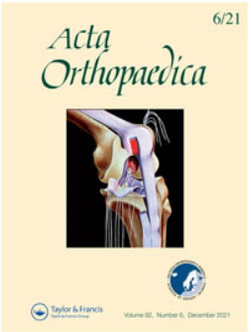

Acta Orthopaedica Scandinavica

ISSN: 0001-6470 (Print) (Online) Journal homepage: informahealthcare.com/journals/iort19

The Short Musculoskeletal Function Assessment Questionnaire (SMFA)Cross-cultural adaptation, validity, reliability and responsiveness of the Swedish SMFA (SMFA-Swe)

Sari Ponzer, Anne Skoog & Gunnar Bergström

To cite this article: Sari Ponzer, Anne Skoog & Gunnar Bergström (2003) The Short Musculoskeletal Function Assessment Questionnaire (SMFA)Cross-cultural adaptation, validity, reliability and responsiveness of the Swedish SMFA (SMFA-Swe) , Acta Orthopaedica Scandinavica, 74:6, 756-763, DOI: 10.1080/00016470310018324

To link to this article: https://doi.org/10.1080/00016470310018324

The Short Musculoskeletal Function Assessment Questionnaire (SMFA)

Cross-cultural adaptation, validity, reliability and responsiveness of the Swedish SMFA (SMFA-Swe)

Sari Ponzer[1], Anne Skoog[1] and Gunnar Bergström[2]

[1]Department of Orthopedics, Karolinska Institutet at Stockholm Söder Hospital, Stockholm, [2]Section for Personal Injury Prevention, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden
Correspondence: sari.ponzer@sos.sll.se
Submitted 03-01-29. Accepted 03-08-07

ABSTRACT The Short Musculoskeletal Function Assessment Questionnaire (SMFA) is designed to measure the functional status of patients with a broad range of musculoskeletal injuries and disorders. It has previously been validated for an American population. We have translated the SMFA into Swedish and tested the translated version (SMFA-Swe) as regards validity, reliability and responsiveness. Acute and elective cases (n = 298) were included in the study. The Swedish version of the SF-36 was used in the validation. We found that the SMFA-Swe was easy to use, that its reliability (internal consistency and stability) was good, that it correlated well with the SF-36 physical scores and that it was also sensitive to changes in musculoskeletal function over time.

Orthopedic studies frequently use disease-specific outcome measures because they are sensitive to the disorder studied. However, there may also be a need to evaluate function in groups of patients with different or multiple musculoskeletal disorders and, in such situations, the SMFA can be useful. We conclude that the SMFA-Swe is a valid instrument and can be used in clinical research as well as clinical practice when focusing on patients with various musculoskeletal disorders.

In clinical research, and clinical practice, one often needs a standardized assessment of the patient's current physical limitations due to an injury or a musculoskeletal disorder. Disease-specific, such as WOMAC (Bellamy et al. 1988), or region-specific methods, such as DASH (Hudak et al. 1996), are commonly used for this purpose. However, it can be difficult to use different instruments for each disorder or region and impossible when one wishes to evaluate and compare outcomes among patients with different or multiple injuries. Generic scales producing a "Health Profile"—e.g., the SF-36 (Ware and Sherbourne 1992, Sullivan et al. 1995), or an "Index"—e.g., EQ-5D (Dolan et al. 1995, Brooks and the EuroQol group 1996) are used in many cases for this purpose, but may be less sensitive for detecting small, yet important changes in musculoskeletal function.

The Short Musculoskeletal Function Assessment (SMFA) Questionnaire was developed by Swiontkowski et al. (1999) to study differences in the functional status of patients with a broad range of musculoskeletal disorders. It is based on the Musculoskeletal Function Assessment (MFA), a 101-item questionnaire, which is mainly used for research purposes due to its length (Engelberg et al. 1999). The American version of the SMFA has been shown to be a valid, reliable and responsive instrument for clinical assessments (Swiontkowski et al. 1999). It can be used to assess and compare all types of musculoskeletal diseases in the general population and is one of the outcome measures recommended by the American Academy of Orthopaedic Surgeons (AAOS).

The increasing interest in cross-cultural comparisons of—e.g., outcome assessment and treatment efficacy—has created a need for internationally

applicable standardized instruments. It has also been recognized that if a health status measure is to be used across cultures, the questionnaire must not only be well-translated linguistically, but also culturally adapted to maintain its content validity (Guillemin et al. 1993). This adaptation process is also expected to reduce the need for developing new instruments for the same purpose. We evaluated the Swedish version of the SMFA (SMFA-Swe) as regards validity, reliability and responsiveness.

Material and methods

Short Musculoskeletal FunctŠn Assessment QuestŠnnaire

The 46-item SMFA questionnaire comprises two parts: the dysfunction index with 34 items and the bother index with 12 items. The dysfunction index assesses the patients' perceptions of the amount of difficulty they have in the performance of certain functions (25 items) and how often the patients have difficulty when performing certain functions (9 items). The dysfunction items are grouped into four categories: daily activities, emotional status, function of the arm and hand, and mobility. Each item has a 5-point response format (1 point for good function and 5 points for poor function). The bother index asks the patients to assess how much they are bothered by problems in various areas of life (e.g., recreation, work, sleep and rest). These items also have a 5-point response format (1 point for not at all bothered and 5 points for extremely bothered).

The scores of the dysfunction and the bother indices are calculated by summing up the responses to the items and then transforming the scores according to the formula: (actual raw score - lowest possible raw score)/(possible range of raw score) × 100. This transformation formula gives the final scores, which ranged from 0 to 100. The higher scores indicate poorer function. In the case of the dysfunction index, unanswered items in a category are replaced by the individual's mean score for that category, so long as more than 50 per cent of the items in that category have been answered. Substitution with the mean is not appropriate for the bother index as each item addresses a unique area of function (Swiontkowski et al. 1999).

AdaptatŠn process

The American version of the SMFA was adapted for Swedish use according to the recommendations for cross-cultural adaptation of health status measures (Guillemin et al. 1993). Two independent translators with Swedish as their mother tongue (one aware of the concept) first translated the American version into Swedish. The two translations were combined into a synthesis and the differences resolved by consensus. Two independent translators with English as their mother tongue then translated this Swedish version of the SMFA back into English. Both were blinded to the concepts being investigated and had no medical background. A committee consisting of one of the translators, an outcome methodologist and two health care professionals reviewed all the translations and reached a consensus on all discrepancies. This pre-final version was tested on 30 orthopedic outpatients and on 30 healthy persons (health care professionals and their relatives and friends) and a few minor adjustments were made. The final version of the Swedish SMFA was then used to evaluate its validity, reliability and responsiveness.

Characteristics of the patients

298 patients (mean age 52 (16–94) years), having various orthopedic injuries and disorders were recruited for this study (Table 1). Those with an acute fracture or soft tissue injury were included during hospitalization, while those with chronic symptoms and/or undergoing elective surgery were contacted by letter and included at an outpatient visit. 51% of the patients were males, 83% were born in Sweden, 56% were married or cohabiting and 28% had had a university education. 40% of the patients were employed full-time, 29% had retired, 11% were on sick leave, 8% worked part-time, 3% were unemployed and data were missing in 9%.

In the test-retest analyses, 63 patients with a stable orthopedic condition were asked to complete the questionnaires twice (at the time of inclusion and after 1 month). Fifty-two patients also returned the second questionnaire in a median of 25 days: 14 with a previous hip replacement, 19 with an arthrodesis due to a rheumatoid hind foot, and 19

Table 1. Patients included by diagnostic groups. SMFA–Swe dysfunction and bother indices at baseline in all patients

Variable	n	Total score mean (SD)	Range (points)	Patients with 0 points, %	Cronbach's alpha
Dysfunction index (all)	295	29 (17)	0–78	1.4	0.94
Hand disorder [a]	7	18 (9)	4–31	0	
Hip replacement [c]	19	19 (16)	0–51	5.3	
Knee injury (meniscus)	46	22 (13)	6–60	0	
Humeral shaft fracture [c]	22	23 (20)	0–75	4.5	
Multi-trauma [d]	49	26 (17)	0–60	4.1	
Knee injury (ACL)	28	26 (14)	6–58	0	
Knee injury (other) [e]	29	28 (15)	2–56	0	
Ankle fracture [f]	29	35 (11)	16–65	0	
Distal radius fracture [g]	16	39 (16)	20–78	0	
Osteoarthritis (hip/knee) [h]	30	38 (18)	10–76	0	
Rheumatoid hind foot [i]	20	46 (18)	10–71	0	
Bother index (all)	297	30 (19)	0–81	3.7	0.90
Hand disorder [a]	7	19 (12)	2–35	0	
Hip replacement [b]	20	20 (21)	0–63	10	
Knee injury (meniscus)	46	24 (16)	0–63	0	
Humeral shaft fracture [c]	22	21 (21)	0–77	18	
Multi-trauma [d]	48	27 (21)	0–81	10.4	
Knee injury (ACL)	28	29 (15)	4–58	0	
Knee injury (other) [e]	30	31 (17)	2–67	0	
Ankle fracture [f]	29	39 (16)	17–75	0	
Distal radius fracture [g]	17	37 (18)	15–72	0	
Osteoarthritis (hip/knee) [h]	30	37 (18)	6–67	0	
Rheumatoid hind foot [i]	20	38 (17)	10–73	0	

[a] 2 ulnar collateral ligament injuries, 3 carpal tunnel syndromes, 2 others
[b] follow-up 3–10 years after a hip replacement
[c] follow-up 2 years after a humeral shaft fracture
[d] follow-up 1 year after multi-trauma (Injury Severity Score mean = 16.2, median = 14)
[e] 19 osteoarthrosis of the knee joint, 8 patella–related disorders, 3 others
[f] 11 lateral, 15 bimalleolar/ trimalleolar, 2 medial malleolar fractures, 6–8 weeks after injury
[g] 11 external fixation and 6 plaster fixation, 4–6 weeks after injury
[h] 16 osteoarthritis of the hip joint and 14 osteoarthritis of the knee joint
[i] follow-up 3–10 years after hind foot arthrodesis

with a healed humeral shaft fracture. We found no significant differences between the patients returning both questionnaires and all others as regards age or sex.

40 of the 46 patients with a condition that was expected to improve—i.e., 26 patients with an ankle fracture and 14 with a fracture of the distal radius—completed the questionnaire during their acute hospitalization and after about 2–4 months (median 87 days).

This study was done in accord with the ethical standards for human experimentation and the Helsinki Declaration of 1975, as revised in 1983. All patients gave their informed consent to participation after the local ethics committee had approved the study (244/99; 30 August 1999).

ReliabÕity

The internal consistency of a scale concerns the association of the included items with one another (Oppenheim 1992). Items devised to assess a single underlying continuum should yield consistent responses, and the scale items should correlate closely with one another. Cronbach's alpha coefficient was used to estimate the internal consistency (Ghiselli et al. 1981). Test-retest (stability), defined as the consistency of scores over time among respondents, assumed not to be changed, was assessed with the intraclass correlation coefficient (Dunn 1989). The intraclass coefficient is more useful than Pearson's r since it takes into account the actual magnitude of scores and the agreement between ratings, not only the correlation and linear association among variables.

Validity

Validity has been defined as the ability of a method to measure what it is intended to measure (Cohen et al. 1996). Three forms of validity were evaluated—i.e., face validity, content validity and construct validity. Face validity was judged by the patients. This type of validity refers to how relevant the test appears to be from the respondent's point of view. Content validity pertains to how well a test (e.g., items in an index) can be thought to measure the dimension it is intended to measure. The expert group evaluated the content validity.

Construct validity was assessed by testing predefined hypotheses concerning the expected relationship between the SMFA and 3 other measures—i.e., the SF-36, ratings by clinicians of mobility and everyday function and clinical measures of function by a physician or a research nurse. First, all patients rated their health-related quality of life according to the SF-36. The latter is a generic instrument comprising 36 statements concerning physical and mental dysfunction with 8 subscores: physical functioning, role limitations due to physical function, bodily pain, general health, vitality, social functioning, role limitations due to emotional problems and mental health. The raw scores are transformed and the final subscores for each category range from 0 to 100 (optimal health). It should be noted that a SMFA score of 100 indicates the worst possible condition. A mental score, a physical score and a total score for the SF-36 can also be calculated (Beaton et al. 1997, Bronfort and Bouter 1999). The SF-36 has been shown to have an acceptable internal consistency and construct validity among Swedish respondents (Sullivan et al. 1995). Second, a physician and/or research nurse rated the patient's function as regards mobility of the upper or lower extremities, the patient's ability to carry out activities of daily living and leisure time activities, as well as the patient's emotional coping. This simple rating scale ranges from 0 to 10 (10 being optimal) and has been used previously (Engelberg et al. 1999). Third, the same physician and/or research nurse measured the range of motion (ROM) of the injured/affected joint. In the analyses, the ROM was categorized as a 0–10 degree difference, a 11–30 degree difference or a more than 30 degrees difference between the affected and unaffected side.

We hypothesized that the SMFA-Swe scores would correlate negatively with the SF-36 scores and that the correlations would be stronger between the SF-36 physical scores than between the mental scores since the SMFA-Swe focuses primarily on perceived physical functioning. We also hypothesized that the SF-36 item "In general, would you say that your health is excellent/very good/good/poor?" would discriminate between low and high SMFA-Swe indices. Moreover, we expected that the physician ratings and measures of ROM would show a positive correlation with the SMFA scores.

Responsiveness

Responsiveness, which measures the ability of the questionnaire to detect clinical change, was evaluated by comparing the SMFA values at baseline and follow-up in patients with an expected clinical improvement—i.e., those with acute injuries would have higher/worse SMFA scores at baseline than at follow-up. Responsiveness was assessed with standardized response mean (SRM) statistics—i.e., the change observed was divided by the standard deviation of the observed change. The SRM is regarded as large (> 0.8), moderate (0.5–0.8), or small (< 0.5).

Other statistical methods

We used the statistical software SPSS 10.1 for Windows. Differences between group means were evaluated with the Student's t-test or ANOVA for independent groups and differences between distributions by the Chi-square test or Fisher's exact test. A paired samples t-test was used to compare scores at follow-up and baseline. Spearman's rho was used as a measure of the correlation between variables (difference from null). All tests were two-sided. The results were considered significant at $p < 0.05$.

Results

AdaptatŠn and face and content validity

Most SMFA items and the instructions in the questionnaire were translated without difficulty. Minor changes in 10 of the 46 items were made during the translation and before a consensus was reached.

Table 2. SMFA scores for the dysfunction and bother indices at baseline and at the time of the 2nd assessment in patients who filled in the SMFA-Swe twice (test-retest group)

Variable	n	Baseline		2nd assessment	
		dysfunction mean (SD)	bother mean (SD)	dysfunction mean (SD)	bother mean (SD)
Hip replacement	14	20 (17)	19 (20)	18 (18)	20 (22)
Humerus shaft fracture	19	25 (21)	23 (22)	26 (22)	23 (24)
Rheumatoid arthritis	19	45 (19)	37 (17)	46 (18)	39 (19)
Total	52	31 (22)	27 (21)	31 (23)	28 (23)

The pretest population (age 18–60 years, about 60% females) found it easy to fill the Swedish SMFA questionnaire. Some persons in the pretest population, who were not patients, commented on the layout, which was changed in its final form, and two of the items were then changed slightly.

As regards the validation of the Swedish SMFA, most patients stated that they had no difficulty in filling out the form, which took about 5–10 minutes. Item 15 ("How difficult is it for you to drive?") and item 22 ("How much difficulty are you having with sexual activity?") were not filled in as often as the other items. Since not everybody in Sweden has a driving license, this question was not applicable to all patients and resulted in a missing value in 9 of 298. Item 22 asks the patient about his/her difficulty as regards sexual activity as compared to his/her normal activity and some patients (9 of 298) preferred not to answer this question. However, the rate of missing values was low for these 2 items (3%), but it was even lower for all other items (1–1.5%). On the basis of the patients' opinions, the face validity can be considered good. The expert group validated the content validity and judged the questionnaire to be usable in an orthopedic population. The final version of the translation has been accepted by the American Association of Orthopaedic Surgeons as the official Swedish translation of the SMFA (i.e., SMFA-Swe). Mean values for the SMFA-Swe dysfunction and bother indices in various groups of patients are given in Table 1.

ReliabÕity

The analyses of internal consistency gave Cronbach alphas of 0.9 or higher for the dysfunction and bother indices (Table 1). These mean values of the indices for patients in whom no change over time was expected (follow-up of patients with hip replacements, humerus shaft fractures and rheumatoid arthritis) are shown in Table 2. The test-retest analyses yielded an intraclass correlation coefficient of 0.93 for the dysfunction index and 0.88 for the bother index.

Construct validity

The SMFA-Swe dysfunction index and the bother index at baseline correlated well with the total SF-36 score at baseline ($r = –0.62$ and –0.58, respectively, $p < 0.001$). The correlations were higher for the SF-36 physical health components ($r = –0.67$ and –0.61, respectively, $p < 0.001$) than for the SF-36 mental health components ($r = –0.48$ and –0.46, respectively, $p < 0.001$). In patients with acute injuries (ankle and distal radius fractures), the correlations between the total SF-36 and the SMFA-Swe dysfunction and bother indices were –0.78 and –0.76 at baseline ($p < 0.001$) and –0.91 and –0.80 ($p < 0.001$) at follow-up. The corresponding correlations in the elective cases were –0.50 for the dysfunction and bother indices on both occasions ($p < 0.001$). Patients who regarded their health as excellent/very good/good ($n = 216$) had a significantly better/lower SMFA-Swe dysfunction index (25 (SD 16) points) than those who regarded their health as fair or poor ($n = 79$, 40 (SD 16) points, $p < 0.001$). The same was true of the bother index where the former had a mean of 26 (SD 18) and the latter 40 (SD 17) ($p < 0.001$).

The relationship between the SMFA-Swe dysfunction and bother indices and the physician ratings (exact measures not shown) is given in Table 3. As can be seen, the SMFA-Swe indices correlated more closely with the physician's ratings as regards activities of daily living and emotional

Table 3. SMFA-Swe indices in relation to the physician's ratings at baseline

SMFA indices, physician's ratings	Spearman's Rho
Dysfunction index	
Physical function: mobility, lower extremities	–0.39
Physical function: mobility, upper extremities	–0.55
Activities of daily living	–0.55
Leisure time activities	–0.36
Emotional function	–0.52
Bother index	
Physical function: mobility, lower extremities	–0.33
Physical function: mobility, upper extremities	–0.41
Activities of daily living	–0.42
Leisure time activities	–0.34
Emotional function	–0.44

Table 4. Difference in range of motion (ROM) between the affected and unaffected side in relation to the SMFA-Swe indices (n = 224)

Difference in ROM [a]	Dysfunction index mean (SD)	Bother index mean (SD)
0–9 degrees	22 (14)	24 (17)
10–30 degrees	32 (16)	32 (18)
> 30 degrees	39 (15)	39 (16)

[a] $p < 0.001$ in the 3 groups

Table 5. SMFA-Swe scores for dysfunction and bother indices at baseline and at the time of the 2nd assessment for patients completing the SMFA-Swe twice (responsiveness group)

Variable	n	Baseline dysfunction mean (SD)	Baseline bother mean (SD)	2nd dysfunction mean (SD)	2nd bother mean (SD)	SRM [a] dysfunction	SRM [a] bother
Ankle fracture [b]	26	35 (11)	39 (6)	14 (10)	16 (15)	2.3	1.4
Distal radius fracture [b]	14	41 (17)	39 (18)	30 (18)	27 (18)	2.1	1.0
Total [b]	40	37 (14)	39 (16)	19 (15)	20 (17)	1.9	1.2

[a] Standardized response mean. Mean (test 1–test 2)/standard deviation of change observed
[b] $p < 0.001$, baseline compared to 2nd assessment.

function than the ratings of physical function and leisure time activities. Table 4 shows the relation between the SMFA-Swe and the ROM in patients with an injury/disorder affecting a joint (ankle, shoulder, wrist, hip, n = 224). Patients with a difference of less than 10 degrees compared to the unaffected side had significantly lower SMFA-Swe indices than those with a larger difference.

Responsiveness

We tested responsiveness by comparing the SMFA-Swe values at baseline and at follow-up in patients with a clinically expected improvement. As shown in Table 5, patients with fractures of the ankle and distal radius had SMFA-Swe scores significantly lower/better at follow-up than at baseline. Moreover, responsiveness was assessed by using the standardized response mean (SRM) (Table 5). The results showed SRM values between 1.0 and 2.3. The SRM is rated as large (> 0.8), moderate (0.5–0.8), or small (< 0.5).

Discussion

Cross-cultural adaptation of a self-administered health status questionnaire for use in a country other than that where it was developed is necessary to ensure its content validity (Guillemin et al. 1993; Beaton et al. 2000). However, to make sure that its psychometric qualities are retained, testing of validity, reliability and responsiveness should be done again after the translation (Ware and Gandek 1998). We did the translation and testing of the Swedish version of the SMFA (SMFA-Swe) in this way and our study showed that the SMFA-Swe is a valid, reliable and responsive method for clinical assessment of patients with musculoskeletal disorders. Our findings regarding the psychometric qualities were comparable to those reported by those who developed the questionnaire (Swiontkowski et al. 1999).

The face validity of the SMFA-Swe was considered good by the patients and the content validity

was regarded as satisfactory by the expert group. The cross-cultural adaptation, including translations, back translations and consensus discussions, was a time-consuming process, but we found this to be necessary. Most of the differences between the translations were minor, but important for the correct understanding of the questions. Answers to questions concerning driving a car and being able to have normal sexual activity were the ones missing most often, although both items can be replaced with the mean value for that category (Swiontkowski et al. 1999). Possessing a driving license and driving a car are probably commoner in the USA than in Sweden, but we chose to leave the questions unchanged. Not having sexual activity and therefore not being able to answer if it was affected by the injury is problematic, but not related to cross-cultural adaptation. However, the willingness to answer this question probably differs between different countries/cultures. The rate of missing values was generally low, which indicated that the questions were easy to understand.

Moreover, our results also showed that the SMFA-Swe had good internal consistency, meaning that these questions yielded consistent replies and the scale items correlated well with one another. The test-retest procedure was used to evaluate the stability of the questionnaire.

In our study, the second questionnaire was filled in after a median of 25 days, which is a longer period of time than in most studies (Deyo et al. 1991, Daltroy et al. 1996). Nonetheless, the results showed high and about the same intraclass correlations reported for the American version of the SMFA (Swiontkowski et al. 1999).

Validity—i.e., that the instrument measures what it is supposed to measure—is of vital importance when testing a questionnaire. The analyses yielded good results for the construct validity of the SMFA-Swe. First, the SMFA-Swe correlated, as hypothesized, with the total score of the SF-36. Second, as also hypothesized, the SMFA-Swe indices showed a stronger correlation with physical health than with the mental health components of the SF-36. We also, as expected, found a relationship between the ratings made by a health professional and the SMFA-Swe scores. As anticipated, these associations were weaker than the associations between the SMFA-Swe and SF-36 since a physician's ratings constitute a different method of gathering data than a patient's self-report on function (part of the associations between the SMFA and SF-36 can be attributed to a shared method variance). The SMFA-Swe dysfunction and bother indices showed a relationship with the measurements of the range of motion of the injured limb.

In the test of responsiveness, we chose patients with acute injuries as the test group to ensure that they would improve clinically during the period before completion of the second questionnaire and it was shown that the SMFA-Swe could detect this change. Therefore, the responsiveness of the SMFA-Swe was acceptable.

The original SMFA was evaluated in a population of 420 patients who had a musculoskeletal injury or disease. Our study population consisted of 298 patients with acute or chronic injuries or disorders, which can be regarded as a sufficiently large study sample. One shortcoming of our study is that we included only patients treated at a single hospital and no patients from primary care, unlike the study done by the developers of this questionnaire whose patients were recruited from several centers (Swiontkowski et al. 1999). However, we believe that since the aim of this study was to evaluate the questionnaire in a Swedish population and not primarily to obtain norms for various patient groups, our patient selection is acceptable.

In orthopedic studies, it is common to use disease-specific outcome measures, as these are sensitive to the disorder studied. However, there may also be a need to evaluate and compare patients with various or multiple musculoskeletal disorders and, in such situations, the SMFA can be useful. We conclude that the SMFA-Swe is a reliable and valid instrument and that it is also sensitive to changes in musculoskeletal function over time. We therefore recommend its use in clinical research and clinical practice, especially when focusing on patients with various musculoskeletal disorders.

This study was supported by research grants from the AFA Research Foundation.

The authors thank Marc F. Swiontkowski, MD, and Julie Agel, MA, for their invaluable advice and support during the translation and validation process and for devising the

SMFA, which served as an interesting subject for research. We also thank Irene Jensen, PhD for her scientific advice and registered nurses Agnes Ek, Maja Duvner and Tina Levander for their help with the patients.

Beaton D E, Hogg-Johnson S, Bombardier C. Evaluating changes in health status: reliability and responsiveness of five generic health status measures in workers with musculoskeletal disorders. J Clin Epidemiol 1997; 50: 79-93.

Beaton D E, Bombardier C, Guillemin F, Ferraz M B. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine 2000; 15 (25): 3186-91.

Bellamy N, Buchanan W W, Goldsmith C H, Campbell J, Stitt L W. Validation study of WOMAC: a health status instrument for measuring clinically important patient-relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988; 15: 1833-40.

Bronfort G, Bouter L M. Responsiveness of general health status in chronic low back pain: a comparison of the COOP charts and the SF-36. Pain 1999; 83: 201-9.

Brooks R with the EuroQol Group. EuroQol: the current state of play. Health Policy 1996; 37: 53-72.

Cohen R J, Swerdlik M E, Phillips S M. Psychological testing and assessment. Chapter 6: Validity, Mayfield Publishing Company, Mountain View, California, 1996: 174-217.

Daltroy L H, Cats-Baril W L, Katz J N, Fossel A H, Liang M H. The North American Spine Society lumbar spine outcome assessment instrument: reliability and validity tests. Spine 1996; 21: 741-9.

Deyo R A, Diehr P, Patrick D L. Reproducibility and responsiveness of health status measures: statistics and strategies for evaluation. Control Clin Trials 1991; 12: 142S-58S.

Dolan P, Gudex C, Kind P, Williams A. A social tariff for EuroQol: results from a UK general population survey. Centre for Health Economics. York Health Economics Consortium. HHS Centre for Reviews & Disseminations: York 1995.

Dunn G. Design and analysis of reliability studies. Oxford University Press, New York 1989.

Engelberg R, Martin D P, Agel J, Swiontkowski M F. Musculoskeletal Function Assessment: reference values for patients and non-patient samples. J Orthop Research 1999; 17: 101-9.

Ghiselli E E, Campbell J P, Zedeck S. Measurement theory for the behavioral sciences. Chapter 9: Estimating reliability of measurement, W. H. Freeman and Company, San Francisco 1981: 230-64.

Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol 1993; 46: 1417-32.

Hudak P L, Amadio P C, 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.

Oppenheim A N. Questionnaire design, interviewing and attitude measurement, Pinter Publisher, London and New York 1992.

Sullivan M, Karlsson J, Ware J E, Jr. The Swedish SF-36 Health Survey: evaluation of data quality, scaling assumptions, reliability and construct validity across general populations in Sweden. Soc Sci Med 1995; 41:1349-58.

Swiontkowski M F, Engelberg R, Martin D P, Agel J. Short Musculoskeletal Function Assessment Questionnaire: validity, reliability and responsiveness. J Bone and Joint Surg (Am) 1999; 8: 1245-60.

Ware J E, Jr., Sherbourne C D. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992; 30: 473-83.

Ware J E Jr, Gandek B. Methods for testing data quality, scaling assumptions and reliability: the IQOLA Project Approach. International Quality of Life Assessment. J Clin Epidemiol 1998; 51: 945-52.