W, Goldsmith CH, Grace E, Hanna B. The MACTAR Patient Preference Disability Questionnaire. An individualized functional priority approach for assessing improvement in physical disability in clinical trials in rheumatoid arthritis. J Rheumatol 1987;14:446-51.
13. Carr A, Higginson IJ. Measuring quality of life: Are quality of life measures patient centred? BMJ 2001;322:1357-60.
14. Guyatt GH, Berman LB, Townsend M, Pugsley SO, Chambers LW. A measure of quality of life for clinical trials in chronic lung disease. Thorax 1987;42:773-8.
15. O'Boyle CA, McGee H, Hickey A. The Schedule for the Evaluation of Individual Quality of Life (SEIQoL): Administration manual. Dublin: Department of Psychology, Royal College of Surgeons in Ireland;1993.
16. Carr AJ. A patient-centred approach to evaluation and treatment in rheumatoid arthritis: the development of a clinical tool to measure patient-perceived handicap. Br J Rheumatol 1996;35:921-32.
17. Ruta D, Garratt A, Leng M, Russell IT, Macdonald LM. A new approach to the measurement of quality of life — the Patient Generated Index. Med Care 1994;32:1109-26.
18. Garratt AM, Ruta DA. The Patient Generated Index. In: Joyce CRB, O'Boyle CA, McGee HM, editors. Individual quality of life: Approaches to conceptualisation and assessment. Amsterdam: Harwood Academic Publishers; 1999:105-18.
19. Ruta DA, Garratt AM, Russell IT. Patient centred assessment of quality of life for patients with four common conditions. Qual Health Care 1999;8:22-9.
20. Jenkinson C, Stradling J, Petersen S. How should we evaluate health status? A comparison of three methods in patients presenting with obstructive sleep apnoea. Qual Life Res 1998;7:95-100.
21. Herd RM, Tidman MJ, Ruta D, Hunter JAA. Measurement of quality of life in atopic dermatitis: correlation and validation of two different methods. Br J Dermatol 1997;136:502-7.
22. van der Linden SJ, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis — a proposal for modification of the New York Criteria. Arthritis Rheum 1984;27:361-8.
23. Whalley D, McKenna SP, de Jong Z, van der Heijde D. Quality of life in rheumatoid arthritis. Br J Rheumatol 1997;36:884-8.
24. Bowling A. Research methods in health. Investigating health and health services. Buckingham, UK: Open University Press; 1997.
25. Jenkinson C, Fitzpatrick R, Peto V. The Parkinson's Disability Questionnaire. User manual for the PDQ-39, PDQ-8 and PDQ summary index. Oxford, UK: Health Services Research Unit, Department of Public Health, University of Oxford; 1998.
26. Garratt AM, Ruta DA, Abdalla MI, Russell IT. Responsiveness of the SF-36 and a condition-specific measure of health for patients with varicose veins. Qual Life Res 1996;5:223-34.
27. Garrett S, Jenkinson T, Kennedy G, Whitelock H, Gaisford P, Calin A. A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis Disease Activity Index. J Rheumatol 1994;12:2286-91.
28. Abbott CA, Helliwell PS, Chamberlain MA. Functional assessment in ankylosing spondylitis — Evaluation of a new self-administered

questionnaire and correlation with anthropometic variables. Br J Rheumatol 1994;33:1060-6.
29. Dziedzic KSG. The body chart: A further sketch towards a fuller picture of ankylosing spondylitis [PhD thesis]. Staffordshire: University of Keele; 1997, 251 p.
30. EuroQol Group. EuroQol: a new facility for the measurement of health-related quality of life. Health Policy 1990;16:199-208.
31. Ware JE, Kosinski M, Keller SD. SF-12: How to score the SF-12 physical and mental health summary scales. Boston: The Health Institute, New England Medical Centre; 1995.
32. Hurst NP, Kind P, Ruta D, Hunter M, Stubbings A. Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). Br J Rheumatol 1997;36:551-9.
33. Hurst NP, Ruta DA, Kind P. Comparison of the MOS Short Form-12 (SF-12) health status questionnaire with the SF-36 in patients with rheumatoid arthritis. Br J Rhematol 1998;37:62-9.
34. Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures. Statistics and strategies for evaluation. Controlled Clin Trials 1991;12:142S-58S.
35. Fitzpatrick R, Zieblans S, Jenkinson C, Mowat A. Transition questions to assess outcomes in rheumatoid arthritis. Br J Rheumatol 1993;32:807-11.
36. Streiner DL, Norman GR. Health measurement scales. A practical guide to their development and use. 2nd ed. Oxford: Oxford Medical Publications; 1995.
37. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. Psychol Bull 1979;86:420-8.
38. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. Health Technol Assess 1998;2:1-74.
39. Carbone LD, Cooper C, Michet CJ, Atkinson EJ, O'Fallon WM, Melton LJ. Ankylosing spondylitis in Rochester, Minnesota, 1935-1989: Is the epidemiology changing? Arthritis Rheum 1992;35:1476-82.
40. McColl E, Jacoby A, Thomas L, et al. Design and use of questionnaires: a review of best practice applicable to surveys of health service staff and patients. Health Technol Assess 2001; 5:1-256.
41. Fitzpatrick R. The measurement of health status and quality of life in rheumatological disorders. Ballieres Clin Rheumatol 1993; 7:297-317.
42. Ware JE. Preface — International Quality of Life Assessment (IQOLA) Project. J Clin Epidemiol 1998;51:891-2.
43. Guillemin F, Challier B, Urlacher F, Vancon G, Pourel J. Quality of life in ankylosing spondylitis: validation of the Ankylosing Spondylitis Arthritis Impact Measurement Scales 2, a modified Arthritis Impact Measurement Scales Questionnaire. Arthritis Care Res 1999;2:157-62.
44. Jenkinson C, Ruta D, Peterson S, Mowat A, Stradling J. Should respondents be allowed to nominate new areas at follow-up in individualised quality of life assessment? An evaluation of two scoring methods using the Patient Generated Index (PGI) [abstract]. Qual Life Res 1998;7:612.
45. Macduff C, Russell E. The problem of measuring change in individual health-related quality of life by postal questionnaire: use of the patient-generated index in a disabled population. Qual Life Res 1998;7:761-9.
46. Haywood KL, Garratt AM, Jordan K, Dziedzic K, Dawes PT. Disease-specific patient-assessed measures of health outcome in ankylosing spondylitis: reliability, validity and responsiveness. Rheumatology 2002;41:1295-302.
47. Haywood KL, Garratt AM, Dziedzic K, Dawes PT. Generic measures of health-related quality of life in ankylosing spondylitis: reliability, validity and responsiveness. Rheumatology 2002;41:1380-7.