Table W1: content and construct validity of the shoulder disability questionnaires

questionnaire		*content validity* item selection*	item reduction*	level of reading examined*	dimensionality studied?	internal consistency	*construct validity* hypothesis	(main) results	floor / ceiling effect	study size
SDQ-UK	w1	patients experts investigator	no	no	?	?	yes	score GP-patients > score community ; restricted ROM -> higher disability	ceiling†	54; 67
SIQ	w2	patients investigator	yes	yes	?	α = 0.91	yes	Constant: r = -0.56 Rowe: r = -0.51 SF36 physical: r = -0.71	no	92
OSQ	w3	patients investigator	yes	yes	?	α = 0.89	yes	Constant: r = -0.74 SF36 physical: r = -0.61 HAQ disability: r = 0.86	no	111
SDQ-NL	w4	experts investigator	yes	yes	?	?	?	?	ceiling‡	180
RC-QOL	w5	patients experts investigator	yes	yes	?	?	yes	SF36: r = 0.78; ASES: r = 0.84	no	70
DASH	w6	patients	yes	yes	?	?				
	w7	experts					yes	SPADI function: r = 0.85	no	138
	w8	investigator					yes	SF36 physical: r=-0.73	no	23
	w9						no	Constant: r = -0.76	?	23
WOSI	w10	patients experts investigator	yes	yes	?	?	yes	DASH: r = 0.77; Constant: r = 0.59; Rowe: r = 0.61; ASES: r = 0.55; SF12 physical: r = 0.66	?	47
SSRS	w11	investigator	no	no	?	?	no	Constant: r = 0.83	no	200
	w12						yes	SF36 physical: r = 0.12; SST: r = 0.47; SPADI: r = 0.50; m-ASES: r = 0.50; SSI: r = 0.48	no	90
SRQ	w13	patients investigator	yes	yes	?	α = 0.77-0.90¶	yes	AIMS: r = -0.84	?	97
SST	w14	patients	no	no						
	w12	investigator					yes	SF36 physical: r = 0.58; SSRS: r = 0.47; SPADI: r = 0.74; m-ASES: r = 0.73; SSI: r = 0.80	no	90
	w9						no	Constant: r = 0.49	?	23
	w15				yes	α = 0.85	no	SPADI: r = -0.80	?	192

questionnaire		*content validity* item selection*	item reduction*	level of reading examined*	dimensionality studied?	internal consistency	*construct validity* hypothesis	(main) results	floor / ceiling effect	study size
WOOS	w16	patients experts investigator	yes	yes	?	?	yes	Constant: r = 0.73; ASES: r = 0.59; SF12 physical: r = 0.65	?	41
SSI	w12	?	?	?	?	?	yes	SF36 physical: r = 0.59; SSRS: r = 0.48; SST: r = 0.80; SPADI: r = 0.79; m-ASES: r = 0.79	no	90
UEFS	w17	experts	yes	no	yes	α = 0.83-0.93**	no	significant difference between levels of severity	no	
ASES	w18	experts	yes	no	?					
	w9	investigator					no	Constant: r = 0.87	?	23
	w12						yes	SF36 physical: r = 0.60; SST: r = 0.73; SSRS: r = 0.50; SPADI: r =0.77; SSI: r = 0.79	no	90
	w10						yes	WOSI: r = 0.55	?	
	w16						yes	WOOS: r = 0.59	?	
	w19						no	Rowe: r = 0.82; UCLA: r = 0.50	?	
	w20					α = 0.90				
SPADI	w21	experts	yes	no	yes	α = 0.93††	yes	ROM: r = -0.54 - -0.80	?	37
	w22						no	HAQ: r = 0.61; SF-20 physical: r = -0.50	no	102
	w23						yes	SIP: r = 0.21 - 0.57	no	94
	w15				yes	α = 0.95††	no	UCLA function: r = -0.64	?	192
	w12						yes	SF36 physical: r = 0.58; SST: r = 0.74; SSRS: r = 0.50; m-ASES: r = 0.77; SSI : r = 0.79	no	90
	w7						yes	DASH: r = 0.85	?	138
	w20					α = 0.94††				
UEFL	w24	investigator	no	no	yes	?	yes	prevalence of self-reported difficulty at each level of functional limitation	floor	1002

* results based on first reference; † ceiling effect in community sample with shoulder disorders; ‡ ceiling effect in people who got physiotherapy treatment for soft tissue disorders; floor effect in healthy community dwelling woman and moderately to sever disabled woman (age > 65 years); ¶ subscales daily activities, recreational and athletic activities, and work; ** range across study groups; †† value of subscale disability; α = chronbach's alpha; ? = no data published; r = correlation coefficient; AIMS = Arthritis Impact Measurement Scales Health Status Questionnaire [w25]; Constant = Constant Score [w26]; Rowe Rating Scale [w27]; SF-36 = Medical Outcome Study Short-Form 36 [w28]; HAQ = Health Assessment Questionnaire [w29]; SIP = Sickness Impact Profile [w30]; UCLA = University California - Los Angeles Shoulder Scale [w31;w32]

Table W2: reproducibility of the shoulder disability questionnaires

questionnaire		reliability (I)	time interval	agreement (II)	study size*
SDQ-UK	w1	?	?	?	?
SIQ	w2	r = 0.97	24 hours	CoR = 5.7	34 (I, II)
OSQ	w3	?	24 hours	CoR = 6.8; MD = -0.12 (out of score 1-5)	60 (II)
SDQ-NL	w4	?	?	?	?
RC-QOL	w5	?	2 weeks	MD = 5.05 (out of score 0-100)	30 (II)
DASH	w6	ICC = 0.96	3-5 days	SEM = 4.6 (score 0-100)	73 (I); 56 (II)
WOSI	w10	ICC = 0.91†	2 weeks	?	51 (I)
SSRS	w33	ICC = 0.71	1 week	% = 63 [71]	41 (I, II)
SRQ	w13	r = 0.89-0.96‡	± 3 days	Kappa = 0.73 - 0.97	40 (I)
SST	w33	ICC = 0.99	1 week	% = 80 [95]	41 (I, II)
	w15			SEM = 11.65 (score 0-100)	192 (II)
WOOS	w16	ICC = 0.94†	3 months	?	22 (I)
SSI	w33	ICC = 0.97	1 week	% = 24 [NA]	41 (I, II)
UEFS	w17	?	?	?	
ASES	w33	ICC = 0.96	1 week	% = 31 [51]	41 (I, II)
	w20	ICC = 0.78 (0.59-0.89) post-surgical ; ICC = 0.86 (0.72-0.94) non-surgical	1 week		31(I) 25 (I)
SPADI	w21	ICC = 0.64¶	24 hours		23 (I)
	w15			SEM = 5.78¶	192 (II)
	w33	ICC = 0.91	1 week	% = 5 [23]	41 (I, II)
	w20	ICC = 0.57 (0.27-0.77) post-surgical ; ICC = 0.84 (0.66-0.92) non-surgical	1 week		31 (I) 25 (I)
UEFL	w24	?	?	?	?

* study size for study of study of reliability (I) and study of agreement (II); † subscale "sport, recreation and work"; ‡ subscales "daily activities", "recreational and athletic activities", and "work"; value of subscale function/disability with confidence intervals (in brackets) for post- and non-surgical patients"; ¶ value of subscale disability; ? = no data published; r = correlation coefficient; CoR = coefficient of reliability [w34]; MD = mean difference; SEM = standard error of measurement; Kappa = the proportion of the observed agreement that exceeds the agreement that is expected by chance alone; % perfect agreement: percent of subjects having identical scores; % perfect agreement within 1 response category (in brackets); NA = not applicable

Table W3: responsiveness and interpretability of the shoulder disability questionnaires

questionnaire		*responsiveness* treatment	time to follow-up	Hypothesis	(main) results	study size	*interpretability* attention for interpretability	baseline and follow up scores	scores of relevant subgroups	MCID
SDQ-UK	w1	?		?	?	?	no	?	?	no
SIQ	w2	physiotherapy / surgery	6 months	no	ES = 0.8 sign. difference between improved - not-improved	64	no	36.6; 95%CI: 34.4-38.8 (baseline); 28.3; 95%CI: 35.6-31.1 (follow-up)*	comparison of change scores with regard to the patients assessment of change	no
OSQ	w3	surgery	6 months	no	ES = 1.2 sign. difference between improved - not-improved	56	no	36.3; 95%CI: 34.6-37.9 (baseline); 26.0; 95%CI: 23.0-28.9 (follow-up)*	comparison of change scores with regard to the patients assessment of change	no
SDQ-NL	w4	physiotherapy	6 weeks	no	CRR = 1.14 ROC curve; AUC = 0.72	180	no	74 (63, 85) (stable); 70 (58, 78) (improved)†	median + percentiles of stable / improved patients for shoulder pain, chief complaint, symptoms and mobility	no
	w35	general practice	1 and 6 months	no	MCS: 20; 35 CRR = 2.22; 1.89 ROC curve; AUC = 0.84; 0.88	308	yes	67 (±23) (baseline); 47 (±31) (1 month); 32 (±31) (6 months)‡	mean change scores for clinical stable, improved and deteriorated patients	3 items
RC-QOL	w5	surgery	42 months (range 25-71)	?	?	?	no	69.9 (4.4-100) (follow up)	scores of large and massive rotator cuff tears	no
DASH	w7	surgery	3 months	yes	MCS: -13.4 (SD 16.6) SRM = 0.81 ES = 0.64 functional status: r =0.69 ROC curve	138	no	48.8 (±21.0) (baseline); 35.3 (± 21.3) (follow-up)‡	mean + SD wrist/hand patients; transition scale; comparison of change scores with regard to the patients assessment of change	no
	w6	?	3 months	yes	SRM =0.71 WOSI: r = 0.76	47	no	?	?	NA
	w9	surgery	57.8 weeks (±15.7)‡	no		23	no	49.6 (±8.5) (pre-operative); 21.6 (±13.0) (post-operatiave)	?	no
WOSI	w10	non specified treatment OA	3 months	yes	SRM = 0.93 DASH: r = 0.76 Constant: r =0.69 ASES: r = 0.50	47	no	?	?	no
SSRS	w11	surgery	3 and 12 months	no		?	yes	47(pre), 83 (post)¶; SA** 72 (pre),95 (post)¶; Bankart** 42 (pre), 52 (post)¶; MUA**	median+range diagnostic groups; comparison of change scores with regard to the patients assessment of treatment results	no
	w33	surgery	6 months	yes	MCS: 16.4 SRM = 0.65	33	no	52.2 (baseline); 69.1 (follow up)	?	no
SRQ	w13	surgery	12 months	no	SRM = 1.9 (1.1 - 1.8)* MCS: 26.7 (1.7 - 4.9)* IoR = 1.6 (1.1 - 2.0)*	30	yes	61.6 (±13.4) (pre-operative); 88.3 (±10.0) (post-operative)‡	overall score, scale-scores; initial score; score at one year follow up	2 points / domain

questionnaire		*responsiveness* treatment	time to follow-up	hypothesis	(main) results	study size	*Interpretability* attention for interpretability	baseline and follow up scores	scores of relevant subgroups	MCID
SST	w14			no	% progress per item	9 - 29	yes	% item score	% score diagnostic groups	no
	w33	surgery	6 months	yes	MCS: 17.2 SRM = 0.87	33	no	36.0 (baseline); 53.8 (follow up)	?	no
	w15	surgery	57.8 weeks (±15.7)‡	no		23	no	3.30 (±1.82) (pre-operative), 6.97 (±1.80) (post-operative)	?	no
WOOS	w16	surgery	3 months	yes	SRM = 1.91 Constant: r = 0.69 ASES: r = 0.43	41	no	?	?	no
SSI	w33	surgery	6 months	yes	MCS: 20.1 SRM = 1.05	33	no	47.0 (baseline); 67.3 (follow up)	?	no
UEFS	w17	?	19 months (12-24)	no	SRM = -1.33 average pain: r = 0.58	16	no	43.3 (3.3-75.9) (baseline); 31.5 (0.0-62.0) (follow up)	working status; duration symptoms	no
ASES	w9	surgery	57.8 weeks (±15.7)‡	no		23	no	33.9 (± 15.9) (pre-operative); 71.9 (± 16.8) (post-operative)	?	no
	w33	surgery	6 months	yes	MCS: 17.6 SRM = 0.93	33	no	49.4 (baseline); 68.0 (follow up)	?	no
	w10	non-defined treatment OA	3 months	yes	SRM = 0.54 WOSI: r = 0.50	47	no	?	?	no
	w16	surgery	6 months	yes	SRM = 1.29 WOOS: r = 0.43	41	no	?	?	no
	w20	?	?	?	?	31; 25	no	65.7 (± 22.7) (post-surgical); 66.4 (± 22.9) (non-surgical)‡	?	
SPADI	w21	medication or injection	30 days	no	ROM: r = -0.52 - -0.70 MCS: -25.3††	30	no	?	?	no
	w22	?	2, 4 and 12 weeks	no	overall status: r = 0.73; r = 0.76; r = 0.79 ROC curve; AUC = 0.91	75	yes	57.6 (22.5) (baseline)‡; -21.9-6.5 (change score)	change score 2-4-12 weeks / overall status (improved, same, worse).	>10 points
	w23	physiotherapy	±10 weeks	no	MCS: -28.4†† SRM = 1.04)††	34	no	33.9 (±28.1) (baseline); -28.4 (±27.2) (change score)‡	consensus between therapist and patient judgement on meaningful improvement in shoulder function	no
	w33	surgery	6 months	yes	MCS: 25.6 SRM = 1.23	33	no	39.9 (baseline); 66.4 (follow up)	?	no
	w7	surgery	3 months	yes	SRM = 0.71††	138	no	?	?	no
	w20	?	?	?	?	31; 25	no	28.5 (±25.6) (post-surgical); 47.9 (±24.6) ((non-surgical)‡	?	
UEFL	w24	?		?	?	?	no	?	?	no

* mean and 95% confidence interval; † mean and 25th and 75th percentiles;‡ mean and SD; mean and range; ¶ median score pre-operatiave (pre) and post-operative (post); ** SA = subacromial decompression; Bankart = Bankart repair of anterior shoulder reconstruction; MUA = manipulation under anesthesia; †† disability scale; MCID = minimal clinically important difference; IoR = Index of Responsiveness [w36]; CRR = calibrated responsiveness ratio; MCS = mean change score; SRM = standardized response mean; ES = effect size

Reference List

w1. 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;**53**:525-8.

w2. Dawson J, Fitzpatrick R, Carr A. The assessment of shoulder instability. The development and validation of a questionnaire. *J Bone Joint Surg Br* 1999;**81**:420-6.

w3. Dawson J, Fitzpatrick R, Carr A. Questionnaire on the perceptions of patients about shoulder surgery. *J Bone Joint Surg Br* 1996;**78**:593-600.

w4. Heijden GJvd, Leffers P, Bouter LM. Shoulder disability questionnaire design and responsiveness of a functional status measure. *J Clin Epidemiol* 2000;**53**:29-38.

w5. Hollinshead RM, Mohtadi NG, Vande Guchte RA, Wadey VM. Two 6-year follow-up studies of large and massive rotator cuff tears: comparison of outcome measures. *J Shoulder Elbow Surg* 2000;**9**:373-81.

w6. Hudak PL, Amadio PC, Bombardier C. Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand). The Upper Extremity Collaborative Group (UECG). *Am J Ind Med* 1996;**29**:602-8.

w7. Beaton DE, Katz JN, Fossel AH, Wright JG, Tarasuk V, Bombardier C. Measuring the whole or the parts? Validity, reliability, and responsiveness of the Disabilities of the Arm, Shoulder and Hand outcome measure in different regions of the upper extremity. *J Hand Ther* 2001 ;**14**:128-46.

w8. SooHoo NF, McDonald AP, Seiler JG, McGillivary GR. Evaluation of the construct validity of the DASH questionnaire by correlation to the SF-36. *J Hand Surg [Am]* 2002;**27**:537-41.

w9. Skutek M, Fremerey RW, Zeichen J, Bosch U. Outcome analysis following open rotator cuff repair. Early effectiveness validated using four different shoulder assessment scales. *Arch Orthop Trauma Surg* 2000;**120**:432-6.

w10. Kirkley A, Griffin S, McLintock H, Ng L. The development and evaluation of a disease-specific quality of life measurement tool for shoulder instability. The Western Ontario Shoulder Instability Index (WOSI). *Am J Sports Med* 1998;**26**:764-72.

w11. Kohn D,.Geyer M. The subjective shoulder rating system. *Arch Orthop Trauma Surg* 1997;**116**:324-8.

w12. Beaton DE,.Richards RR. Measuring function of the shoulder. A cross-sectional comparison of five questionnaires. *J Bone Joint Surg Am* 1996;**78**:882-90.

w13. L'Insalata JC, Warren RF, Cohen SB, Altchek DW, Peterson MG. A self-administered questionnaire for assessment of symptoms and function of the shoulder. *J Bone Joint Surg Am* 1997;**79**:738-48.

w14. Lippit SB, Harryman DTI, Matsen FAI. A practical tool for evaluation of function: the simple shoulder test. *The Shoulder: a Balance of Mobility and Stability.Matsen FA III, FU FH, Hawkins RJ (ed).Rosemont, Illinios, The American Academy of Orthopaedic Surgeons* 1993;501-18.

w15. Roddey TS, Olson SL, Cook KF, Gartsman GM, Hanten W. Comparison of the University of California-Los Angeles Shoulder Scale and the Simple Shoulder Test with the shoulder pain and disability index: single-administration reliability and validity. *Phys Ther* 2000;**80**:759-68.

w16. Lo IK, Griffin S, Kirkley A. The development of a disease-specific quality of life measurement tool for osteoarthritis of the shoulder: The Western Ontario Osteoarthritis of the Shoulder (WOOS) index. *Osteoarthritis Cartilage* 2001;**9**:771-8.

w17. Pransky G, Feuerstein M, Himmelstein J, Katz JN, Vickers-Lahti M. Measuring functional outcomes in work-related upper extremity disorders. Development and validation of the Upper Extremity Function Scale. *J Occup Environ Med* 1997;**39**:1195-202.

w18. Richards RR, An K-N, Bigliani LU, Friedman RJ, Gartsman GM, Gristina AG *et al.* A standardized method for the assessment of shoulder function. *Journal of Shoulder and Elbow Surgery* 1994;**3**:347-52.

w19. Romeo AA, Bach BR, O'Halloran KL. Scoring systems for shoulder conditions. *Am J Sports Med* 1996;**24**:472-6.

w20. Cook KF, Roddey TS, Olson SL, Gartsman GM, Valenzuela FF, Hanten WP. Reliability by surgical status of self-reported outcomes in patients who have shoulder pathologies. *J Orthop Sports Phys Ther* 2002;**32**:336-46.

w21. Roach KE, Budiman-Mak E, Songsiridej N, Lertratanakul Y. Development of a shoulder pain and disability index. *Arthritis Care Res* 1991;**4**:143-9.

w22. Williams JW, Holleman DR, Simel DL. Measuring shoulder function with the Shoulder Pain and Disability Index. *J Rheumatol* 1995;**22**:727-32.

w23. Heald SL, Riddle DL, Lamb RL. The shoulder pain and disability index: the construct validity and responsiveness of a region-specific disability measure. *Phys Ther* 1997;**77**:1079-89.

w24. Simonsick EM, Kasper JD, Guralnik JM, Bandeen-Roche K, Ferrucci L, Hirsch R *et al*. Severity of upper and lower extremity functional limitation: scale development and validation with self-report and performance-based measures wof physical function. *J Gerontol B Psychol Sci Soc Sci* 2001;**56**:10-9.

w25. Meenan RF, Mason JH, Anderson JJ, Guccione AA, Kazis LE. AIMS2. The content and properties of a revised and expanded Arthritis Impact Measurement Scales Health Status Questionnaire. *Arthritis Rheum* 1992;**35**:1-10.

w26. Constant CR,.Murley AH. A clinical method of functional assessment of the shoulder. *Clin Orthop* 1987;160-4.

w27. Rowe CR, Patel D, Southmayd WW. The Bankart procedure: a long-term end-result study. *J Bone Joint Surg Am* 1978;**60**:1-16.

w28. Ware JE,.Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;**30**:473-83.

w29. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980;**23**:137-45.

w30. Bergner M, Bobbitt RA, Carter WB, Gilson BS. The Sickness Impact Profile: development and final revision of a health status measure. *Med Care* 1981;**19**:787-805.

w31. Amstutz HC, Sew Hoy AL, Clarke IC. UCLA anatomic total shoulder arthroplasty. *Clin Orthop* 1981;7-20.

w32. Ellman H, Hanker G, Bayer M. Repair of the rotator cuff. End-result study of factors influencing reconstruction. *J wBone Joint Surg Am* 1986;**68**:1136-44.

w33. Beaton D,.Richards RR. Assessing the reliability and responsiveness of 5 shoulder questionnaires. *J Shoulder Elbow Surg* 1998;**7**:565-72.

w34. Bland JM,.Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;**1**:307-10.

w35. Windt DAvd, Heijden GJvd, Winter AFd, Koes BW, Deville W, Bouter LM. The responsiveness of the Shoulder Disability Questionnaire. *Ann Rheum Dis* 1998;**57**:82-7.

w36. Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis* 1987;**40**:171-8.