EXTENDED REPORTS

The responsiveness of the Shoulder Disability Questionnaire

Daniëlle A W M van der Windt, Geert J M G van der Heijden, Andrea F de Winter,
Bart W Koes, Walter Devillé, Lex M Bouter

Abstract

Objective—To evaluate the responsiveness of the Shoulder Disability Questionnaire (SDQ).

Methods—The study was conducted within the framework of an observational study on shoulder disorders in primary care. After first presentation of their complaints to the general practitioner and after one and six months, participants completed the SDQ, a single question on functional status (FSQ), and an ordinal 11 point scale for the severity of pain (PSS). Responsiveness of the SDQ was evaluated compared with that of the FSQ and PSS, by calculating responsiveness ratios and by plotting receiver operating characteristic (ROC) curves. Recovery according to the patient was used as an external criterion for clinically relevant improvement (complete recovery or much improved on a six point Likert scale was denoted as clinically relevant improvement).

Results—A total of 349 consecutive patients with shoulder disorders were enrolled in the observational study. Response rates ranged between 96% and 89%. Responsiveness ratios were slightly higher for the PSS compared with the SDQ (2.53 versus 2.22 at one month, 2.24 versus 1.89 at six months). The area under the ROC curve was 0.84 for both the SDQ and the PSS, and 0.72 for the FSQ.

Conclusion—The results of this study confirm the responsiveness of the SDQ, making it a useful instrument to assess functional disability in longitudinal studies.

(*Ann Rheum Dis* 1998;57:82–87)

Institute for Research in Extramural Medicine, Vrije Universiteit, Amsterdam, the Netherlands
D A W M van der Windt
A F de Winter
B W Koes
W Devillé
L M Bouter

Department of Epidemiology, Maastricht University, the Netherlands
G J M G van der Heijden

Institute for Rehabilitation Research, Hoensbroek, the Netherlands
G J M G van der Heijden

Department of Epidemiology and Biostatistics, Faculty of Medicine, Vrije Universiteit, Amsterdam, the Netherlands
W Devillé
L M Bouter

Correspondence to:
Mrs D A W M van der Windt, Institute for Research in Extramural Medicine, Faculty of Medicine, Vrije Universiteit, Van der Boechorststraat 7, 1081 BT Amsterdam, the Netherlands.

According to several epidemiological studies, the prevalence of shoulder pain in the general population may be as high as 6–11% under the age of 50 years, increasing to 16–25% in the elderly.[1–5] Shoulder pain is often considered to be of a benign nature with a favourable prognosis, but symptoms may be persistent or recurrent in many patients.[6–10] Restricted range of motion and shoulder pain interferes considerably with activities of daily living, and may be associated with increased sick leave and incapacity in the elderly.[3 11 12]

Therefore, the treatment of shoulder disorders is usually aimed at improvement of functional status, next to pain reduction. Consequently, outcome measurements should include an instrument for the evaluation of functional status, in addition to the assessment of perceived benefit and pain severity. Over 50 randomised clinical trials have been reported in the medical literature, evaluating the efficacy of medication, physiotherapy or corticosteroid injections for shoulder disorders.[13–15] Approximately one third of these studies included some assessment of functional status, usually consisting of a single question on the ability to perform daily activities. Only a few trials have included a more formalised assessment of functional status.[16 17]

Until recently, a functional status questionnaire for shoulder disability was not available. During the past five years, several instruments have been developed nearly simultaneously,[18–21] including the Shoulder Disability Questionnaire (SDQ). The responsiveness of the SDQ has been described for a comparatively homogeneous population of patients with shoulder pain, participating in a randomised, placebo controlled trial of physiotherapy.[21] In this trial the SDQ was completed by the patient under the guidance of a research assistent. Additional evidence for the responsiveness of the SDQ is not yet available, although the questionnaire is presently being used in research and clinical practice in several countries. As the properties of an outcome measure may vary between populations and study settings, it is important not to rely on the results of a single study. The objective of our study was to investigate the responsiveness of the SDQ in an observational study in a primary care population.

Methods

THE SHOULDER DISABILITY QUESTIONNAIRE

The SDQ is a pain related disability questionnaire, which contains 16 items describing common situations that may induce symptoms in patients with shoulder disorders. All items refer to the preceding 24 hours. An English translation of the SDQ can be found in the Appendix. Response options are either 'yes', 'no', or 'not applicable'. The 'not applicable'

category should be used when the situation at issue has not occurred during the preceding 24 hours. A final score is calculated by dividing the number of positively scored items by the total number of applicable items, and subsequently multiplying the score by 100, resulting in a final score ranging between 0 (no disability) and 100 (all applicable items positive).

PATIENTS

The study was carried out within the framework of an observational study of shoulder disorders in general practice, described in full elsewhere.[10 22] Selection criteria for participation were: (1) informed consent; (2) age 18 years or older; (3) ability to complete questionnaires (no dementia, sufficient knowledge of the Dutch language); (4) shoulder pain originating from within the shoulder girdle (no known neurological or vascular disorders, neoplasms, referred pain from internal organs or systemic rheumatic conditions); (5) no fractures or luxations; (6) no consultation for the afflicted shoulder in the preceding 12 months.

MEASUREMENTS AND INSTRUMENTS

At inclusion, the participants completed a baseline questionnaire, containing questions on demographic variables, previous complaints, precipitating events, and duration of symptoms at presentation. In addition to the SDQ, the participants scored the average severity of their shoulder pain during the day (referring to the preceding week) on an 11 point ordinal scale, ranging from 0 'no pain' to 10 'very severe pain' (pain severity score). The scores were linearly transformed to a score between 0 and 100 to facilitate comparisons between the SDQ and the pain severity score. The questionnaire finally contained a single question on the ability to perform daily activities (functional status question (FSQ)). The three response options were: (1) 'little discomfort during daily activities'; (2) 'much discomfort during daily activities', and (3) 'unable to perform daily activities'.

Recovery since baseline was recorded by the patient at one and six months follow up on a six point Likert scale, ranging from 'complete recovery' to 'much worse'. This global measure of improvement was used as an external criterion for the evaluation of responsiveness, as a gold standard for clinical improvement in shoulder pain is not available. Complete recovery and much improved were denoted as clinical improvement, little improved and no change as clinical stability, and little worse and much worse as clinical deterioration.

RESPONSIVENESS

As the SDQ was designed for use in longitudinal studies, we were particularly interested in its responsiveness: the ability to detect clinically relevant changes over time.[23] The responsiveness of the SDQ was studied in comparison with that of a single question on functional disability: the FSQ, to evaluate if a single question would suffice for the assessment of shoulder disability. The responsiveness of the SDQ was furthermore compared with that of the severity

of pain, scored on the pain severity scale (PSS). The following two methods were used for the evaluation of responsiveness.

Responsiveness ratio
Guyatt *et al*[24–26] illustrated the concept of responsiveness by demonstrating the analogy with signal to noise ratios. The signal constitutes the clinically relevant change one wishes to detect. The noise represents the measurement error over which this change should be detected—that is, the within subject variability unrelated to true clinical change. Thus, according to Guyatt *et al*, the most appropriate measure of responsiveness relates a clinically relevant change to the variability of the change score in stable patients.[25]

Responsiveness ratios were calculated for the SDQ and PSS at one and six months follow up, by calculating the ratio of the mean change score in clinically improved patients (a clinically relevant change) to its variability (SD) in clinically stable patients. (As the three point ordinal scale of the FSQ does not enable the calculation of a SD, responsiveness ratios were not calculated for the FSQ.) If the responsiveness ratio is larger than 1, the mean change score in clinically improved patients exceeds the measurement error and the instrument may be considered to be responsive, to an extent that is proportional to the magnitude of the responsiveness ratio.

Receiver operating characteristic (ROC) curves
The ability to detect changes over time can also be quantified by the construction of ROC curves.[27] ROC curves synthesise information on the sensitivity and specificity for discriminating between patients reporting clinical improvement and patients reporting clinical stability. ROC curves for change scores of the SDQ, PSS, and FSQ at one month and six months follow up were created by plotting the true positive proportion (sensitivity) versus the false positive proportion (100-specificity) for multiple cut off points. Two potential cut off points for changes of the SDQ were highlighted: (1) the mean change of SDQ scores in clinically improved patients (the numerator of the responsiveness ratio), and (2) the point closest to the upper left corner of the ROC curve, which is assumed to represent the optimal trade off between sensitivity and specificity for detecting clinical improvement.[28]

The area under the curve can be interpreted as the probability of correctly identifying an improved patient from randomly selected pairs of improved and stable patients.[27 29] An area under the curve of 1.0 means perfect discrimination between these two health states. An instrument that does not discriminate for improvement will have an area under the curve of 0.5. The further the curve is in the upper left corner, the higher the responsiveness of the instrument.[30]

MISSING VALUES

To improve response rates in the observational study, participants who reported complete recovery after six months were allowed to skip a

Table 1 Patient characteristics of the study population. Response rate to the baseline questionnaire was 96% (n = 335)

	Patients (n)	%
Sex (male)	146	44
Previous episodes of shoulder complaints	154	46
Duration of current episode at baseline		
< 1 week	48	14
1 week–1 month	122	37
1 month–6 months	115	34
> 6 months	50	15
Acute onset	160	48
Diagnosis at baseline ★		
Capsular syndrome	75	22
Acute bursitis	56	17
Acromioclavicular syndrome	14	4
Subacromial syndrome	159	48
Unclear	31	9
Age (y)	Mean 49.6	SD14.4

★ Diagnostic criteria were based on guidelines issued by the Dutch College of General Practitioners. Detailed information can be found elsewhere.[22]

number of questions, including the SDQ, PSS, and FSQ. However, it is possible that participants who reported complete recovery still experienced symptoms during some of their daily activities. Results at one month follow up, at which time a complete date set was available for the SDQ, show that 74% of all completely recovered patients scored either 0 or 1 items positive on the SDQ, resulting in a median score of 0, and a mean score of only 8. We conservatively decided to substitute missing SDQ scores with 8, whenever patients reported complete recovery after six months (n=157). As similar data at one month follow up were not available for the PSS, missing severity scores in case of complete recovery at one month (n=74) and at six months (n=157) were, conservatively, substituted with 10. Analogously, missing values for the FSQ of completely recovered patients were substituted with 1 (that is, 'little discomfort during daily activities').

Results

During the recruitment period of one year, 349 patients met the selection criteria and were included in the observational study. The baseline questionnaire was returned by 335 participants (96%). Table 1 gives patient characteristics of the study population, including age, sex, diagnosis, and the duration of symptoms at presentation.

Items 2, 3, 4, 6, and 10 of the SDQ (see Appendix) were scored positive most frequently at baseline: lying on the afflicted side (79% of all participants), putting on a coat or sweater (88%), performing usual daily activities (83%), moving the arm (87%), and reaching above

shoulder level (89%). Items 11 and 16: writing or typing, and irritability were scored positive by only 28% and 29%, respectively.

Response rates to the follow up questionnaires were 92% (n = 321) after one month and 89% (n = 312) after six months. Table 2 presents the severity of shoulder symptoms at baseline and at follow up.

After one month, 44% of all participants reported much improvement or complete recovery (that is, clinical improvement). This proportion increased to 68% after six months. Deterioration of symptoms was reported by few participants (7%). The mean SDQ score improved from 67 to 32, the mean PSS score from 72 to 30 at six months follow up (after substitution of missing values). The proportion of patients who were able to perform daily activities with little discomfort (FSQ) increased from 26% at baseline to 77% after six months.

RESPONSIVENESS RATIOS

The mean change of SDQ scores in clinically improved patients at one and six months follow up was 40 and 51, respectively (table 3). In patients reporting no change or little improvement (clinical stability), mean changes of SDQ scores were slightly larger than zero, and a slightly negative mean change was seen in patients with deteriorating conditions. The mean change score in clinically improved patients was somewhat smaller for the SDQ than for the PSS, resulting in slightly lower responsiveness ratios: 2.22. versus 2.53 at one month and 1.89 versus 2.24 at six months (table 3).

ROC CURVES

Figure 1 presents the ROC curves generated for changes of the SDQ, PSS, and FSQ at one month follow up. True positive proportions (sensitivity) and false positive proportions (100-specificity) for the discrimination between clinical improvement and clinical stability are plotted for multiple cut off points. The area under the curve was 0.84 for both the SDQ and the PSS, and 0.72 for the FSQ. At six months follow up the area under the curve was slightly higher for all instruments: 0.88 for the SDQ, 0.86 for the PSS, and 0.79 for the FSQ.

Figure 1 highlights two potential cut off points for changes of the SDQ after one month. The use of a cut off point of 40 (the mean change in clinically improved patients; that is an improvement of at least six items) showed a very high specificity to detect improvement (98%), but would imply a considerable number of false negatives (sensitivity is only 46%). A cut off point of 18.75 (an improvement of at least three items on the SDQ) approximates the optimal trade off between sensitivity (74%) and specificity (77%).

Discussion

In our experience, patients find the 16 items of the SDQ easy to complete within a few minutes time. This confirms the suitability for use in observational studies, but also in randomised

Table 2 Severity of shoulder symptoms at baseline and after one and six months

	Baseline (n = 335)	1 month (n = 319)	6 months (n = 308)
Recovery since baseline★			
Clinical improvement		142 (44)	209(68)
Clinical stability		156 (49)	77(25)
Clinical deterioration		21 (7)	22(7)
FSQ †	85 (25)	202 (63)	235(76)
PSS ‡	72 (24)	43 (28)	30(27)
SDQ ¶	67 (23)	47 (31)	32(31)

★ Frequency (%), † functional status question, three point ordinal scale: 'only little discomfort during daily activities', frequency (%), ‡ pain severity score, 100 point scale, mean (SD), ¶ Shoulder Disability Questionnaire, 100 point scale, mean (SD).

Table 3   Mean change scores (SD) and responsiveness ratios for the Shoulder Disability Questionnaire (SDQ) and the pain severity score (PSS) after one and six months

	1 month			6 months		
	Number	SDQ	PSS	Number	SDQ	PSS
Clinical improvement*	142	40 (30)	48 (28)	209	51 (26)	56 (26)
Clinical stability	156	4 (18)	14 (19)	77	5 (27)	16 (25)
Clinical deterioration	21	−2 (21)	0.5 (21)	22	−5 (12)	−0.5 (16)
Responsiveness ratio †		2.22	2.53		1.89	2.24

* Substitution of missing values was conducted for patients reporting complete recovery: for 74 patients at one month (PSS only) and for 157 patients at six months (PSS and SDQ),
† Responsiveness ratio: the ratio of the mean change score in clinically improved patients to the variability (SD) of change scores in clinically stable patients.
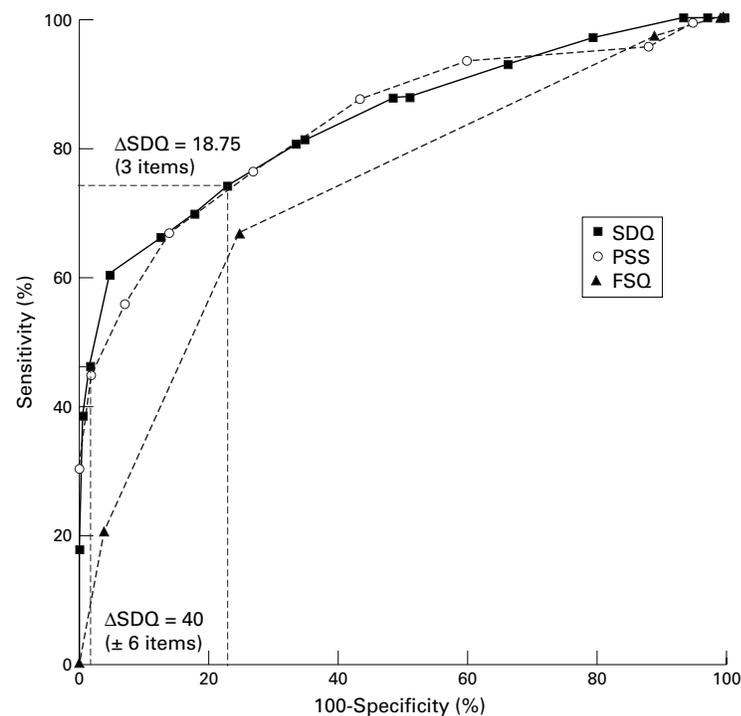

Figure 1   Receiver operator characteristic (ROC) curves for ΔSDQ (Shoulder Disability Questionnaire, ΔPSS (Pain Severity Scale) and ΔFSQ (Functional Status Question) at one month. True positive rate (sensitivity) and false positive rate (100-specificity) are for discriminating between patients reporting clinical improvement or clinical stability. Potential cut off points for the SDQ: ΔSDQ = 18.75: sensitivity 74%, specificity 77% (optimal trade off); ΔSDQ = 40: (mean change in clinically improved patients) sensitivity 46%, specificity 98%.

separate scales; one for pain (five items) and one for disability (eight items). All items refer to the preceding week and are scored on visual analogue scales. The Shoulder Pain and Disability Index has been made suitable for telephone administration by converting the visual analogue scales to a 0–10 numeric scale.[31] Which of these three instruments shows the best performances with respect to responsiveness can only be answered by a direct comparison between the instruments in a single study population.

Our observational study included a large number of consecutive patients consulting a general practitioner for shoulder pain. We tried to optimise response rates to the mailed questionnaires by allowing participants who reported complete recovery to skip parts of the questionnaire. A high response rate was necessary to ensure a valid and reliable study of the course of shoulder disorders and to investigate prognostic indicators of outcome.[10] Although this measure seemed to be successful (given the high response rates), it caused some difficulties for our evaluation of responsiveness. Our assumptions for the replacement of missing values in recovered patients were based on the complete data set of the SDQ at one month follow up. We presume that our rather conservative algorithm (a score of 8 for the SDQ and 10 for the PSS in case of complete recovery) has not introduced major bias, but we decided to conduct sensitivity analyses to evaluate alternative assumptions.

An even more conservative algorithm, using a replacement value of 16 (two items positively scored), resulted in a slightly lower responsiveness ratio for the SDQ after six months (1.69 compared with 1.89). A sensitivity analysis, using the median score of 0 as a replacement value, instead of the mean score of 8, showed, of course, a higher responsiveness ratio after six months for the SDQ (2.14). A complete data set would probably have yielded responsiveness ratios between these two values. In a final sensitivity analysis we assessed the distribution of SDQ scores in patients reporting recovery at one month (in the complete data set) and replaced the missing values resulting from recovery at six months using the same distribution. Using this method, the responsiveness ratio for the SDQ at six months was calculated at 1.91, which is clearly quite similar to the 1.89 presented in our paper.

Our results confirm the responsiveness of the SDQ. The area under the ROC curve for the SDQ was 0.84 after one month and 0.88 after six months, which is not much different from other results reported for the SDQ (0.72)[21] and for the Shoulder Pain and Disability Index (0.91).[31] The responsiveness ratios were clearly larger than 1, confirming the ability of the SDQ to detect a clinically important change.

The definition of a clinically important change depends on a subjective judgement by either a patient or clinician. In the absence of a gold standard for clinical change, we decided to use the improvement of symptoms as reported by the patient as an external criterion. A more objective assessment of recovery was not

trials that may include a wide array of outcome measures, demanding considerable time and devotion of the participants.

The SDQ shows similarity with the Disability Questionnaire designed by Croft *et al*,[19] which consists of 22 items, also refers to the preceding 24 hours and is scored on a dichotomous scale. In a general practice population, the most frequently reported items on the Disability Questionnaire were sleep disturbances, moving the arm or hand, and difficulties in dressing, which is consistent with the results of our study. The 'not applicable' response category of the SDQ, which has not been included in the Disability Questionnaire, may be a useful addition to a dichotomous item scale. It may prevent difficulties and missing values for items on activities that have not been performed during the relevant period of time.

The Shoulder Pain and Disability Index[18] has a different scoring system. It consists of two

available as a physical examination was not part of our observational study. None the less, we feel that improvement as perceived by the patients is a valid estimation of true recovery. Patients who reported deterioration of symptoms were excluded from the analysis of responsiveness, analogously to the definition of responsiveness by Guyatt *et al*.[25] Including patients who have deteriorated will produce an increase of the noise in the signal to noise ratio, which depends on the number of such patients in the population at issue. Yet, it may give a more adequate reflection of clinical change in the study population at issue. A repeated analysis of responsiveness including all patients in our study resulted in negligible differences of ROC curves and responsiveness ratios, because of the small proportion of patients reporting deterioration (7%) in our study population.

The mean change score for the SDQ in clinically improved patients was 40. This rather large change may not be the smallest change that might be considered to be clinically relevant. A change score of 18.75 (three items on the SDQ) might approximate the smallest, yet clinically relevant improvement. The responsiveness ratio for this cut off point is nearly equal to 1 (18.75/18). A cut off point of 18.75 proved to be close to the optimal trade off of sensitivity and specificity to detect clinical improvement. However, the selection of an optimal cut off point also depends on the purpose of the study. The selection of a high cut off point (for example 40) will result in few false positives, but will miss a large proportion of clinically improved patients. The reverse holds for a low cut off point.

The responsiveness of the SDQ was studied in contrast with that of a single question on functional status (FSQ). The assessment of functional status might be substantially simplified if a single question would perform as well as the SDQ. The smaller area under the ROC curve for the FSQ, however, showed that this was clearly not the case.

Both the responsiveness ratio and the ROC curve enabled direct comparisons between the PSS and the SDQ. The results of the responsiveness ratios were somewhat more favourable for the PSS than for the SDQ, because of the larger mean change score of the PSS in clinically improved patients. Such differences were not reflected by the ROC curves; depending on the selected cut off point, the ROC curves show only very small differences between instruments. The fact that differences between instruments may become more apparent when expressed by responsiveness ratios compared with ROC curves, can also be noticed in other clinimetric studies.[28 32]

Consistent with the results of Van der Heijden *et al*,[21] we expected the responsiveness of the PSS to be slightly more favourable compared with the SDQ. Pain reflects a different aspect of symptom severity. Although these outcome measures will often correlate, they may also differ systematically. When pain has subsided, patients may still have difficulties with daily activities, such as driving a car, doing household chores, etc. Both outcome measures are clearly

relevant and can be used next to each other in longitudinal studies, as they will often provide complementary information on recovery.

It would be interesting to evaluate the test-retest reliability and the construct validity of the SDQ, especially when the SDQ will be used for discriminative purposes; for example to discriminate between patients with different severity levels of shoulder disability. However, our study was not particularly suited to evaluate these properties, as many of our participants showed a rapid recovery. This might be more adequately evaluated in a study population consisting of patients with chronic conditions, of which the severity of shoulder disability ranges from slight to very severe.

In conclusion, the SDQ seems to be a valuable functional status instrument for both intervention studies and observational studies using mailed questionnaires, and is easy to complete within a few minutes. The results of our study confirm the responsiveness in a primary care population, emphasising the usefulness of the SDQ in longitudinal studies. The properties of the SDQ should be further explored in different settings and different study populations. As yet, there is insufficient evidence to justify a preference for one particular instrument: the Shoulder Pain and Disability Index,[18 20 31] the Disability Questionnaire,[19] or the SDQ. Future studies should directly compare these instruments and assess their relative performances, before embarking on the development of yet another instrument for shoulder disability.

1 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.
2 Bjelle A. Epidemiology of shoulder problems. Baillière's Clin Rheumatol 1989;3:437-51.
3 Chakravarty K, Webley M. Shoulder joint movement and its relationship to diability in the elderly. J Rheumatol 1993;20:1359-61.
4 Chard MD, Hazleman R, Hazleman BL, King RH, Reiss BB. Shoulder disorders in the elderly: a community survey. Arthritis Rheum 1991;34:766-9.
5 Croft P. Soft tissue rheumatism. In: Silman AJ, Hochberg MC, eds. *Epidemiology of the rheumatic diseases*. Oxford: Oxford Medical, 1993:375-421.
6 Binder AI, Bulgen DY, Hazleman BL, Roberts S. Frozen shoulder: a long-term prospective study. Ann Rheum Dis 1984;43:361-4.
7 Chard MD, Satelle LM, Hazleman BL. The long-term outcome of rotator cuff tendinitis - A review study. Br J Rheumatol 1988;27:385-9.
8 Reeves B. The natural history of the frozen shoulder syndrome. Scand J Rheumatol 1975;4:193-6.
9 Vecchio PC, Kavanagh RT, Hazleman BL, King RH. Community survey of shoulder disorders in the elderly to assess the natural history and effects of treatment. Ann Rheum Dis 1995;54:152-4.
10 Windt DAWM van der, Koes BW, Boeke AJP, Devillé W, Jong BA de, Bouter LM. Shoulder disorders in general practice. Prognostic indicators of outcome. Br J Gen Pract 1996;46:519-23.
11 Cunningham DA, Paterson DH, Himann JE, Rechnitzer PA. Determinants of independence in the elderly. Can J Appl Physiol 1993;18:243-54.
12 Bassey EJ, Morgan K, Dallosso HM, Ebrahim SBJ. Flexibility of the shoulder joint measured as range of abduction in a large representative sample of men and women over 65 years of age. Eur J Appl Physiol 1989;58:353-60.
13 Windt DAWM van der, Heijden GJMG van der, Scholten RJPM, 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.
14 Heijden GJMG van der, Windt DAWM van der, Kleijnen J, Koes BW, Bouter LM, Knipschild PG. Steroid injections for shoulder disorders: a systematic review of randomized clinical trials. Br J Gen Pract 1996;46:309-16.

15 van der Heijden GJMG, Bouter LM, Beckerman H, Bie RA de, Oostendorp RAB. Effectiviteit van fysiotherapie bij schouderklachten. Nederlands Tijdschrift voor Fysiotherapie 1992;102:38-46.
16 Downing DS, Weinstein A. Ultrasound therapy of subacromial bursitis. A double blind trial. Phys Ther 1986;66:194-9.
17 Nykänen M. Pulsed ultrasound treatment of the painful shoulder. A randomized, double-blind, placebo-controlled study. Scand J Rehabil Med 1995;27:105-8.
18 Roach KE, Budiman-Mak E, Songsiridej N, Lertratanakul Y. Development of a Shoulder Pain and Disability Index. Arthritis Care Res 1991;4:143-9.
19 Croft P, Pope D, Zonca M, O'Neill TO, Silman A. Measurement of shoulder related disability: results of a validation study. Ann Rheum Dis 1995;53:525-8.
20 Beaton DE, Richards RR. Measuring function of the shoulder. A cross-sectional comparison of five questionnaires. J Bone Joint Surg 1996;78[A]:882-90.
21 Van der Heijden GJMG. Shoulder disability questionnaire: design and responsiveness of a functional status measure. In: Shoulder disorder treatment: efficacy of ultrasound and electrotherapy [PhD thesis]. Maastricht: Datawyse/University Press Maastricht, 1996: 79-91.
22 Windt DAWM van der, Koes BW, Jong BA de, Bouter LM. Shoulder disorders in general practice. Incidence, patient characteristics and management. Ann Rheum Dis 1995;54:959-64.
23 Liang MH. Evaluating measurement responsiveness. J Rheumatol 1995;22:1191-2.
24 Guyatt GH, Bombardier C, Tugwell P. Measuring disease-specific quality of life in clinical trials. Can Med Assoc J 1986;134:889-95.
25 Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. J Chron Dis 1987;40:171-8.
26 Guyatt GH, Kirschner B, Jaeschke R. Measuring health status: what are the necessary measurement properties? J Clin Epidemiol 1992;45:1341-5.
27 Deyo RA, Centor RM. Assessing the responsiveness to change of functional scales to clinical change: an analogy to diagnostic test performance. J Chron Dis 1986;39:897-906.
28 Stucki G, Liang MH, Fossel AH, Katz JN. Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. J Clin Epidemiol 1995;48:1369-78.
29 Hanley JA, McNeil BJ. The meaning and use of the area under a Receiver Operating Characteristic (ROC) curve. Radiology 1982;143:29-36.
30 Schouw YT van der, Verbeek ALM, Ruijs JHJ. ROC curves for the initial assessment of new diagnostic tests. Fam Pract 1992;9:506-11.
31 Williams JW, Holleman DR, Simel DL. Measuring shoulder function with the Shoulder Pain and Disability Index. J Rheumatol 1995;22:727-32.
32 Beurskens AJHM, Vet HCW de, Köke AJA. Responsiveness of functional status in low back pain. A comparison of different instruments. Pain 1996;65:71-6.

Appendix

The Shoulder Disability Questionnaire (SDQ)

Instructions

When your shoulder hurts, you may find it difficult to do certain things you normally do. This list contains 16 sentences that people have used to describe themselves when they have shoulder pain. When you read the sentences, you may find that some stand out because they describe you today (last 24 hours). As you read the list, think of yourself today (last 24 hours). Ask yourself if you performed the activity.

Examples for completion

- You did not perform the activity in the last 24 hours, for example: you did not lie on your shoulder in the last 24 hours: put a check mark in the box for NA (not applicable).

NA YES NO
☑ ☐ ☐

My shoulder hurts when I lie on it.

- You did perform the activity in the last 24 hours, for example: you opened or closed a door in the last 24 hours. If your shoulder was painful during opening or closing a door; put a check mark in the box for YES.

NA YES NO
☐ ☑ ☐

My shoulder hurts when I open or close a door.

- You did perform the activity in the last 24 hours, for example: you did lean on your elbow or hand in the last 24 hours. If your shoulder did not hurt during leaning on your elbow or hand; put a check mark in the box for NO.

NA YES NO
☐ ☐ ☑

My shoulder hurts when I lean on my elbow or hand.

SDQ items

1 I wake up at night because of shoulder pain.
2 My shoulder hurts when I lie on it.
3 Because of pain in my shoulder it is difficult to put on a coat or a sweater.
4 My shoulder hurts during my usual daily activities.
5 My shoulder hurts when I lean on my elbow or hand.
6 My shoulder hurts when I move my arm.
7 My shoulder hurts when I write or type.
8 My shoulder is painful when I hold the driving wheel of my car or handle bars of my bike.
9 When I lift and carry something my shoulder hurts.
10 During reaching and grasping above shoulder level my shoulder hurts.
11 My shoulder is painful when I open or close a door
12 My shoulder is painful when I bring my hand to the back of my head.
13 My shoulder is painful when I bring my hand to my buttock.
14 My shoulder is painful when I bring my hand to my low back.
15 I rub my painful shoulder more than once during the day.
16 Because of my shoulder pain I am more irritable and bad tempered with people than usual.