

ORIGINAL REPORT

HETEROGENEOUS ASSESSMENT OF SHOULDER DISORDERS: VALIDATION OF THE STANDARDIZED INDEX OF SHOULDER FUNCTION

Arnaud Dupeyron, MD, PhD^{1,5}, Anthony Gelis, MD^{1,5}, Philippe Sablayrolles, MD², Philippe-Jean Bousquet, MD, PhD³, Marc Julia, MD², Christian Herisson, MD^{2,5}, Jacques Pélissier, MD^{1,5} and Philippe Codine, MD⁴

From the ¹Département de Médecine Physique & Réadaptation, Centre Hospitalier Universitaire de Nîmes, Nîmes, ²Service Central de Rééducation, Centre Hospitalo-Universitaire Lapeyronie, Montpellier, ³BESPIIM, Centre Hospitalier Universitaire de Nîmes, Nîmes, ⁴Clinique La Pinede Centre Médical sur route de Peyrestortes, St Esteve and ⁵Movement to Health, Montpellier-1 University, EuroMov, Montpellier, France

Objective: Although 40 assessment tools are described in the literature, very few of them have been correctly validated. The Standardized Index of Shoulder Function (FI2S) encompasses pain, mobility, strength and function. The aim of this work is to describe the FI2S and to study its construct validity, reliability and responsiveness to change.

Patients: Fifty-nine patients with non-surgical (rotator cuff lesions, frozen shoulders, osteoarthritis) or post-surgical (acromioplasty, repairs of rotator cuff tears, arthroplasty) shoulder disorders were included.

Methods: The FI2S was compared with the Disabilities of the Arm, Shoulder and Hand questionnaire (DASH), with the Constant-Murley Score (CMS), and with a visual analogue scale for pain.

Results: Inter-test reliability and inter-rater reliability are excellent, with intra-class correlation coefficient of 0.93 (0.88–0.96) and 0.94 (0.90–0.96), respectively. Under a convergent hypothesis, the Spearman's correlation coefficients with the CMS and DASH score are 0.91 ($p < 0.0001$) and -0.64 ($p < 0.0001$), respectively. Correlations between the FI2S and the CMS are excellent for mobility and strength, but moderate for pain and functional capacities. Under a divergent hypothesis, no correlation is observed between the FI2S total score and age. Responsiveness to change is excellent.

Conclusion: The FI2S appears to be a proper assessment tool for pain, mobility, strength and function in shoulder disorders, easy to administer and of good metric value.

Key words: shoulder; assessment; heterogeneous score; validation.

J Rehabil Med 2010; 42: 967–972

Correspondence address: Arnaud Dupeyron, Département de Médecine Physique & Réhabilitation, CHU de Nîmes, Place du Pr Robert Debré, FR-30 029 Nîmes, cedex 09, France. E-mail: arnaud.dupeyron@chu-nimes.fr

Submitted January 3, 2010; accepted August 31, 2010

INTRODUCTION

In order to assess the efficacy of any treatment and compare it with another one, particularly in different study designs, we need accurate, reliable and widely used tools to assess pain, mo-

tor function and impact on physical and participatory activities. This is especially true for shoulder disorders (1). For shoulder pathologies, more than 40 assessment tools are available (2). Some of them assess shoulder function in specific lesions (instability, osteoarthritis) (3, 4). Others, such as the Disabilities of the Arm, Shoulder and Hand (DASH) questionnaire score, measure the general function of the upper arm (5), regardless of the original pathological cause. Out of 43 evaluation tools, only 9 have undergone a correct validation process for reliability and validity (2, 6), 4 of them are specific to shoulder pathologies (Western Ontario Shoulder Instability index (WOSI) or Shoulder Instability Questionnaire (SIQ) for instability, Western Ontario Osteoarthritis of the Shoulder (WOOS) for osteoarthritis, Oxford Shoulder Questionnaire (OSQ) for surgery) and 5 are non-specific (American Shoulder and Elbow Society scores (ASES), DASH, self-reported Flexilevel Scale of Shoulder Function (FLEX-SF), Shoulder Pain and Disability Index (SPADI), Simple Shoulder Test (SST)). A validated tool designed for all shoulder pathologies in general is lacking.

In order to completely describe “shoulder outcomes”, clinicians usually take into account self-reported pain (7), range of motion (ROM), strength and function. Some argue that physical impairments such as strength or mobility, are not closely related to function (8, 9), which requires specific functional tests. Other indexes, such as self-administered questionnaires, based either on functional abilities or quality of life, are limited by the lack of an objective evaluation of mobility and strength. In this context, the Standardized Index of Shoulder Function (FI2S) was designed to measure both objective and subjective data. In order to be widely and easily accessible, it needs acceptable psychometric properties, especially reproducibility, construct validity, and a good sensibility to responsiveness to change (10). Therefore, the objective of this study is: (i) to describe this new shoulder assessment tool; (ii) to test its reliability and responsiveness; and (iii) to compare it with other tools to partly assess concurrent validity.

Standardized Index of Shoulder Function description

In order to build the most pertinent assessment tool for shoulder disorders, a panel of clinical and surgical shoulder experts was selected based on their clinical expertise, critical literature and review of existing scales.

The following criteria were distributed to the panel of experts:

- to assess more than one dimension of perceived pain;
- to address tasks of functional relevance and measure anatomical restrictions at the same time;
- to be easy to use even for severe or moderate impairments;
- to give an idea of strength that can be compared between age and sex groups; and
- to divide the index into subgroups with global and ideally inherent clinical relevance.

In consequence, the FI2S items were selected and divided into 4 subgroups: pain, function, mobility and strength, as summarized in Appendix I. The total score is 100:

- *pain* is attributed 28 points with a qualitative and quantitative evaluation (1, 7, 11) at night, at rest and during activity and also with analgesic drug administration;
- *active mobility* is attributed 24 points. ROM is measured with a goniometer (°) and thumb-C7 distance is also measured (cm);
- *function* (or limitation of activities) is attributed 30 points. The selected items correspond to different kinds of activities, such as activities of daily living (dressing, catching an object, open a door, etc.) that can be easily performed during clinical examination;
- *strength* assessment is essential since it has been demonstrated that it is directly related to quality of life (12, 13) and is attributed 18 points.

According to the Constant method of strength assessment (14), the patient holds the handle of a spring balance in his or her hand, with the palm of the hand facing the floor, at arm's length with 90° of forward flexion in the sagittal plane. The patient resists the force applied by the examiner and is asked to maintain this position for 5 s. Three tests are performed and the average (kg) is noted. The strength value has to be adjusted for gender and age as it has been addressed in the CMS (15, 16). This adjustment has been determined using strength measured with an electronic Kinedyne-type dynamometer (Kinetec®, Tournes, France) in 86 control subjects with no upper arm disorders (Table I). Due to the lack of significant differences in subjects under the age of 50 years, and according to the linear decrease in isometric strength with age (17), only 3 age groups were retained: under 50 years, between 50 and

60 years, and over of 60 years; in order to reach the score of 18 attributed to strength, and according to these normal values depending on sex and age, adjustment coefficients were used to calculate the value of the subgroup strength from strength measured with the balance spring (see Appendix I). These coefficients are in accordance with those issued by the Copenhagen City Heart Study (17). The total score in the control sample was always 100, except for one person who achieved a total score of 93/100. Thus, the mean score for the 86 control subjects was 99.9.

Finally, the FI2S was in French, and translation was based on the guidelines of translation/back-translation in order to compile an English version in accordance with the initial French one (18).

Standardized Index of Shoulder Function comparison

To address the second objective of this study, the FI2S was compared with the Constant-Murley score (CMS) and the DASH questionnaire, which are widely used in shoulder disease. The CMS has the advantage of including pain, function, motion and strength for shoulder assessment (19). It has been demonstrated to have insufficient reliability during clinical follow-up (20), probably because shoulder pain is assessed by only a single visual analogue scale, function is evaluated by global discomfort in daily life activities, range of motion is partially assessed by functional tests and shoulder strength assessment requires the use of a weighting table. The DASH is a 30-item, validated, self-report questionnaire designed to measure physical function and symptoms in people with any of several musculoskeletal disorders of the upper limb. Whereas this questionnaire is not specific for shoulder disorders, its usefulness for clinicians who wish to monitor arm pain and function in individual patients has been widely demonstrated (21).

MATERIAL AND METHODS

The FI2S was validated in a prospective multicentre study. Patients with one of the following shoulder pathologies were recruited: rotator cuff lesions with or without rupture, frozen shoulder, or osteoarthritis. Patients who had surgery for acromioplasty, rotator cuff repair, or shoulder arthroplasty were also included. Other pathologies, such as shoulder pain caused by cancer, fracture, rheumatoid arthritis, septic arthritis, or shoulder instability, as well as acute painful shoulder caused by calcified tendinitis and shoulder pain due to neurological diseases were not considered. Patients who were unable to answer questions or complete the questionnaires, or who did not give their consent, were also excluded. The local ethics committee authorized this study and a signed consent form was obtained from all recruited patients.

Reliability was assessed by practitioners specialized in physical medicine and rehabilitation (PRM). The intra-rater reliability was tested by administering the FI2S index at D0 (day 0) and D1 (day 1) by the same examiner, assuming that results observed at D1 were not influenced by the examination performed at D0. The inter-rater reliability was evaluated by administering the FI2S twice to two examiners in random order, thus minimizing potential bias caused by the influence of the first examination on the second. The first administration was carried out at the time of inclusion and the second an hour later. The patients were asked not to report previous examinations and results to the examiners, and examiners were blinded to the results of other examinations. Reliability was tested by the intra-class correlation coefficient, the Wilcoxon paired test and Bland and Altman graphic analysis in order to observe potential fixed and proportional biases. A linear

Table I. Strength in the control group, according to age and gender (n=86)

Age (years)	n (F/M)	Strength (kg)			
		Female (n=47)		Male (n=39)	
		Right arm	Left arm	Right arm	Left arm
20–29	11/9	5.99	5.48	11.50	10.60
30–39	9/7	6.50	5.90	9.55	9.30
40–49	10/8	5.55	5.53	11.75	11.19
50–59	9/8	5.43	5.18	8.72	8.42
>60	8/7	3.48	3.26	7.85	7.15

F: female; M: male.

regression tested for a linear relation between the means of the two measurements and the difference between the two measurements.

Construct validity of the FI2S was assessed by correlating the overall score with scores on variables supposedly assessing similar dimensions or concepts (10). We hypothesized that the FI2S score would have: (i) strong to moderate associations with the CSM and DASH scores, both in general and for sub-domains; (ii) weaker associations with pain at rest, and pain during activities and age. Spearman's correlation coefficients were interpreted as excellent (≥ 0.90), good ($0.70 < 0.89$), moderate ($0.50 < 0.69$), fair ($0.30 < 0.49$), or little or no correlation (< 0.30).

Responsiveness was evaluated by calculating the effect size and the standardized response mean (SRM). Since this analysis was not the main target of the present study, the responsiveness was studied only on the first 25 patients included.

All analyses were performed under SAS v8.1 (SAS Institute, Cary, NC, USA). The alpha level was set at 5%.

RESULTS

Population

Fifty-nine shoulders corresponding to 59 patients (24 (41%) were men whose mean age and standard deviation (SD) at evaluation was 60.3 years (SD 10.6)) were evaluated (Table II). The majority of the patients were right-handed (55 (93%)). The shoulder disorder was located on the right side in 34 (58%) cases, and corresponded to the dominant side in 32 (54%) cases (31 right-handed and 1 left-handed). Symptom duration lasted a median of 24 months (8–60). For 27 (46%) of the patients, there was no surgery: 16 rotator cuff lesions including 8 with ruptures, 8 frozen shoulders, and 3 cases of osteoarthritis. Thirty-two patients underwent surgery: 9 for acromioplasty, 18 repairs of rotator cuff ruptures, and 5 for arthroplasty. The mean age at surgery was 59.9 (SD 11.7) years; the median evaluation for patients after the surgery was 1.5 months (0.9–1.8) after the surgery. Discharge occurred from 18 to 61 days after inclusion (mean 31 (SD 11)). Visual analogue scale (VAS) values for pain, as well as the FI2S, CMS and DASH scores at inclusion are summarized in Table II. At discharge, the mean CMS was 57.54 (SD 19.05) and the mean FI2S was 68.71 (SD 17.71).

Reproducibility

Intra-rater reliability. This was first evaluated by the intra-class correlation coefficient (ICC) (1, 3) calculation. ICC was

Table III. Overall and sub-score reproducibility (interclass correlation coefficient) and confidence intervals

Examination	First examiner	First and second examiner
	At day 0 and 1	At day 0
Pain	0.84 (0.74–0.90)	0.81 (0.71–0.89)
Mobility	0.92 (0.87–0.95)	0.87 (0.79–0.92)
Function	0.90 (0.85–0.94)	0.82 (0.72–0.89)
Strength	0.93 (0.86–0.96)	0.80 (0.65–0.89)
Overall	0.94 (0.90–0.96)	0.93 (0.88–0.96)

excellent, and over 90% (ICC = 0.94 (0.90–0.96)). The mean difference between the two measurements was 1.4 (SD –5.5), and not statistically different ($p = 0.06$). Results for mobility, function, and strength are excellent and good for pain (Table III). All but one point (2%) out of 59 were observed within the 95% confidence interval (CI) of Bland and Altman graphic analysis for the intra-rater comparison, thus indicating a very good reliability between the two examinations (Fig. 1A). Moreover, the graphic analysis did not report any fixed or proportional bias. Differences between measurements were stable, even with extreme values. The absence of bias was confirmed by a non-significant linear regression between the mean measurements and the difference between the two measurements (slope = 0.02, $p = 0.66$).

Inter-rater reliability. Similar findings were observed for the inter-rater comparison. The ICC was set at 0.93 (0.88–0.96) and results for each sub-score were good (Table III). The mean difference between the two measurements was 0.02 (SD 6.2), and was not statistically significant ($p = 0.87$). All but one point (2%) out of 59 were within the 95% CI of Bland and Altman graphic analysis (Fig. 1B). No fixed or proportional biases were graphically observed and differences were stable, including extreme values. The linear regression was not significant (slope = 0.04, $p = 0.42$).

Construct validity

Graphic analysis shows a linear correlation between the CMS and the FI2S; Spearman's correlation coefficients were 0.91 and 0.93 ($p < 0.0001$) at the first and second examinations, respectively.

Table II. Description of the studied population at inclusion (n = 59)

	Non-surgery			Surgery			
	Whole group	Osteoarthritis	Rotator cuff	Adhesive capsulitis	Arthroplasty	Acromioplasty	Rotator cuff
Total, n	59	3	16	8	5	9	18
Women, n	35	2	7	6	2	5	13
Men, n	24	1	9	2	3	4	5
Age, years, mean (SD)	60.3 (10.6)	62.3 (8.1)	60.8 (8.5)	57.4 (14.2)	64.2 (1)	58.6 (12.1)	59.4 (10.5)
Disease duration, months, mean (SD)	41.8 (48.9)	132 (95.2)	48.6 (51.9)	15.6 (18.2)	40.8 (13.7)	19.6 (17.8)	43.7 (48.2)
Pain at rest, mean (SD)	22.4 (24.3)	37 (32.1)	37.4 (28.7)	34 (22.4)	5 (6.2)	10.3 (15.8)	12.2 (15.3)
Pain during activity, mean (SD)	49.1 (24.7)	40.7 (16.2)	54.6 (28.8)	58.4 (27.6)	41.6 (19.6)	43.2 (18)	46.6 (25.2)
FI2S, mean (SD)	55.5 (15.9)	42 (8.8)	62.7 (17.5)	44.9 (12.7)	53.9 (12.4)	60.1 (16.2)	54.3 (14.7)
CMS, mean (SD)	43.2 (15.0)	27.8 (6.5)	52.2 (15)	31.1 (9.5)	44.2 (8.4)	46.2 (13.4)	41.3 (15.1)
DASH, mean (SD)	47.8 (20.7)	50.1 (17.9)	49.3 (25.5)	59.1 (13.1)	39 (14.2)	39.4 (19.2)	47.6 (20.8)

SD: standard deviation; DASH: Disabilities of the Arm, Shoulder and Hand questionnaire; CMS: Constant-Murley Score; FI2S: Standardized Index of Shoulder Function.

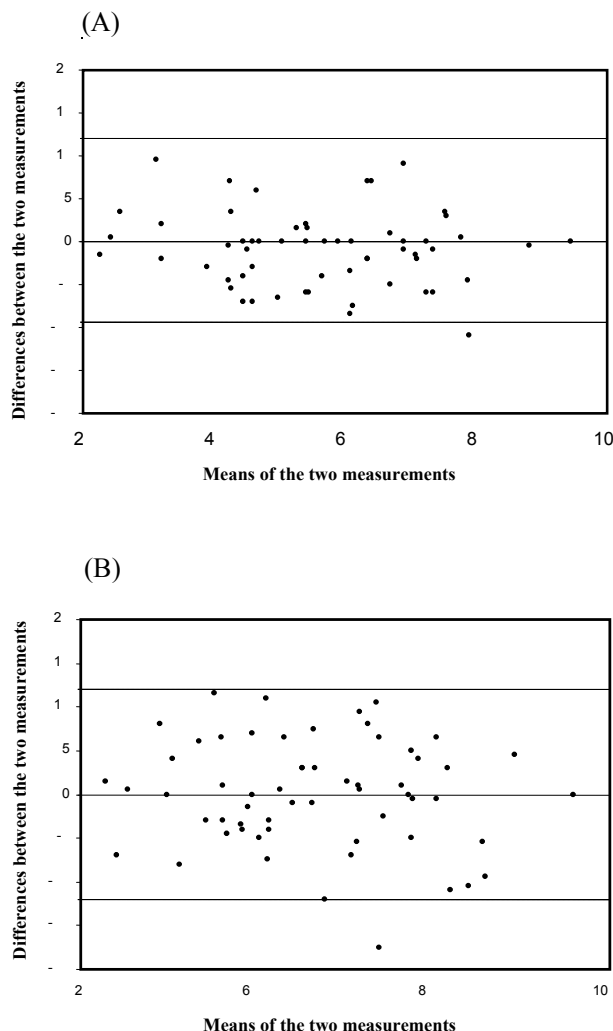


Fig. 1. (A) Intra-rater and (B) inter-rater reliability of the Standardization Index of Shoulder Function.

Table IV shows the correlation between the FI2S and the CMS, which is excellent for mobility and strength, but moderate for pain and function. For pain, the overall FI2S score is fairly correlated with the VAS pain score during activity ($\rho = -0.45, p < 0.001$ and $\rho = -0.42, p = 0.001$) and/or at rest ($\rho = -0.40, p = 0.002$

Table IV. Spearman correlation coefficients between Standardized Index of Shoulder Function and Constant-Murley Score components and overall estimation at different times

Examination	Day 0 First examiner	Day 0 Second examiner	Day 1 First examiner
Pain	0.55*	0.58*	0.65*
Mobility	0.90*	0.91*	0.89*
Function	0.66*	0.67*	0.74*
Strength	0.92*	0.90*	0.94*
Overall	0.91*	0.93*	0.92*

Day 0: first examination.

* $p < 0.001$.

and $\rho = -0.37, p = 0.004$, respectively). The DASH score is moderately correlated with the total FI2S score ($\rho = -0.53, p < 0.001$ and $\rho = -0.60, p < 0.001$ at the first and second examination on the first day, respectively) and the function subscores ($\rho = -0.50, p < 0.0001$ and $\rho = -0.64, p < 0.0001$). No correlation was observed between the FI2S total score and age ($\rho = 0.14, p = 0.28$ and $\rho = 0.10, p = 0.43$, respectively).

Responsiveness to change

Responsiveness to change was calculated by comparing FI2S at inclusion and at discharge (mean 8 weeks (SD 1.2)) in a subgroup of 25 patients. In this subgroup there were 12 women (48%), with a mean age of 58.6 (SD 9.5) years. The mean symptom duration was 49.3 (SD 53.0) months. Ten (40%) patients had surgery. A large effect size ES was observed (1.5, with a mean change of 20.5 and an initial standard deviation of 13.6). The SRM was 1.26 (with a 20.5 change (SD 16.2) in the FI2S total score), and could be considered a “wide” change in score.

DISCUSSION

Based on our study of 59 patients with important major shoulder disorders, the FI2S appears to be a relevant, reproducible assessment tool with a good responsiveness to change. The intra and inter-rater reliability were very good with both ICCs over 90% (0.94 (0.90–0.96) and 0.93 (0.88–0.96), respectively), and with nearly all the patients falling within the 95% CI of the Bland and Altman plot graphic analysis.

This study has some limitations. First, the FI2S was designed specifically to assess musculoskeletal disorders of the shoulder. The FI2S has not been tested for neurological shoulders, arthritis or infections not included in this study. Secondly, as concerns acceptability, it has been argued that the FI2S was easy to use both in clinical practice or clinical trials. However, the time needed to administer the FI2S has not been studied.

The CMS is widely used, but failed to demonstrate any metric properties and became a gold standard with use. Furthermore, the weight of some items or the measurement method has been criticized (22). These drawbacks limit its use in clinical research. Another example, the American Shoulder and Elbow Surgeons (ASES) scale, demonstrated good metric qualities (23), but whereas pain and function are well described, objective measurements, such as range of motion or strength, are lacking (24). Subjective indices, although easy to administer and focused on the patient, are in fact difficult to analyse; many studies have shown a frequent lack of correlation between the self-reported disability and functional performances (13, 25, 26).

The FI2S was issued following analysis by a group of experts with the standardization of 4 assessment scales commonly used in shoulder disorders, thus enhancing its content relevance. The differences and similarities based on the convergence and divergence hypothesis were examined. The FI2S had a good convergent validity with CMS, moderate with DASH and poor with pain. This can be explained by the content of

the pain subgroup in the FI2S, which takes into account pain intensity and its variation during the day, the consumption of analgesic drugs, while the VAS measures the intensity of pain at a certain given point. The correlation with the DASH is average, which is not surprising because the DASH focuses on pain and shoulder stiffness (beyond the shoulder on the whole upper arm mobility) in addition to function. The correlation between this functional subgroup of the CMS and the FI2S is also average, which can be related to different abilities of the shoulder at different times after lesions or surgeries. In the FI2S, strength is tested in a more comfortable position and the measure is less disturbed by pain. The importance of strength in the total score has been lowered in relation to its importance in the CMS, where a higher value can distort the total score when rest and function recovery are reached, corresponding to treatment objectives.

The FI2S has excellent psychometric properties: inter-test and inter-examiner reproducibility, correlation with the CMS, both qualities that make it a good tool to assess shoulder disorders and evaluate treatment results. For orthopaedic research, the good responsiveness of the FI2S will make it an essential tool in clinical trials for calculating sample size and power estimates. The equal importance given to the subscales make it more adapted to the practice of clinical orthopaedics and PRM than the CMS; more focus on pain and function, less on range of motion and strength makes it different from the CMS. Yet function (described through 5 activities) and pain (analysed through intensity, duration and analgesic consumption) are highly valued parameters in PRM and clinical orthopaedics, where the aim is not to achieve a range of motion or acquire more strength, but rather to improve the patient's autonomy and quality of life.

The FI2S was tested for main shoulder pathologies, in rehabilitation settings as well as post-surgical settings, but not for shoulder stability; it therefore cannot be used in that case. It will take its place alongside CMS and ASES, but seems more adapted to a rehabilitation context without losing its value in post-operative follow-up. The major interest of this heterogeneous scale is to obtain an overall value from objective and subjective data that can be used in comparative studies. Using some items independently can certainly be relevant for the follow-up of an isolated patient, but may lead to a loss of reproducibility, since the reproducibility of the total score does not necessarily correspond to that of each subgroup. Therefore, as suggested by Angst et al. (27), who compared metric properties and especially responsiveness to change of 6 shoulder evaluation scales, do we need to choose the most adapted tool for the objective of the assessment? For an overall, simple and relatively fast assessment, closely similar to CMS and ASES, FI2S with its own qualities fits this objective.

In conclusion, the FI2S is a heterogeneous index for assessing pain, mobility, strength and function, and gives greater importance to pain and function than the CMS; for these reasons it seems well tailored to PRM practice in shoulder evaluation. Compared with other tools, such as the CMS, DASH, and VAS at rest and during activity, it has a good constructed validity. Its inter-test and inter-rater reproducibility is also good, much

like its responsiveness to change. It appears to be an easy to administer, simple assessment tool with good metric value for shoulder disorders.

ACKNOWLEDGEMENTS

The authors would like to thank Bénédicte Clément and Dr Carey Suehs for their English revision of the manuscript.

REFERENCES

- Green S, Buchbinder R, Glazier R, Forbes A. Systematic review of randomized controlled trials of interventions for painful shoulder: selection criteria, outcome assessment, and efficacy. *BMJ* 1998; 316: 354–360.
- Fayad F, Mace Y, Lefevre-Colau MM. Shoulder disability questionnaires: a systematic review. *Ann Readapt Med Phys* 2005; 48: 298–306.
- 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: 771–778.
- Rowe CR, Patel D, Southmayd WW. The Bankart procedure: a long-term end-result study. *J Bone Joint Surg Am* 1978; 60: 1–16.
- 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: 602–608.
- Fayad F, Mace Y, Lefevre-Colau MM, Poiraudou S, Rannou F, Revel M. Measurement of shoulder disability in the athlete: a systematic review. *Ann Readapt Med Phys* 2004; 47: 389–395.
- Dinnes J, Loveman E, McIntyre L, Waugh N. The effectiveness of diagnostic tests for the assessment of shoulder pain due to soft tissue disorders: a systematic review. *Health Technol Assess* 2003; 7: 1–166.
- MacDermid JC, Ghobrial M, Quirion KB, St-Amour M, Tsui T, Humphreys D, et al. Validation of a new test that assesses functional performance of the upper extremity and neck (FIT-HaNSA) in patients with shoulder pathology. *BMC Musculoskelet Disord* 2007; 8: 42.
- MacDermid JC, Chesworth BM, Patterson S, Roth JH. Validity of pain and motion indicators recorded on a movement diagram of shoulder lateral rotation. *Aust J Physiother* 1999; 45: 269–277.
- Fermanian J. Validation of assessment scales in physical medicine and rehabilitation: how are psychometric properties determined? *Ann Readapt Med Phys* 2005; 48: 281–287.
- Kuijpers T, van der Windt DA, van der Heijden GJ, Bouter LM. Systematic review of prognostic cohort studies on shoulder disorders. *Pain* 2004; 109: 420–431.
- Bankes MJ, Crossman JE, Emery RJ. A standard method of shoulder strength measurement for the Constant score with a spring balance. *J Shoulder Elbow Surg* 1998; 7: 116–121.
- MacDermid JC, Ramos J, Drosdowech D, Faber K, Patterson S. The impact of rotator cuff pathology on isometric and isokinetic strength, function, and quality of life. *J Shoulder Elbow Surg* 2004; 13: 593–598.
- Moseley HF. *Shoulder lesions*. 3rd edn. Edinburgh: E&S Livingstone; 1969, p. 28–29.
- Katolik LI, Romeo AA, Cole BJ, Verma NN, Hayden JK, Bach BR. Normalization of the Constant score. *J Shoulder Elbow Surg* 2005; 14: 279–285.
- Yian EH, Ramappa AJ, Arneberg O, Gerber C. The Constant score in normal shoulders. *J Shoulder Elbow Surg* 2005; 14: 128–133.
- Danneskiold-Samsøe B, Bartels EM, Bülow PM, Lund H, Stockmarr A, Holm CC, et al. Isokinetic and isometric muscle strength

in a healthy population with special reference to age and gender. *Acta Physiologica* 2009; 197: 1–68.

18. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)* 2000; 25: 3186–3191.
19. Constant CR, Murley AH. A clinical method of functional assessment of the shoulder. *Clin Orthop Relat Res* 1987; 214: 160–164.
20. Conboy VB, Morris RW, Kiss J, Carr AJ. An evaluation of the Constant-Murley shoulder assessment. *J Bone Joint Surg Br* 1996; 78: 229–232.
21. 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: 128–146.
22. Bankes MJ, Emery RJ. An evaluation of the Constant-Murley shoulder assessment. *J Bone Joint Surg Br* 1997; 79: 696.
23. Kocher MS, Horan MP, Briggs KK, Richardson TR, O'Holleran J, Hawkins RJ. Reliability, validity, and responsiveness of the American Shoulder and Elbow Surgeons subjective shoulder scale in patients with shoulder instability, rotator cuff disease, and glenohumeral arthritis. *J Bone Joint Surg Am* 2005; 87: 2006–2011.
24. Sallay PI, Reed L. The measurement of normative American Shoulder and Elbow Surgeons scores. *J Shoulder Elbow Surg* 2003; 12: 622–627.
25. Bostrom C, Harms-Ringdahl K, Nordemar R. Relationships between measurements of impairment, disability, pain, and disease activity in rheumatoid arthritis patients with shoulder problems. *Scand J Rheumatol* 1995; 24: 352–359.
26. Geertzen JH, Dijkstra PU, van Sonderen EL, Groothoff JW, ten Duis HJ, Eisma WH. Relationship between impairments, disability and handicap in reflex sympathetic dystrophy patients: a long-term follow-up study. *Clin Rehabil* 1998; 12: 402–412.
27. Angst F, Goldhahn J, Drerup S, Aeschlimann A, Schwyzer HK, Simmen BR. Responsiveness of six outcome assessment instruments in total shoulder arthroplasty. *Arthritis Rheum* 2008; 59: 391–398.

APPENDIX I. Standardized Index of Shoulder Function (FI2S)

Standardized Index of Shoulder Function					
Pain	Analgesics drug consumption	Daily 0	Irregular 3.5	Never 7	Total pain /28
	Pain at rest	Unbearable 0	Bearable or Inconstant 3.5	None 7	
	Pain for usual activities	Unbearable 0	Bearable or Inconstant 3.5	Normal 7	
	Night-time pain	Sleep is very disturbed 0	Sleep is moderately disturbed 3.5	Normal 7	
Range of motion	Forward flexion	Abduction	External rotation (at abduction 90°)	Thumb/C7 Distance (cm)	Total ROM /24
	0–60 1	0–60 1	0–30 1	>60 0	
	61–80 2	61–80 2	31–45 2	41–60 2	
	81–100 3	81–100 3	46–60 4	21–40 4	
	101–120 4	101–120 4	61–90 6	<20 6	
	121–140 5	121–140 5			
141–160 6	141–160 6				
Function	Ask the patient to perform every activity: 6 points if performed without any compensation 3 points if possible with compensation or performed with difficulties 0 pt if impossible				Total function /30
	To comb hair back from the forehead To pull on or off a sweater To catch an object at eye-level To open a door To pull up pants or skirt				
Strength $F = kg \times p$	The patient must hold the handle of a spring balance in his hand, palm of the hand facing the floor, at arm's length in both 90° of forward flexion and 90° of abduction in the coronal plane. The patient is asked to maintain this elevation for 5 s, repeated 3 times, the average (kg) is noted and multiplied by the applicable adjustment coefficient below to obtain the total strength score.				Total strength /18
	<i>p</i>	<50 years old	50–60 years old	>60 years old	
	Men	2	2.5	3	
Women	3	4	5		
					Total /100