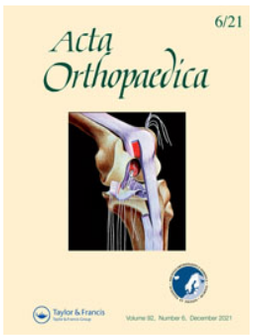

Acta Orthopaedica Scandinavica

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

Translation and validation of the Oxford-12 Item Knee Score for use in Sweden

Michael J Dunbar, Otto Robertsson, Leif Ryd & Lars Lidgren

To cite this article: Michael J Dunbar, Otto Robertsson, Leif Ryd & Lars Lidgren (2000) Translation and validation of the Oxford-12 Item Knee Score for use in Sweden, Acta Orthopaedica Scandinavica, 71:3, 268-274, DOI: 10.1080/000164700317411861

To link to this article: https://doi.org/10.1080/000164700317411861

Translation and validation of the Oxford-12 Item Knee Score for use in Sweden

Michael J Dunbar, Otto Robertsson, Leif Ryd and Lars Lidgren

Department of Orthopaedics, Lund University Hospital, SE-221 85 Lund, Sweden. Tel +46 46–17 15 10. Fax –17 71 67. Correspondence: Dr. Michael J. Dunbar, Suite 4882, Division of Orthopaedics, QE II Health Sciences Centre, New Halifax Infirmary Hospital Site, Halifax, Nova Scotia, B3H 3A7, Canada. Fax +1 902 443 9987
Submitted 99-05-28. Accepted 99-12-28

ABSTRACT – The Oxford-12 Item Knee Score is a recently developed and validated patient-completed outcome measure designed specifically for use with knee arthroplasty in the United Kingdom. We have translated this questionnaire into Swedish and tested the validity and reliability of the translated version in a cross-sectional study by a postal survey to 1200 randomly selected patients from the Swedish Knee Arthroplasty Register. Swedish versions of the WOMAC, Nottingham Health Profile, SF-36, SF-12, and the Sickness Impact Profile were employed in the validation process. We also tested feasibility and patient-burden parameters.

The translated version appeared to be linguistically and culturally equivalent to the original version with good validity and reliability. Indirect measures of responsiveness indicated that it is at least as responsive to relevant knee arthroplasty patient states as the previously validated Swedish version of the WOMAC. Application of the translated questionnaire to this population is feasible with minimal imposed patient-burden.

The Swedish translation of the Oxford-12 Item Knee Score is a valid and reliable tool for outcome studies on knee arthroplasty patients.

■

Outcome questionnaires for use with knee arthroplasty have much value (Kantz et al. 1992, Bombardier et al. 1995, Ritter et al. 1995, Rissanen et al. 1996). To date, the WOMAC is the only disease/site-specific questionnaire applicable to knee arthroplasty that has been translated and validated for use in Sweden (Roos et al. 1998). However, the WOMAC was not designed to measure knee arthroplasty outcomes (Bellamy et al. 1988). Questionnaires specific for a disease process or intervention have less "noise" than non-specific questionnaires (Dawson et al. 1996).

The Oxford-12 Item Knee Score is a new and well-validated outcome questionnaire, designed for use with knee arthroplasty patients (Dawson et al. 1998). We previously found that the Swedish translated version of the Oxford-12 performed optimally across multiple parameters in a cross-sectional study (Dunbar et al. 1999).

It is insufficient solely to translate a questionnaire into a foreign language, without validating the translated version (Guillemin et al. 1993, Guyatt 1993). Therefore, the purpose of this study was to translate and validate the Oxford-12 for use in Sweden.

Methods

Translation

The translation processes followed general guidelines from the literature (Guillemin et al. 1993, Mathias et al. 1994). The Oxford-12 was independently translated into Swedish and back-translated by a professional translator and bilingual orthopedic surgeon. Adequacy of the translated versions was assessed and a final translated version was agreed on. A pilot study was conducted on 8 bilingual subjects who completed in random order the Swedish and English versions of the Oxford-12, separated by a 5-day interval, to assess the translation further.

Patients

We randomly selected 1,200 patients from the Swedish Knee Arthroplasty Register after mortality data were updated from the Swedish National Statistics Register. Patients with a diagnosis of primary osteoarthrosis, age ≥ 55 years at time of surgery, age ≤ 95 years at the time of mail-out and prosthesis type of medial uni-compartmental, lateral uni-compartmental, bilateral (same knee) uni-compartmental and total knee arthroplasty were eligible. Patients who had undergone revision surgery were also eligible, provided they had not had an extraction arthroplasty, amputation or arthrodesis. The 1,200 patients represented a subset of 3,600 patients studied and reported on previously (Dunbar et al. 1999).

The 1,200 patients were divided into 4 groups of 300 each, receiving a combination of 1 of 4 general health questionnaires along with the Oxford-12. A cover letter was included with a post-age-paid return envelope and a third questionnaire asking how much time was required and whether patients needed help to complete the questionnaires. A reminder letter was sent at 2 weeks to non-responders. At 3 weeks, 120 patients were randomly selected from those who completed the Oxford-12 and were sent a WOMAC.

The average patient age at the time of mail-out was 78 (58–94) years and 71 (55–90) years at the time of index surgery. The average follow-up time was 6.7 (1.4–21) years. 70% (n 840) of the sample were female. 94% were primary arthroplasties. 59% of all patients had tri-compartmental replacements, 35% had medial uni-compartmental replacements, and 6.0% had either a lateral uni-compartmental or both compartments of the same knee replaced with a uni-compartmental prosthesis.

The Swedish Knee Arthroplasty Register has permission from the Swedish Health Authority (Socialstyrelsen) and the National Controlling Body for Computer Registers (Datainspektionen) to obtain and record patient factors related to knee arthroplasty.

Questionnaires

The Oxford-12 was tested against 5 other questionnaires in the validation process. All 5 questionnaires are available in a Standard Swedish version, which have been validated, with the exception of the SF-12 which is currently being validated. A brief description of each questionnaire follows. A domain refers to a series of items in a questionnaire that pose questions around a single concept. A summary score utilizes items from more than one domain.

Disease/site specific questionnaires

Oxford-12 Item Knee Score (Oxford-12) (Dawson et al. 1998). 12 questions are asked relating specifically to the knee. Each question has a Likert-box response key from 1 to 5. A single score is produced ranging from 12 to 60, with 12 indicating the best possible health state.

Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) (Bellamy et al. 1988, Roos et al. 1998). The WOMAC consists of 24 Likert-box questions divided into 3 domains: pain (5 questions), stiffness (2 questions) and physical function (17 questions). Scores range from 0–20 for pain, 0–8 for stiffness and 0–68 for physical function. A score of 0 represents the best possible health.

General health questionnaires

Nottingham Health Profile (NHP) (Hunt et al. 1980, 1981, Wiklund et al. 1988). The NHP asks 45 questions organized into two parts, to which a response of yes or no is given. In Part 1, 38 questions are used to generate weighted scores for 6 domains (Table 1), while in Part 2, 7 non-weighted questions are generated regarding perceived health problems affecting activities of daily life. Part 2 was not utilized in this study. Scores in part 1 range from 0–100, with 0 representing the best possible health.

12-Item Short-Form Health Survey (SF-12) (Ware et al. 1996). The SF-12 consists of 12 questions with a Likert-box response key. Scores are transformed into two weighted summary scores called Physical Component Summary and Mental Component Summary. The weights are calculated so that an average population sample will record a score of 50 for each summary and a score change of 10 points represents one standard deviation. A score above 50 represents a perception of better health than in that of the average population.

36-Item Short-Form Health Survey (SF-36) (Brazier et al. 1992, Ware and Sherbourne 1992,

Table 1. Spearman's correlation coefficients (non-parametric) for the Oxford-12-Item Knee Score against the domains and summary scores of 5 questionnaires (p < 0.001 for all correlations)

Questionnaire domains	Spearman Rho	95% CI lower limit	95% CI upper limit	n
Nottingham Health Profile				
Pain	0.82	0.77	0.86	211
Physical mobility	0.78	0.73	0.83	221
Energy	0.64	0.55	0.71	222
Emotional reaction	0.52	0.41	0.61	208
Sleep	0.46	0.35	0.56	219
Social isolation	0.33	0.21	0.45	218
SF-12				
Physical component summary [a]	–0.56	–0.65	–0.45	174
Mental component summary [a]	–0.50	–0.60	–0.38	174
SF-36				
Body pain	–0.64	–0.71	–0.56	224
Physical component summary [a]	–0.57	–0.67	–0.46	157
Physical functioning	–0.57	–0.65	–0.47	207
Vitality	–0.54	–0.63	–0.44	212
General health	–0.53	–0.62	–0.42	211
Social functioning	–0.51	–0.60	–0.40	211
Mental component summary [a]	–0.45	–0.57	–0.31	157
Role–physical	–0.42	–0.52	–0.30	219
Role-emotion	–0.41	–0.51	–0.29	217
Mental health	–0.40	–0.51	–0.28	210
Sickness impact profile				
Physical dimension [a]	0.55	0.45	0.64	196
Body care and movement	0.54	0.43	0.63	205
Ambulation	0.53	0.43	0.63	208
Total score [a]	0.52	0.38	0.64	121
Home management	0.52	0.41	0.61	208
Mobility	0.46	0.35	0.57	200
Sleep and rest	0.41	0.29	0.52	209
Psychosocial dimension [a]	0.35	0.22	0.47	203
Alertness behavior	0.34	0.21	0.46	207
Recreation and pastimes	0.30	0.17	0.42	201
Social interaction	0.29	0.16	0.41	208
Emotional behavior	0.29	0.16	0.41	207
Communication	0.26	0.13	0.39	206
Work	0.16	–0.02	0.31	135
Eating	0.14	0.00	0.27	205
WOMAC				
Physical function	0.87	0.80	0.92	78
Pain	0.83	0.76	0.89	92
Stiffness	0.74	0.63	0.82	91

[a] Summary scores

Sullivan et al. 1995). The SF-36 consists of 36 questions with Likert-box response keys. 8 domain scores are generated ranging from 0-100 (Table 1). A score of 100 represents the best possible health state. Two summary scales are also generated for the SF-36 (Physical and Mental Component Summary) and their scoring is similar to that for the summary scores of the SF-12.

Sickness Impact Profile (SIP) (Pollard et al. 1976, Sullivan 1985). The SIP is a 136-item questionnaire that asks patients to place a simple check mark beside each question if it applies. Otherwise, the question response key is left blank (Damiano 1996). The questionnaire produces weighted results for 12 domains and 3 summary scores (Table 1). Scores range from 0-100 with 0 representing the best possible health.

Patient-burden and feasibility

We investigated patient-burden by inquiring as to how much time was required to complete the questionnaire and whether the patient required assistance to do so (Lohr et al. 1996).

Feasibility was determined by calculating the percentage of questionnaires returned and the percentage of questionnaires that were returned fully completed. Missing responses were not imputed.

Psychometric properties

Convergent and divergent construct validity were tested by comparing the Spearman's correlation coefficients of the Oxford-12 scores to the domains of the four general health questionnaires and the WOMAC. We hypothesized that Oxford-12 should show the highest correlations with the physical and pain domains of the 5 other questionnaires (convergent validity) and the lowest with the eating domain of the SIP and the psychosocial domains of the general health questionnaires (divergent validity). Spearman's correlations were used because of the non-parametric nature of the data.

Content validity was studied by examining the skew of the distribution as well as floor and ceiling effects. Floor effect reflects the percentage of

patients recording the best possible score (12 of 60) while ceiling effect represents the percentage of patients recording the worst possible score (60 of 60).

To determine test-retest reliability, 60 patients were randomly selected from those who had completed the Oxford-12. Each was mailed a repeat Oxford-12 at 4 weeks. Both the intra-class correlation coefficient (Bland and Altman 1996a) and the coefficient of reliability were calculated (Bland and Altman 1986).

Internal consistency was determined by calculating Cronbach's Alpha (Cronbach and Meehl 1955) for the Oxford-12 and by determining the effect on this value of removing any single question. A value for Cronbach's Alpha greater than 0.8 was considered "good" while a value greater than 0.9 was considered "excellent" (Feinstein 1987).

Discriminative ability was tested by comparing the Oxford-12 scores generated for revised and unrevised knees with the Mann-Whitney U-test and by calculating the area under the Receiver Operating Characteristic Curve (Hanley and McNeil 1982, Centor 1991). The same tests were performed for the WOMAC. We hypothesized that the WOMAC and Oxford-12 should have similar discriminative ability.

Results

Translation

The two translated versions of the Oxford-12 were very similar, and a common version was accepted which incorporated aspects of both translations. Back translation of the accepted version was stable. The original and translated versions were judged to be culturally the same, especially since the questionnaire concentrated on concrete concepts related to the knee. The bilingual test-retest results and similarity between forward and back translations showed language equivalence.

Patient burden and feasibility

Patients required median 10 (0.5–60) minutes to complete the Oxford-12. 23% of patients stated that they needed help to complete the questionnaire.

Of the 1,200 Oxford-12 questionnaires posted, 2 were returned by the post office incorrectly addressed and 3 were returned with a note by a family member or caregiver indicating that the patient had died. 1,026 questionnaires were returned at least partly completed, yielding a response rate of 86%. Of these, 89% were complete. The net response rate therefore was 77%.

Psychometric properties

The Oxford-12 correlated most closely with the physical domains and less so with the mental and social domains in all general health questionnaires (Table 1). Correlations with the WOMAC domains were the highest (WOMAC Pain rho = 0.83, WOMAC Stiffness rho = 0.74 and WOMAC Physical Function rho = 0.87). Negative correlations were noted with the SF-36 and SF-12 due to the reciprocal nature of their scoring systems. The Oxford-12 correlated poorly with the eating domain of the SIP (rho = 0.14) thus demonstrating good divergent construct validity.

6.8 % of patients surveyed who completed the questionnaire showed a floor effect. Only 0.1% showed a ceiling effect. The frequency distribution of the score was positively skewed, with a skew value of 0.73. Of the commonest transformations (Bland and Altman 1996b, c), the logarithmic transformation yielded the most Normalized distribution of scores with a skew value of 0.09.

The intra-class correlation coefficient for the Oxford-12 was high at 0.94 (95% CI 0.89–0.96). The mean difference between the two sets of scores was –0.7 (95% CI –2.0–0.6), which was not significantly different from 0 (one sample t-test). The coefficient of repeatability was 9.6 and 95% of the values were within -0.7 ± 9.6.

The internal consistency was excellent with a Cronbach's Alpha of 0.93 (95% CI 0.63–0.84). Removal of any of the 12 items in the calculation of Cronbach's Alpha did not result in a value greater than 0.93.

All three domains of the WOMAC showed a difference between the revised and unrevised groups, both with the Mann-Whitney U test and the area under the Receiver Operating Characteristic Curve (Table 2, Figure). The Oxford-12 displayed similar ability, using the same methods.

Table 2. Ability of Oxford-12 and WOMAC to distinguish between revised and unrevised knee arthroplasty patients

Questionnaire	n	Mann-Whitney U-test	Area under ROC [a] curve	95% CI for ROC curve	Asymptotic sig. ROC curve
Oxford-12	917	p < 0.0001	0.64	(0.58–0.70)	p < 0.001
WOMAC pain	967	p < 0.0001	0.70	(0.64–0.76)	p < 0.001
WOMAC stiffness	986	p < 0.0001	0.66	(0.60–0.72)	p < 0.001
WOMAC physical function	862	p < 0.0001	0.67	(0.60–0.74)	p < 0.001

[a] Receiver operating characteristic curve

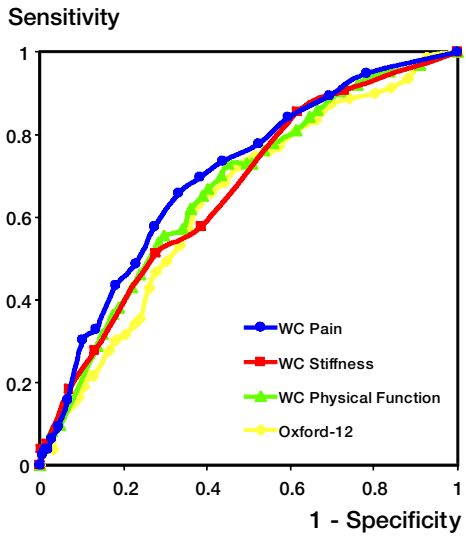

Figure 1. Receiver Operating Characteristic Curve demonstrating comparable ability of the Oxford-12 and WOMAC to discriminate between patients with unrevised and revised knee arthroplasties. Discriminative ability is related to the area under the curve.

Discussion

It is insufficient simply to translate a questionnaire into another language (Guillemin et al. 1993, Guyatt 1993). Instead, a more extensive approach is required in which cultural and language equivalence, as well as psychometric soundness are checked. The Oxford-12 is a relatively concrete questionnaire, hence, cultural and language equivalence were anticipated and subsequently found to be maximal.

Patient-burden imposed by administering the Oxford-12 was minimal, while the feasibility properties were maximal. This has been reported on in more detail in a previous study (Dunbar et al. 1999).

The Swedish translation of the Oxford-12 has proved to be psychometrically sound. As expected, good convergent and divergent construct validity were shown by the Spearman's correlations to the other questionnaires tested. Such correlations mirror those reported by Dawson et al. (1998) for the English validation of the Oxford-12.

The translated version of the Oxford-12 had a definite floor effect but little ceiling effect and a moderate skew to the right. This reflects the overall favorable postoperative status given to the patients by the arthroplasty intervention. The floor effect and skew, however, were acceptable (Brazier et al. 1992, McHorney et al. 1994, Martin et al. 1997, Dunbar et al. 1999). However, logarithmic transformation of the scores should be considered when performing statistical tests (Bland 1995).

Both the intra-class correlation coefficient and the coefficient of repeatability (Bland and Altman 1986) showed good test-retest reliability. The coefficient of repeatability was higher than that published by Dawson et al, but may reflect the larger sample size and higher average patient age in this study. The internal consistency of the translated version of the Oxford-12 was excellent (Feinstein 1987). We found exactly the same value as Dawson et al. did for their postoperative patients (Dawson et al. 1998).

Because of the cross-sectional nature of this study, classic measures of responsiveness are not applicable (Hays et al. 1993). We have instead used the Receiver Operating Characteristic Curve as an indirect measure of responsiveness (Essink-

Bot et al. 1997). The WOMAC and Oxford-12 have comparable discriminative ability. Since the WOMAC has been found to be responsive using more conventional calculations (Roos et al. 1998), these similarities suggest that the Oxford-12 would be equally responsive.

Dawson et al. were able to compute directly responsiveness using the effect-size (Kazis et al. 1989) with preoperative and postoperative Oxford-12 scores (Dawson et al. 1998). An effect size of > 0.8 is considered large, and Dawson et al. reported a large effect size of 2.0. Because of the psychometric similarities between the English and Swedish Oxford-12 Knee Scores, an effect size greater than 0.8 between pre- and postoperative applications of the Swedish Oxford-12 is likely. Therefore, the lack of a direct responsiveness statistic should not preclude the general use of the Oxford-12 in Sweden at this time. Validity is usually a matter of degree rather than an all-or-none characteristic, and validation is an unending process (Nunnally and Bernstein 1994).

In conclusion, we found that the Swedish translation of the Oxford-12 Knee Score is linguistically and culturally equivalent to the English version and that it has solid psychometric characteristics, in keeping with the original questionnaire. This translated version is appropriate for general use with knee arthroplasty patients in Sweden.

We thank Jonas Ranstam for statistical consultation.
Funding has been provided by grants from the Arthritis Society of Canada, the Medical Research Council of Sweden (Project 9509), Stiftelsen för bistånd åt vanföra i Skåne, the Medical Faculty of the University of Lund, and Socialstyrelsen.

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 (12): 1833-40.

Bland M. An introduction to medical statistics. Oxford University Press, New York 1995.

Bland J M, Altman D G. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1 (8476): 307-10.

Bland J M, Altman D G. Measurement error and correlation coefficients. Br Med J 1996a; 313 (7048): 41-2.

Bland J M, Altman D G. Transformations, means, and confidence intervals. Br Med J 1996b; 312 (7038): 1079.

Bland J M, Altman D G. Transforming data. Br Med J 1996c; 312 (7033): 770.

Bombardier C, Melfi C A, Paul J, Green R, Hawker G, Wright J, Coyte P. Comparison of a generic and a disease-specific measure of pain and physical function after knee replacement surgery. Med Care (Suppl 4) 1995; 33: AS131-44.

Brazier J E, Harper R, Jones N M, O'Cathain A, Thomas K J, Usherwood T, Westlake L. Validating the SF-36 health survey questionnaire: new outcome measure for primary care [see comments]. Br Med J 1992; 305 (6846): 160-4.

Centor R M. Signal detectability: the use of ROC curves and their analyses. Med Decis Making 1991; 11 (2): 102-6.

Cronbach L J, Meehl, P.E. Construct validity in psychological tests. Psych Bull 1955; 52 : 281-302.

Damiano A M. Sickness Impact Profile user's manual and interpretation guide. John Hopkins University, Baltimore 1996.

Dawson J, Fitzpatrick R, Murray D, Carr A. The problem of "noise" in monitoring patient-based outcomes: generic, disease-specific and site-specific instruments for total hip replacement. J Health Serv Res Policy 1996; 1 (4): 224-31.

Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. J Bone Joint Surg (Br) 1998; 80 (1): 63-9.

Dunbar M J, Robertsson O, Ryd L. Orthopaedic Research Society Annual Meeting. Anaheim, California 1999.

Essink-Bot M L, Krabbe P F, Bonsel G J, Aaronson N K. An empirical comparison of four generic health status measures: The Nottingham Health Profile, the Medical Outcomes Study 36-item Short- Form Health Survey, the COOP/WONCA charts, and the EuroQol instrument. Med Care 1997; 35 (5): 522-37.

Feinstein A R. Clinimetrics. Yale University Press, New Haven 1987.

Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines [see comments]. J Clin Epidemiol 1993; 46 (12): 1417-32.

Guyatt G H. The philosophy of health-related quality of life translation. Qual Life Res 1993; 2 (6): 461-5.

Hanley J A, McNeil B J. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982; 143 (1): 29-36.

Hays R D, Anderson R, Revicki D. Psychometric considerations in evaluating health-related quality of life measures. Qual Life Res 1993; 2 (6): 441-9.

Hunt S M, McKenna S P, McEwen J, Backett E M, Williams J, Papp E. A quantitative approach to perceived health status: a validation study. J Epidemiol Community Health 1980; 34 (4): 281-6.

Hunt S M, McKenna S P, McEwen J, Williams J, Papp E. The Nottingham Health Profile: subjective health status and medical consultations. Soc Sci Med (A) 1981; 15 (3 Pt 1): 221-9.

Kantz M E, Harris W J, Levitsky K, Ware J E, Jr., Davies A R. Methods for assessing condition-specific and generic functional status outcomes after total knee replacement. Med Care (Suppl 5) 1992; 30: MS240-52.

Kazis L E, Anderson J J, Meenan R F. Effect sizes for interpreting changes in health status. Med Care (Suppl 3) 1989; 27: S178-89.

Lohr K N, Aaronson N K, Alonso J, Burnam M A, Patrick D L, Perrin E B, Roberts J S. Evaluating quality-of-life and health status instruments: development of scientific review criteria. Clin Ther 1996; 18 (5): 979-92.

Martin D P, Engelberg R, Agel J, Swiontkowski M F. Comparison of the Musculoskeletal Function Assessment questionnaire with the Short Form-36, the Western Ontario and McMaster Universities Osteoarthritis Index, and the Sickness Impact Profile health-status measures. J Bone Joint Surg (Am) 1997; 79 (9): 1323-35.

Mathias S D, Fifer S K, Patrick D L. Rapid translation of quality of life measures for international clinical trials: avoiding errors in the minimalist approach. Qual Life Res 1994; 3 (6): 403-12.

McHorney C A, Ware J E, Jr., Lu J F, Sherbourne C D. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. Med Care 1994; 32 (1): 40-66.

Nunnally J C, Bernstein I H. Psychometric theory. McGraw-Hill, New York 1994.

Pollard W E, Bobbitt R A, Bergner M, Martin D P, Gilson B S. The Sickness Impact Profile: reliability of a health-status measure. Med Care 1976; 14 (2): 146-55.

Rissanen P, Aro S, Sintonen H, Slatis P, Paavolainen P. Quality of life and functional ability in hip and knee replacements: a prospective study. Qual Life Res 1996; 5 (1): 56-64.

Ritter M A, Albohm M J, Keating E M, Faris P M, Meding J B. Comparative outcomes of total joint arthroplasty. J Arthroplasty 1995; 10 (6): 737-41.

Roos E M, Klässbo M, Lohmander L S. WOMAC osteoarthritis index: reliability, validity, and responsiveness in patients with arthroscopically-assessed osteoarthritis. Thesis, Lund University, Sweden 1998.

Sullivan M. (Sickness Impact Profile: introduction of a Swedish version of health status indicators). Lakartidningen 1985; 82 (20): 1861-2.

Sullivan M, Karlsson J, Ware J E, Jr. The Swedish SF-36 Health Survey—I. Evaluation of data quality, scaling assumptions, reliability and construct validity across general populations in Sweden. Soc Sci Med 1995; 41 (10): 1349-58.

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 (6): 473-83.

Ware J, Jr., Kosinski M, Keller S D. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996; 34 (3): 220-33.

Wiklund I, Romanus B, Hunt S M. Self-assessed disability in patients with arthrosis of the hip joint. Reliability of the Swedish version of the Nottingham Health Profile. Int Disabil Stud 1988; 10 (4): 159-63.