

QUESTIONNAIRE ON THE PERCEPTIONS OF PATIENTS ABOUT SHOULDER SURGERY

JILL DAWSON, RAY FITZPATRICK, ANDREW CARR

From the University of Oxford and the Nuffield Orthopaedic Centre, Oxford, England

We developed a 12-item questionnaire for completion by patients having shoulder operations other than stabilisation. A prospective study of 111 patients was undertaken before operation and at follow-up six months later. Each patient completed the new questionnaire and the SF36 form. Some filled in the Stanford Health Assessment Questionnaire (HAQ). An orthopaedic surgeon assessed the Constant shoulder score.

The single score derived from the questionnaire had a high internal consistency. Reproducibility, examined by test-retest reliability, was found to be satisfactory. The validity of the questionnaire was established by obtaining significant correlations in the expected direction with the Constant score and the relevant scales of the SF36 and the HAQ. Sensitivity to change was assessed by analysing the differences between the preoperative scores and those at follow-up. Changes in scores were compared with the patients' responses to postoperative questions about their condition. The standardised effect size for the new questionnaire compared favourably with that for the SF36 and the HAQ. The new questionnaire was the most efficient in distinguishing patients who said that their shoulder was much better from all other patients.

The shoulder questionnaire provides a measure of outcome for shoulder operations which is short, practical, reliable, valid and sensitive to clinically important changes.

J Bone Joint Surg [Br] 1996;78-B:593-600.

Received 20 November 1995; Accepted 16 January 1996

J. Dawson, DPhil, Research Officer
R. Fitzpatrick, PhD, University Lecturer
Department of Public Health and Primary Care, University of Oxford,
Gibson Building, Radcliffe Infirmary, Woodstock Road, Oxford OX2
6HE, UK.

A. Carr, FRCS, Consultant Orthopaedic Surgeon
The Nuffield Orthopaedic Centre, Windmill Road, Headington, Oxford
OX3 7LD, UK.

Correspondence should be sent to Dr R. Fitzpatrick.

©1996 British Editorial Society of Bone and Joint Surgery
0301-620X/96/41207 \$2.00

The prevalence of shoulder pain has recently been reported to be approximately 7% overall, rising with age to around 20% in individuals over 70 years (Chakravarty and Webley 1990; Chard et al 1991; Badley and Tennant 1992; Van Schaardenburg et al 1994). It represents a substantial proportion of new referrals to hospital specialists (OPCS 1986; Bamji et al 1990), and in middle-age has important socio-economic implications because of time lost from work (Ekberg et al 1994).

The pain usually arises from disorders of the periarticular soft tissue, particularly the rotator cuff, although it is secondary to arthritis in about 5% of cases (Uthoff et al 1990). There is considerable uncertainty as to the effectiveness of the various methods of management of these problems (Ogilvie-Harris and Demazière 1993; Dalton 1994; Van der Windt et al 1995). Few assessments of the methods which are available use independent outcome measures.

Scoring systems which measure the outcome of orthopaedic surgery generally derive from clinical and radiological data and depend on the surgeon's judgement (MacDonald 1993). Patients and surgeons may differ in their concerns and priorities (Wright, Rudicel and Feinstein 1994), and it is increasingly recognised that methods are required to elicit the patient's perception of the outcome (Amadio 1993).

Research has shown that patients can provide reliable and valid judgements of health status and of the benefits of treatment (Fitzpatrick et al 1992). A number of questionnaires have been developed to examine a wide range of medical and surgical problems. These have proved useful in assessing joint replacement (Kantz et al 1992; Cleary et al 1993), but because they have been designed to evaluate a diverse range of health problems, they are often rather long. They also contain questions which may not relate to a specific problem, which limits their use in assessing the results of a particular intervention (Patrick and Deyo 1989).

There are advantages in using questionnaires designed to address the patient's perception of a single condition. These shorter questionnaires may be just as sensitive to changes of importance to patients (Katz et al 1992; Fitzpatrick et al 1993) and much simpler to use.

We have devised and tested a short questionnaire to assess the outcome of operations on the shoulder, excluding stabilisation.

PATIENTS AND METHODS

Development of the questionnaire. Initially, we interviewed 20 patients attending an outpatient shoulder clinic to identify how they experienced and reported problems with their shoulders. From these results and from established questionnaires, we drafted a 22-item questionnaire and tested it on 20 new patients. They were also given a second copy of the questionnaire, and asked to complete it at home on the following day and return it in a prepaid envelope. They were invited to add their comments to this copy and to include any further shoulder problems which were not addressed by it.

It became clear that there was a distinct group of patients with a tendency towards recurrent dislocation or subluxation of the shoulder. They were characterised by the anticipation of problems arising in relation to very specific activities. We excluded this group from the study in order to concentrate on patients presenting with a painful shoulder related to a degenerative or inflammatory condition.

The original questionnaire was modified after the pilot study and the revised version was tested on two further groups of patients until its final form was established. This contains 12 items, each of which has five response categories (Table I). Each item is scored from 1 to 5, from least to most difficulty or severity, and combined to produce a single score with a range from 12 (least difficulties) to 60 (most difficulties).

We then tested the questionnaire in a prospective study to see whether it was internally consistent, reproducible, valid and sensitive to clinical change (Cox et al 1992). Internal consistency examines whether the items measure a single underlying concept. Reproducibility is concerned with whether the questionnaire yields the same results on repeated trials under the same conditions. Validity determines whether it measures what it aims to measure; this can be examined by two methods. Content validity shows whether items in a questionnaire cover the intended topics clearly. Construct validity, the extent to which the questionnaire supports predefined hypotheses, is assessed by whether it produces an anticipated set of relationships with other variables such as clinical evidence. Sensitivity to change, or responsiveness, reflects the ability to detect clinically significant changes.

Between March 1994 and mid-June 1995 we recruited 111 consecutive patients into a prospective study from the preadmission assessment clinic at the Nuffield Orthopaedic Centre, Oxford. There were 51 women and 60 men with a median age of 57.4 years (19.9 to 87.5; SD 15.0). Of these, 79 (67.5%) were diagnosed as having an impingement syndrome with or without a rotator-cuff tear, or calcified deposits in the rotator-cuff tendon. The existence or extent of rotator-cuff tears could be confirmed only at the time of operation in many cases. Twenty-four (20.5%) were diagnosed as having either primary or secondary osteoarthritis or inflammatory arthritis as the main cause of their symp-

toms, 12 (10.3%) had an adhesive capsulitis ('frozen shoulder') and two (1.7%) were not given a formal diagnosis before operation. Six patients had two diagnoses, giving 117 diagnoses for 111 patients.

Of the 111 patients all but five had an operation during the study period. Two were medically unfit, one had complete relief from symptoms, and two had their surgery delayed beyond the study period. The postoperative data which we present relate to the first 56 patients to reach a six-month assessment. Of these, two had a manipulation of their shoulder under anaesthesia, 28 had arthroscopic surgery including subacromial decompression, distension or exploration, three had an open acromioplasty, 13 repair of a rotator-cuff tear and 13 patients had either a hemi- or a total shoulder arthroplasty. Two patients had more than one type of procedure.

Internal consistency. Internal consistency was tested by using Cronbach's alpha (Cronbach 1951) before operation and at six months. This summarises the internal correlations of all items in a scale. The higher the alpha coefficient (range 0.0 to 1.0) the more consistent is the scale and the greater the likelihood that it is tapping an underlying single variable on the questionnaire. We examined correlations of all items with the overall score and also whether Cronbach's alpha was improved by removal of any item.

Reproducibility. Reproducibility (test-retest reliability) was assessed by asking 60 patients at the preoperative stage to complete and return a second questionnaire 24 hours after the first. The data were examined by the coefficient of reliability according to the method described by Bland and Altman (1986).

Construct validity. This was examined by Pearson correlation coefficients between the total score of the questionnaire and other related measures obtained at the same assessment both before operation and at six months. We considered that the scores for the questionnaire should correlate moderately with Constant shoulder scores rated by an orthopaedic surgeon and with scores from the two other health-status questionnaires (SF36 and Stanford Health Assessment Questionnaire).

The Constant shoulder score (Constant and Murley 1987) contains both subjective and objective elements. The subjective assessments of pain and activities of daily living are allocated 15 and 20 points respectively, out of a total of 100. Objective measurements of the active range of motion and power are allocated 40 and 25 points, respectively. We used a goniometer to make a reliable assessment of the range of motion. Shoulder power was measured in kilograms of resisted abduction using a spring balance, as described by Constant. As with the new 12-item questionnaire, the Constant shoulder score has been found to be inappropriate in conditions of instability (MacDonald 1993).

The SF36 is a 36-item questionnaire widely used to measure health status (Ware and Sherbourne 1992). It provides scores on physical functioning, social functioning,

Table I. Preoperative and postoperative scores obtained using the 12-item shoulder questionnaire

Item	Scoring categories	Preop (n = 111)		Postop (n = 56)	
		Number	Percentage	Number	Percentage
<i>During the past four weeks</i>					
1) How would you describe the worst pain you had from your shoulder?	1 None	0	0.0	6	10.7
	2 Mild	3	2.7	19	33.9
	3 Moderate	31	27.9	20	35.7
	4 Severe	56	50.5	7	12.5
	5 Unbearable	21	18.9	4	7.1
2) Have you had any trouble dressing yourself because of your shoulder?	1 No trouble at all	11	9.9	25	44.6
	2 Little trouble	37	33.3	20	35.7
	3 Moderate trouble	47	42.3	9	16.1
	4 Extreme difficulty	15	13.5	2	3.6
	5 Impossible to do	1	0.9	0	0.0
3) Have you had any trouble getting in and out of a car or using public transport because of your shoulder? (whichever you tend to use)	1 No trouble at all	41	36.9	38	67.9
	2 Very little trouble	36	32.4	12	21.4
	3 Moderate trouble	25	22.5	4	7.1
	4 Extreme difficulty	8	7.2	2	3.6
	5 Impossible to do	1	0.9	0	0.0
4) Have you been able to use a knife and fork - at the same time?	1 Yes, easily	62	55.9	43	76.8
	2 With little difficulty	26	23.4	7	12.5
	3 With moderate difficulty	16	14.4	2	3.6
	4 With extreme difficulty	4	3.6	0	0.0
	5 No, impossible	3	2.7	4	7.1
5) Could you do the household shopping on your own?	1 Yes, easily	26	23.4	22	39.3
	2 With little difficulty	26	23.4	18	32.1
	3 With moderate difficulty	22	19.8	3	5.4
	4 With extreme difficulty	16	14.4	2	3.6
	5 No, impossible	21	18.9	11	19.6
6) Could you carry a tray containing a plate of food across a room?	1 Yes, easily	35	31.5	29	51.8
	2 With little difficulty	29	26.1	13	23.2
	3 With moderate difficulty	21	18.9	5	8.9
	4 With extreme difficulty	9	8.1	3	5.4
	5 No, impossible	17	15.3	6	10.7
7) Could you brush/comb your hair with the affected arm?	1 Yes, easily	7	6.3	19	33.9
	2 With little difficulty	17	15.3	18	32.1
	3 With moderate difficulty	40	36.0	5	8.9
	4 With extreme difficulty	22	19.8	5	8.9
	5 No, impossible	25	22.5	9	16.1
8) How would you describe the pain you usually had from your shoulder?	1 None	1	0.9	9	16.1
	2 Very mild	2	1.8	18	32.1
	3 Mild	25	22.5	8	14.3
	4 Moderate	67	60.4	17	30.4
	5 Severe	16	14.4	4	7.1
9) Could you hang your clothes up in a wardrobe, using the affected arm?	1 Yes, easily	9	8.1	19	33.9
	2 With little difficulty	16	14.4	18	32.1
	3 With moderate difficulty	37	33.3	7	12.5
	4 With great difficulty	20	18.0	3	5.4
	5 No, impossible	29	26.1	9	16.1
10) Have you been able to wash and dry yourself under both arms?	1 Yes, easily	31	27.9	33	58.9
	2 With little difficulty	22	19.8	11	19.6
	3 With moderate difficulty	27	24.3	5	8.9
	4 With great difficulty	19	17.1	1	1.8
	5 No, impossible	12	10.8	6	10.7
11) How much has pain from your shoulder interfered with your usual work (including housework)?	1 Not at all	2	1.8	17	30.4
	2 A little bit	6	5.4	16	28.6
	3 Moderately	36	32.4	12	21.4
	4 Greatly	53	47.7	8	14.3
	5 Totally	14	12.6	3	5.4
12) Have you been troubled by pain from your shoulder in bed at night?	1 No nights	5	4.5	25	44.6
	2 Only 1 or 2 nights	7	6.3	8	14.3
	3 Some nights	30	27.0	9	16.1
	4 Most nights	31	27.9	4	7.1
	5 Every night	38	34.2	10	17.9

role limitations due to physical problems, role limitations due to emotional problems, general mental health, energy, bodily pain and general perceptions of health. Scores for each item range from 0 (poor) to 100 (good).

The Stanford Health Assessment Questionnaire (HAQ) (Fries, Spitz and Young 1982) is a measure of functional limitations in which patients rate, on a four-point scale, the degree of difficulty which they have experienced during the last week with 20 tasks grouped into eight sections covering various aspects of daily life, namely, dressing, rising, hygiene, reach, walking, eating, grip and activities. The responses are based on the American Rheumatism Association functional class rating system of 'normal' (no difficulty = 0), 'adequate' (some difficulty = 1), 'limited' (much difficulty = 2), and 'unable to do' (= 3). For any task for which the respondent has help or uses some aid or device to assist them, the score for that section is recorded as 'limited'. Scores are based on the highest within each group and may also be expressed as an overall mean score ranging from 0 to 3. Pain is measured separately on a simple visual analogue scale (VAS); patients are asked to place a mark on a 15 cm line to represent the severity of pain experienced in the past week. In our study the VAS was modified by specifying 'pain in your shoulder' when normally the site of any pain experienced would remain unspecified.

All patients were asked to complete the new shoulder questionnaire and the SF36 both at the preoperative assessment and at the six-month follow-up. Only 56 patients were asked to complete the HAQ in addition at the preoperative stage.

Sensitivity to change. The sensitivity to change of the study questionnaire was examined by comparing scores before and six months after operation. The scores of 56 patients were available at follow-up for these analyses (n = 28 for the HAQ).

We calculated the effect sizes for the shoulder questionnaire, the SF36 and the HAQ. This is a method of calculating the extent of change measured by a questionnaire in a standardised way which allows comparison between different questionnaires (Kazis, Anderson and

Meenan 1989). In our study this was determined as the difference between the mean preoperative and postoperative scores, divided by the standard deviation of the preoperative scores. An effect size of 1.0 is equivalent to a change of one standard deviation in the sample.

Sensitivity was also examined by one patient satisfaction item and three transition items (Ziebland et al 1992). Patients were asked to say how successful they thought the operation had been and also to assess the change that had occurred concerning their shoulder problems, pain and day-to-day life as a result of the procedure. The first question offered four possible categories of response, and the transitional questions offered five. The patients who gave the most positive response in each case were compared with all others by way of their mean change scores (the preoperative minus the postoperative score) for each of the health-status questionnaires.

Statistical significance was measured by the *t*-test.

RESULTS

Table I gives the individual scores obtained for patients on the shoulder questionnaire before operation and at six months. Before operation, the summed score for the questionnaire had a median value of 33.0 (18 to 55), a mean of 36.3 (95% CI 34.6 to 37.9), and at follow-up a median value of 24 (12 to 54) and a mean of 25.96 (95% CI 23.0 to 28.9).

Internal consistency. Cronbach's alpha for the study questionnaire was 0.89 at the preoperative assessment (n = 111) and 0.92 at the six-month follow-up (n = 56). All items correlated with the total score at >0.4 (Table II).

Reproducibility. In the test-retest sample, the differences in scores between the first and second tests were plotted against their means. The scatter appeared normal and showed the same variability across the range of scores at all levels of disability and severity. The estimated mean (-0.12) of score differences was not significantly different from 0. The coefficient of reliability was calculated as 6.8 using the Bland and Altman (1986) method and 95% of

Table II. Internal consistency of the shoulder questionnaire, preoperatively and postoperatively

Question	Preoperative			Postoperative		
	Mean score (sd)	Item-total correlation	Alpha if item removed	Mean score (sd)	Item-total correlation	Alpha if item removed
1 Worst pain from shoulder	3.86 (0.75)	0.40	0.89	2.71 (1.06)	0.59	0.92
2 Trouble with dressing	2.62 (0.87)	0.65	0.88	1.79 (0.85)	0.84	0.91
3 Trouble with transport	2.03 (0.99)	0.62	0.88	1.46 (0.79)	0.54	0.92
4 Using a knife and fork	1.74 (1.02)	0.64	0.88	1.48 (1.10)	0.71	0.92
5 Doing household shopping alone	2.82 (1.43)	0.72	0.87	2.32 (1.51)	0.64	0.92
6 Carrying a tray of food	2.50 (1.41)	0.67	0.88	2.00 (1.35)	0.80	0.91
7 Brushing/combing hair	3.37 (1.17)	0.70	0.88	2.41 (1.45)	0.71	0.92
8 Usual level of shoulder pain	3.86 (0.71)	0.47	0.89	2.80 (1.24)	0.75	0.91
9 Hanging clothes in wardrobe	3.40 (1.25)	0.69	0.88	2.38 (1.42)	0.72	0.92
10 Washing under both arms	2.63 (1.34)	0.64	0.88	1.86 (1.31)	0.74	0.92
11 Work interference due to pain	3.64 (0.84)	0.60	0.88	2.36 (1.21)	0.75	0.92
12 Pain in bed at night	3.81 (1.12)	0.42	0.89	2.39 (1.55)	0.56	0.92
	Preoperative Cronbach Alpha: 0.89			Postoperative Cronbach Alpha: 0.92		

Table III. Correlation between the 12-item shoulder questionnaire and the Constant shoulder assessment, the SF36 and the HAQ assessments

Test	Correlation coefficient	
	Preoperative (n = 111)	Postoperative (n = 56)
Constant	-0.74*	-0.75*
SF36		
Physical activity	-0.61*	-0.62*
Pain	-0.66*	-0.68*
Mental health	-0.39*	-0.54*
Social function	-0.55*	-0.61*
Role - physical	-0.41*	-0.61*
Role - emotional	-0.37*	-0.51*
Energy	-0.52*	-0.59*
Health perceptions	-0.34*	-0.42*
HAQ		
Pain VAS	0.49*	0.71*
Disability index	0.86*	0.80*

* p < 0.01

score differences fell between 0 ± 6.8. Overall, 83% of score differences lay between 0 ± 4 points.

Construct validity. The study questionnaire correlated well with the Constant scores both before operation and at the six-month follow-up (Table III). There was also significant agreement between the study questionnaire and the scales of the SF36 and the HAQ with related content, particularly in the areas of physical function and pain.

Sensitivity to change. Patients reported a substantial improvement in health status at the six-month follow-up. Effect sizes were larger for the study questionnaire than for any of the scales of the SF36 or HAQ disability index, and equivalent to the HAQ pain VAS which had been presented to patients in a modified, shoulder-specific format (Table IV).

Table V gives further evidence for the sensitivity of the new questionnaire to change in clinical condition. For the patient satisfaction item and each of the three transition questions, the new questionnaire was shown to distinguish clearly between those patients who rated the most positive

change in their shoulder condition after surgery and those who said that the change had been only slightly better or worse. In each case the level of significance achieved was much higher for the new questionnaire than for any of the relevant parts of either the SF36 or the HAQ.

DISCUSSION

We have developed and tested a short 12-item questionnaire which patients find easy to complete and which provides reliable, valid and responsive data regarding their perception of shoulder problems. It is intended for use as an outcome measure during specialist treatment and imposes very little burden on the patients. Few reported any difficulties in completing it.

The finding that the questionnaire is unsuitable for patients with shoulder instability is not surprising and a similar experience has been reported with the Constant shoulder score (MacDonald 1993).

The items are internally consistent and reproducible and the questionnaire may therefore be considered to be at least as reliable as the clinical scores used to assess outcomes (Koran 1975).

To ensure content validity, every effort was made to derive the items to be included in the questionnaire from exploratory interviews with patients, rather than by imposing clinical assumptions. Draft versions of the questionnaire were tested on patients and the final content agreed only when patients understood it and felt that no important items had been omitted. Construct validity was tested by examining the level of agreement of the questionnaire with clinical data and with scales from existing health-status questionnaires. All correlations were in the expected direction with poorer scores on our shoulder questionnaire correlating with poorer scores for the Constant, the SF36 and the HAQ assessments. Correlations were highest in the assessment of pain and physical function.

Responsiveness, or sensitivity to clinically important

Table IV. Comparison by effect size of the 12-item shoulder questionnaire with the SF36 and the HAQ in regard to ability to measure clinical change in the patients after shoulder surgery

Test	Mean score		At six-month follow-up	95% CI	Effect size
	Preoperative	95% CI			
12-item questionnaire (n = 56)	36.3	34.6 to 37.9	26.0	23.0 to 28.9	1.2
SF36 (n = 56)					
Health perceptions	68.7	64.8 to 72.7	66.2	59.8 to 72.7	0.1
Health change	43.9	39.9 to 47.9	62.1	54.8 to 69.3	-0.8
Physical activity	56.5	51.6 to 61.4	63.2	55.9 to 70.4	-0.3
Pain	36.6	33.2 to 40.1	58.0	51.0 to 65.0	-1.1
Role limitation - physical	22.0	16.3 to 27.8	41.8	32.1 to 51.5	-0.7
Role limitation - mental	61.2	52.6 to 78.3	66.1	54.6 to 77.5	-0.1
Mental health	70.6	67.1 to 74.1	73.9	68.4 to 79.4	-0.2
Energy	51.4	47.0 to 55.7	52.5	46.5 to 58.5	-0.1
Social function	70.5	64.9 to 76.0	74.4	67.1 to 81.7	-0.1
HAQ (n = 28)					
Pain VAS	8.5	7.6 to 9.4	4.7	3.3 to 6.2	1.1
Disability index	1.3	1.1 to 1.5	0.8	0.5 to 1.1	0.7

Table V. Comparison of change scores between the 12-item shoulder questionnaire and the SF36 and HAQ in regard to the patients' assessment of change as measured by one patient satisfaction item and three transition items

Question 1 'How successful do you feel your shoulder operation has been?'			
Test	'Very successful' Mean change score (n = 23)	All other responses Mean change score (n = 32)	t-test
12-item shoulder score (n = 56)	14.96	8.63	2.84*
SF36 (n = 56)			
Health perceptions	-1.30	1.63	NS
Health change	-23.91	-14.84	NS
Physical activity	-10.91	-7.00	NS
Pain	-26.26	-15.28	NS
Role limitation - physical	-34.09	-17.19	NS
Role limitation - mental	-3.17	-11.46	NS
Mental health	-2.43	-3.75	NS
Energy	-6.09	-0.78	NS
Social function	-14.49	-2.43	NS
HAQ (n = 28)			
Pain VAS	6.00	1.67	2.70†
Disability index	0.79	0.37	2.11†

Transition item 1 'How have the problems related to your shoulder changed since having the operation?'

	'Much better' Mean change score (n = 31)	All other responses Mean change score (n = 25)	t-test
12-item shoulder score (n = 56)	15.55	5.72	5.11‡
SF36 (n = 56)			
Health perceptions	-0.26	0.40	NS
Health change	-27.42	-7.0	-3.23*
Physical activity	-13.88	-1.96	-2.59†
Pain	-27.04	-10.67	-2.60†
Role limitation - physical	-36.67	-9.00	-2.77*
Role limitation - mental	-14.94	0.00	NS
Mental health	-4.29	-1.14	NS
Energy	-4.29	0.48	NS
Social function	-13.33	2.65	-2.44†
HAQ (n = 28)			
Pain VAS	5.39	0.90	3.08*
Disability index	0.85	0.28	NS

Transition item 2 'Have you experienced a change in the level of pain from your shoulder since having the operation?'

	'Much better' Mean change score (n = 35)	All other responses Mean change score (n = 21)	t-test
12-item shoulder score (n = 56)	14.43	5.71	4.16‡
SF36 (n = 56)			
Health perceptions	1.20	-1.90	NS
Health change	-25.00	-7.14	-2.68†
Physical activity	-10.88	-4.74	NS
Pain	-24.51	-11.64	NS
Role limitation - physical	-30.88	-13.10	NS
Role limitation - mental	-9.09	-6.35	NS
Mental health	-4.34	-1.14	NS
Energy	-4.29	0.48	NS
Social function	-13.33	-2.65	-2.44†
HAQ (n = 28)			
Pain VAS	4.26	1.41	NS
Disability index	0.67	0.23	2.31†

Transition item 3 'How has your shoulder operation changed your day-to-day life so far?'

	'Much better' Mean change score (n = 25)	All other responses Mean change score (n = 31)	t-test
12-item shoulder score (n = 56)	16.84	6.58	5.45‡
SF36 (n = 56)			
Health perceptions	-0.08	0.13	NS
Health change	-26.00	-12.10	-2.09†
Physical activity	-11.04	-6.72	NS
Pain	-28.70	-12.54	-2.55†
Role limitation - physical	-42.71	-9.68	-3.40*
Role limitation - mental	-18.84	0.00	NS
Mental health	-4.96	-1.68	NS
Energy	-5.40	0.16	NS
Social function	-12.00	3.58	NS
HAQ (n = 28)			
Pain VAS	4.99	1.94	NS
Disability index	0.86	0.28	-3.65*

* p < 0.01

† p < 0.05

‡ p < 0.001

change, is least likely to be examined despite its being of the greatest importance in any form of prospective outcome study (Guyatt, Walter and Norman 1987). When the sensitivity to change of our questionnaire was compared with that of the SF36 and the HAQ the standardised effect size was highest for the study questionnaire. The new questionnaire also showed a much greater ability to distinguish between subgroups of patients based on their own assessment of the change in their shoulder after operation, indicating that it is particularly sensitive to improvements obtained by surgery. This reflects its specific design to assess the outcome of operations on the shoulder.

Our questionnaire may be compared with another patient-based shoulder disability questionnaire recently produced for use in primary care (Croft et al 1994). This uses dichotomous (yes/no) response categories throughout, which may increase its repeatability at the expense of reduced sensitivity, but unfortunately the authors did not address this aspect in their paper. It would be interesting to contrast the performance of the two different questionnaires used in similar circumstances.

Patient-based measures are intended to supplement, not replace, conventional measures of outcome. It is becoming increasingly clear, however, that systematic studies are required to examine the effects on outcomes of alternative treatments and surgical techniques in the management of orthopaedic conditions (Murray, Carr and Bulstrode 1993; Van der Windt et al 1995). Because variations in outcomes may be quite modest, such studies will need large sample sizes to detect differences and will therefore have to be multicentred. Standardised patient-based measures of outcome have obvious advantages, particularly if they are highly sensitive to clinical change. Similarly, a particularly long period of follow-up is required to detect differences in outcome for many orthopaedic conditions. Simple questionnaires may enhance the feasibility of long-term studies,

particularly if the responses can be obtained by post.

We acknowledge financial support by grant from Oxford Regional Health Authority (Audit). We would also wish to thank the nursing staff in the preadmission and outpatient clinics, as well as individual patients for giving their valuable time.

No benefits in any form have been received or will be received from a commercial party related directly or indirectly to the subject of this article.

REFERENCES

- Amadio PC.** Editorial. Outcomes measurement: more questions; same answers. *J Bone Joint Surg [Am]* 1993;75-A:1583-4.
- Badley EM, Tennant A.** Changing profile of joint disorders with age: findings from a postal survey of the population of Calderdale, West Yorkshire, United Kingdom. *Ann Rheum Dis* 1992;51:366-71.
- Bamji AN, Dieppe PA, Haslock DI, Shipley ME.** What do rheumatologists do? a pilot audit study. *Br J Rheumatol* 1990;29:295-8.
- Bland JM, Altman DG.** Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;i:307-10.
- Chakravarty KK, Webley M.** Disorders of the shoulder: an often unrecognised cause of disability in elderly people. *Br Med J* 1990;300:848-9.
- Chard MD, Hazleman R, Hazleman BL, et al.** Shoulder disorders in the elderly: a community survey. *Arthritis Rheum* 1991;34:766-9.
- Cleary PD, Reilly DT, Greenfield S, et al.** Using patient reports to assess health-related quality of life after total hip replacement. *Qual Life Res* 1993;2:3-11.
- Constant CR, Murley AHG.** A clinical method of functional assessment of the shoulder. *Clin Orthop* 1987;214:160-4.
- Cox D, Fitzpatrick R, Fletcher A, et al.** Quality of life assessment: can we keep it simple? *J Roy Stat Soc Series A* 1992;155:353-93.
- Croft P, Pope D, Zonca M, O'Neill T, Silman A.** Measurement of shoulder related disability: results of a validation study. *Ann Rheum Dis* 1994;5:525-8.
- Cronbach L.** Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297-34.
- Dalton SE.** The conservative management of rotator cuff disorders. *BJ Rheumatol* 1994;33:663-7.
- Ekberg K, Bjorkqvist B, Malm P, Bjerre-Kiely B, Axelson O.** Controlled two year follow-up of rehabilitation for disorders in the neck and shoulders. *Occup Environ Med* 1994;51:833-8.
- Fitzpatrick R, Fletcher A, Gore S, et al.** Quality of life measures in health care. I: Applications and issues in assessment. *BMJ* 1992;305:1074-7.

- Fitzpatrick R, Ziebland S, Jenkinson C, Mowat A, Mowat A.** A comparison of the sensitivity to change of several health status instruments in rheumatoid arthritis. *J Rheumatol* 1993;20:429-36.
- Fries JF, Spitz PW, Young DY.** The dimensions of health outcomes: the Health Assessment Questionnaire, disability and pain scales. *J Rheumatol* 1982;9:789-93.
- Guyatt G, Walter GS, Norman G.** Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis* 1987;40:171-8.
- Kantz ME, Harris WJ, Levitsky K, Ware JE, Davies AR.** Methods for assessing condition-specific and generic functional status outcomes after total knee replacement. *Med Care* 1992;30(5 Suppl):240-52.
- Katz JN, Larson MG, Phillips CB, Fossel AH, Liang MH.** Comparative measurement sensitivity of short and longer health status instruments. *Med Care* 1992;30:917-25.
- Kazis LE, Anderson JJ, Meenan RF.** Effect sizes for interpreting changes in health status. *Med Care* 1989;27:178-89.
- Koran LM.** The reliability of clinical methods, data and judgements. *N Engl J Med* 1975;293:642-6.
- MacDonald DA.** The shoulder and elbow. In: Pynsent PB, Fairbank JC, Carr A, eds. *Outcome measures in orthopaedics*. Oxford, etc: Butterworth Heinemann, 1993:144-73.
- Murray DW, Carr AJ, Bulstrode C.** Survival analysis of joint replacements. *J Bone Joint Surg [Br]* 1993;75-B:697-704.
- OPCS.** *Morbidity statistics from general practice 1981-82*. London, HMSO, 1986.
- Ogilvie-Harris DJ, Demazière A.** Arthroscopic debridement versus open repair for rotator cuff tears: a perspective cohort study. *J Bone Joint Surg [Br]* 1993;75-B:416-20.
- Patrick DL, Deyo RA.** Generic and disease-specific measures in assessing health status and quality of life. *Med Care* 1989;27:(3 Suppl) 217-32.
- Uthoff HK, Sarkar K.** An algorithm for shoulder pain caused by soft-tissue disorders. *Clin Orthop* 1990;254:121-7.
- Van der Windt DA, Van der Heijden GJ, Scholten RJ, Koes BW, Bouter LM.** The efficacy of non-steroidal anti-inflammatory drugs (NSAIDs) for shoulder complaints: a systematic review. *J Clin Epidemiol* 1995;48:691-704.
- Van Schaardenburg D, Van den Brande KJ, Ligthart GJ, et al.** Musculoskeletal disorders and disability in persons aged 85 and over: a community survey. *Ann Rheum Dis* 1994;53:807-11.
- Ware JE, Sherbourne CD.** The MOS 36-item short-form health survey (SF36). I. Conceptual framework and item selection. *Med Care* 1992;30:473-83.
- Wright JG, Rudicel S, Feinstein AR.** Ask patients what they want: evaluation of individual complaints before total hip replacement. *J Bone Joint Surg [Br]* 1994;76-B:229-34.
- Ziebland S, Fitzpatrick R, Jenkinson C, Mowat A, Mowat A.** Comparison of two approaches to measuring change in health status in rheumatoid arthritis: the health assessment questionnaire (HAQ) and modified HAQ. *Ann Rheum Dis* 1992;51:1202-5.