Osteoarthritis and Cartilage

I C R S International Cartilage Repair Society

OARSI OSTEOARTHRITIS RESEARCH SOCIETY INTERNATIONAL

Short Communication
The 12-item Oxford Knee Score: cross-cultural adaptation into German and assessment of its psychometric properties in patients with osteoarthritis of the knee

F. D. Naal M.D.†*, F. M. Impellizzeri M.S.‡, M. Sieverding M.D.†, M. Loibl M.D.†, F. von Knoch M.D.†, A. F. Mannion Ph.D.‡, M. Leunig M.D.† and U. Munzinger M.D.†
† *Department of Orthopaedic Surgery, Schulthess Clinic, Lengghalde 2, 8008 Zurich, Switzerland*
‡ *Department of Research and Development, Schulthess Clinic, Lengghalde 2, 8008 Zurich, Switzerland*

Summary

Objective: To cross-culturally adapt and validate the Oxford Knee Score (OKS) for use in German-speaking patients with osteoarthritis of the knee.

Methods: After the cross-cultural adaptation (OKS-D), the following metric properties of the questionnaire were assessed in 100 consecutive patients (mean age 66.5 years, 61 women) undergoing total knee replacement: feasibility (percentage of fully completed questionnaires), reliability (Intraclass Correlation Coefficients [ICC] and Bland and Altman's limits of agreement), and construct validity (correlation with the Western Ontario and McMaster Universities Index [WOMAC], Knee Society Score [KSS], Activities of Daily Living Scale [ADLS], and Short Form 12 [SF-12]), floor and ceiling effects, and internal consistency (Cronbach's Alpha, CA).

Results: We received 91.9% fully completed questionnaires. Reliability of the OKS-D was excellent (ICC 0.91). Bland and Altman's limits of agreement revealed no significant bias (−0.2) and a random error of 6.2. Correlation coefficients with the other questionnaires ranged from −0.22 (SF-12 Mental Component Scale [MCS]) to −0.77 (ADLS). We observed no floor or ceiling effects. The CA was 0.83.

Conclusions: The German version of the OKS is a reliable and valid measure for the self-assessment of pain and function in German-speaking patients with osteoarthritis of the knee.
© 2008 Osteoarthritis Research Society International. Published by Elsevier Ltd. All rights reserved.

Key words: Oxford Knee Score, Self-report questionnaires, Total knee arthroplasty, Osteoarthritis, Translation and validation, Clinical assessment, Outcome.

Introduction

In recent years, outcome assessment in orthopaedic surgery has increasingly focused on patient self-report questionnaires[1]. In general, such questionnaires should be short and easy to understand to promote their acceptance by the patient and to reduce the potential for data loss. Moreover, they should be reliable, valid and sensitive to clinical change[1]. The Oxford Knee Score (OKS), a 12-item joint-specific self-administered questionnaire, is one such measure that was developed for the assessment of pain and function in patients with osteoarthritis of the knee[2]. The OKS has proven to be reliable, valid and responsive to clinical change[3,4]. The questionnaire has been translated into different languages and used in several clinical studies[5−8]. However, no German version of the OKS exists, despite the fact that German is spoken by more than 100 million people in Europe. When questionnaires are to be used in other languages, it is not sufficient to simply translate them on a linguistic basis; instead, they must be cross-culturally adapted

following specific guidelines[9] in order to avoid culture-related bias and to retain the original content and construct validity of the questionnaire.

The purpose of the present study was to cross-culturally adapt and validate the OKS for use in German-speaking patients with knee osteoarthritis. We chose to adapt the OKS because it is joint-specific, brief, and has recently proven to be a valid scale that is largely free of bias using Rasch analysis[10]; these qualities are not necessarily guaranteed by other German knee questionnaires.

Material and methods

CROSS-CULTURAL ADAPTATION

The cross-cultural adaptation of the OKS was carried out following the guidelines of the American Association of Orthopedic Surgeons (AAOS) Outcomes Committee[9]. The process comprised five steps, each of which was documented with a written report. Step 1 involved forward translation from English to German by one informed translator (T1, Orthopaedic Surgeon, mother tongue German, fluent in English) and one uninformed translator (T2, mother tongue German, fluent in English). Step 2 comprised synthesis of T1 and T2 into one version (T12), resolving any discrepancies under the supervision of one methodologist. The accuracy of the language in the T12 version was verified by a German language professional. In step 3, two independent back translations of the T12 version from German to English were carried out by native English speakers (BT1 and BT2) who were fluent in German and naive to the outcome measure. Step 4 comprised a consensus meeting of all persons involved in the translation process to resolve any remaining problems,

*Address correspondence and reprint requests to: Dr Florian D. Naal, M.D., Department of Orthopaedic Surgery, Schulthess Clinic, Lengghalde 2, 8008 Zurich, Switzerland. Tel: 41-44-3857322; Fax: 41-44-3857795; E-mail: florian.naal@gmail.com

discrepancies and ambiguities, and to establish the pre-final German version. Step 5 involved pre-testing of the German version in 20 consecutive patients (undergoing total knee arthroplasty (TKA) in our hospital) to examine the accuracy of wording and ease of understanding of the questionnaire.

PATIENT COHORT AND STUDY DESIGN

The validation study involved 100 consecutive patients undergoing primary TKA in our hospital in October and November 2007. The cohort comprised 61 women (61%) and 39 men (39%). The mean age of the patients was 66.5 ± 9.1 years (range 46 to 88). There were no differences in either mean age or gender distribution between the study sample and our routine patient-collective of the last 5 years ($n = 2600$, $P > 0.05$). The age- and gender distribution was as well similar to previous reports[2,7,8]. Our study cohort was hence considered to be representative. The study was approved by the local ethical committee and all patients gave their written informed consent to participate.

A complete set of questionnaires (see outcome tools) accompanied by an explanatory letter was mailed to the patients 1 week prior to their admission for surgery. Patients were requested to fill out the questionnaires at home and bring them along on the day of admission. After completing the first set, 36 patients volunteered to complete a second questionnaire set for the assessment of test−retest reliability. For inter-score correlations all 136 completed questionnaire sets were used.

RELIABILITY

Relative reliability was assessed with the Intraclass Correlation Coefficient ICC(2,1), a two-way random effects model with single measure (absolute agreement) in which variance over the repeated session is considered[11]. Absolute reliability was assessed using the Bland and Altman's 95% limits of agreement, with the mean difference between duplicate scores representing the bias and the 95% confidence interval (CI) representing the random error[12].

CONSTRUCT VALIDITY AND OUTCOME TOOLS

The construct validity of the translated OKS was examined by analysing the strength of the correlation of its scores with those of the Western Ontario and McMaster Universities Index (WOMAC)[13], Knee Society Score (KSS)[14] with knee and function subscales, Activities of Daily Living Scale (ADLS)[15] and Short Form 12 (SF-12)[16] using Spearman's correlation coefficients. We assumed moderate to high correlations between the German OKS and WOMAC, KSS, ADLS and SF-12 Physical Component Scale (PCS) (convergent validity). To examine divergent validity we hypothesised that correlation coefficients between the German OKS and the Mental Component Scale (MCS) of the SF-12 would be lower than those between the OKS and the other scores.

FLOOR AND CEILING EFFECTS

The distribution of floor and ceiling effects of the German OKS was determined by calculating the proportion of individuals obtaining the lowest (12 points) and the highest (60 points) scores, respectively. This indicates the proportion of patients for whom it would not be possible to measure a meaningful improvement (i.e., even lower score) or deterioration (i.e., even higher score) of their condition, as they are already at the extreme of the range.

INTERNAL CONSISTENCY

Internal consistency of the German OKS was examined by calculating Cronbach's Alpha (CA)[17].

STATISTICAL ANALYSIS

All statistical tests described above were carried out using the software-package SPSS (Version 13.0, SPSS Inc, Chicago, IL, USA). *P* values < 0.05 were considered to be statistically significant. Unless otherwise stated, all data are presented as the mean ± standard deviation (SD). Normal distribution of the scores was tested using the Shapiro Wilk *W* test. We only used fully completed questionnaires for the analysis; forms with any missing data were excluded.

Results

TRANSLATION AND CROSS-CULTURAL ADAPTATION

Forward and back translations of the OKS revealed no major problems or language difficulties. Most discrepancies

concerned synonyms for specific expressions, e.g., ''difficulty → Schwierigkeiten → problems''. Pre-testing of the German version in 20 patients revealed no difficulties in comprehension of the items. The German version of the OKS (OKS-D) is shown in Appendix 1.

FEASIBILITY

None of the patients appeared to have had difficulty in completing the OKS-D. Overall, we received 125 fully completed questionnaires (91.8%). There was no specific question that was consistently left unanswered. Missing items appeared to arise randomly.

RELIABILITY

Mean scores for the first and second questionnaires administration were 32.2 ± 6.3 and 31.9 ± 6.4, respectively. The OKS-D demonstrated excellent relative reliability with an ICC for the total score of 0.91 (95% CI 0.82−0.95). Bland and Altman's limits of agreement (absolute reliability) demonstrated no significant bias (−0.2) and a random error of ± 6.2 (total error from −6.4 to 6.0). No heteroscedasticity was found.

CONSTRUCT VALIDITY

The results for all inter-score correlations are presented in Table I. Convergent validity for the OKS-D was demonstrated by its moderate to high correlations in the expected directions with most of the other questionnaire scores. The highest correlation was found between the OKS-D and the ADLS total score (−0.77, $P < 0.001$). The lowest correlation coefficient was found with the MCS of the SF-12 (−0.22, $P = 0.02$).

FLOOR AND CEILING EFFECTS

We found no floor or ceiling effects for the OKS-D. Two patients had scores lying between the lowest value and the random error of measurement (12−18.5 points), and no patients had scores between the highest value and the random error (53.5−60 points).

INTERNAL CONSISTENCY

The OKS-D proved to be internally consistent with a CA of 0.83.

Table I

Mean score values and correlation coefficients between OKS scores and scores of the other measures used

	Score values (mean ± SD)	German OKS (Spearman *R*)	*P* values
WOMAC total	41.4 ± 20.1	0.76	<0.001
WOMAC pain	9.0 ± 4.2	0.71	<0.001
WOMAC stiffness	3.9 ± 1.8	0.43	<0.001
WOMAC function	28.5 ± 15.0	0.76	<0.001
KSS total	125.5 ± 27.4	−0.50	<0.001
KSS knee score	61.4 ± 16.3	−0.28	0.024
KSS function score	64.1 ± 15.9	−0.57	<0.001
ADLS	36.4 ± 12.2	−0.77	<0.001
SF-12 (PCS)	34.1 ± 7.9	−0.57	<0.001
SF-12 (MCS)	55.4 ± 9.5	−0.22	0.023

Discussion

During the past two decades it has become evident that evaluation from the patient's perspective is necessary to allow a complete assessment of disease and treatment effects[1]. The use of practicable, reliable and valid self-assessment questionnaires is hence a prerequisite for successful outcomes research in orthopaedic surgery. The present study demonstrated that the German version of the OKS (OKS-D) is feasible to implement and is a reliable and valid tool for the assessment of pain and function in German-speaking patients with osteoarthritis of the knee.

Our patients had no major difficulties in completing the OKS-D as revealed by detailed interviewing of the 20 patients in the pretest phase and the subsequent high completion rate in the main study of more than 90%. We could not identify any single item that was responsible for non-completion of the questionnaire. In the present study, only fully completed questionnaires were used for subsequent analysis. As suggested by Murray et al., missing items can be managed by transforming the results of the total questionnaire to a 100% scale, as long as no more than two items are missing[4]. We used the scoring procedure originally described by Dawson and colleagues[2]; however, using Murray's calculation method, the proportion of questionnaires not able to be evaluated in the present study would have been only approximately 1%.

The psychometric properties of the OKS-D compared well with those reported for the English version by the original developers of the instrument[2], and with the cross-culturally adapted Dutch, Swedish, and Italian version, respectively[5–7]. The absence of any floor or ceiling effects was reported also for the Chinese, Singapore English, Dutch and Italian versions[6–8].

The construct validity of the OKS-D was obvious by the moderate to high correlations in the expected directions with the other questionnaire scores. The correlation coefficients we found between the OKS-D and the KSS, WOMAC, and SF-12 were comparable with those values previously reported[2,5–8]. To our knowledge, no previous studies have examined the correlation between the OKS and ADLS; in the present study this was the strongest correlation observed. This might be related to the knee-specific characteristics of both measures; the WOMAC, in contrast, is not joint-specific but disease-specific.

It is noteworthy that compared with previous reports[2,7,8], we observed lower preoperative scores on the German OKS for our patients. This was not the result of age-related or gender-related effects since the mean age and gender distribution were comparable among these various studies. It might be that waiting time for surgery plays an important role. Most previous studies were carried out on National Health Service (NHS) patients in the United Kingdom, where waiting times average 12–18 months[18]; in our hospital, in contrast, waiting times for total knee arthroplasty are normally in the range of 6–12 weeks. A deterioration of the clinical status whilst on the waiting list for knee joint replacement surgery has been shown before[19]. Another confounding variable might be related to geographical and cultural differences. Lingard et al. have shown that patient expectations differed between the United States, United Kingdom and Australia[20]. The notion that sociocultural differences may be responsible for different mean score values is also supported by the differences in the general health SF subscale values, seen, for example, between our study and those reported in the Italian study[7]. Therefore, when comparing questionnaire scores across clinical studies carried out in different countries, differences in sociocultural factors and healthcare systems should be borne in mind.

Conclusion

The present study demonstrated that the German version of the OKS (OKS-D) is a practicable, reliable, valid and internally consistent questionnaire for the self-assessment of pain and function in German-speaking patients with osteoarthritis of the knee.

Conflict of interest

No benefit of any kind will be received either directly or indirectly by the authors.

Acknowledgements

Appendix 1. The German version of the Oxford Knee Score (OKS-D)

Oxford Knie Score
Bitte beantworten Sie die folgenden 12 Fragen, indem Sie bei jeder Frage die zutreffende Zahl ankreuzen. Wählen Sie nur eine Antwort pro Frage.

Während der letzten 4 Wochen…

1. Wie würden Sie die Schmerzen beschreiben, die Sie üblicherweise in Ihrem Knie hatten?

(1) Keine
(2) Sehr Gering
(3) Gering
(4) Mässig
(5) Stark

2. Hatten Sie wegen Ihrem Knie Schwierigkeiten, sich selbst zu waschen und abzutrocknen (am ganzen Körper)?

(1) Überhaupt keine Schwierigkeiten
(2) Sehr geringe Schwierigkeiten
(3) Mässige Schwierigkeiten
(4) Extreme Schwierigkeit
(5) Unmöglich zu tun

3. Hatten Sie wegen Ihrem Knie Schwierigkeiten, in ein, bzw. aus einem Auto zu steigen oder öffentliche Verkehrsmittel zu benutzen?
(welches Sie eher benutzen)

(1) Überhaupt keine Schwierigkeiten
(2) Sehr geringe Schwierigkeiten
(3) Mässige Schwierigkeiten
(4) Extreme Schwierigkeit
(5) Unmöglich zu tun

4. Wie lange konnten Sie gehen, bevor Sie starke Schmerzen in Ihrem Knie bekamen (mit oder ohne Stock)?

(1) Keine Schmerzen /> 30 Minuten
(2) 16 bis 30 Minuten

(3) 5 bis 15 Minuten
(4) Nur zu Hause
(5) Gar nicht

5. Wie schmerzhaft war es für Sie wegen Ihrem Knie, nach einer Mahlzeit wieder vom Tisch aufzustehen?

(1) Gar nicht schmerzhaft
(2) Ein wenig schmerzhaft
(3) Mässig schmerzhaft
(4) Sehr schmerzhaft
(5) Unerträglich

6. Haben Sie wegen Ihrem Knie beim Gehen gehinkt?

(1) Selten/nie
(2) Manchmal oder nur am Anfang
(3) Oft, nicht nur am Anfang
(4) Die meiste Zeit
(5) Die ganze Zeit

7. Konnten Sie sich hinknien und danach wieder aufstehen?

(1) Ja, leicht
(2) Mit geringen Schwierigkeiten
(3) Mit mässigen Schwierigkeiten
(4) Mit extremen Schwierigkeiten
(5) Nein, unmöglich

8. Wurden Sie nachts im Bett durch Schmerzen in Ihrem Knie gestört?

(1) Nie
(2) Nur 1 oder 2 Nächte
(3) Einige Nächte
(4) Die meisten Nächte
(5) Jede Nacht

9. Wie sehr haben Schmerzen in Ihrem Knie Ihre normale Arbeit (einschliesslich Hausarbeit) beeinträchtigt?

(1) Gar nicht
(2) Ein wenig
(3) Mässig
(4) Erheblich
(5) Vollständig

10. Hatten Sie das Gefühl, dass Ihr Knie plötzlich ''nachgeben'' oder Sie nicht mehr tragen könnte?

(1) Selten/nie
(2) Manchmal oder nur am Anfang
(3) Oft, nicht nur am Anfang
(4) Die meiste Zeit
(5) Die ganze Zeit

11. Konnten Sie die Haushaltseinkäufe selbst erledigen?

(1) Ja, leicht
(2) Mit geringen Schwierigkeiten
(3) Mit mässigen Schwierigkeiten
(4) Mit extremen Schwierigkeiten
(5) Nein, unmöglich

12. Konnten Sie eine Treppe hinunter gehen?

(1) Ja, leicht
(2) Mit geringen Schwierigkeiten
(3) Mit mässigen Schwierigkeiten
(4) Mit extremen Schwierigkeiten
(5) Nein, unmöglich

References

1. Marx RG. Knee rating scales. Arthroscopy 2003;19:1103−8.
2. 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-B:63−9.
3. Garratt AM, Brealey S, Gillespie WJ, DAMASK Trial Team. Patient-assessed health instruments for the knee: a structured review. Rheumatology (Oxford) 2004;43:1414−23.
4. Murray DW, Fitzpatrick R, Rogers K, Pandit H, Beard DJ, Carr AJ, et al. The use of the Oxford hip and knee scores. J Bone Joint Surg (Br) 2007;89-B:1010−4.
5. Dunbar MJ, Robertsson O, Ryd L, Lidgren L. Translation and validation of the Oxford-12 item knee score for the use in Sweden. Acta Orthop 2000;71:268−74.
6. Haverkamp D, Breugem SJ, Sierevelt IN, Blankevoort L, van Dijk CN. Translation and validation of the Dutch version of the Oxford-12 item knee questionnaire for knee arthroplasty. Acta Orthop 2006;76:347−52.
7. Padua R, Zanoli G, Ceccarelli E, Romanini E, Bondi R, Campi A. The Italian version of the Oxford 12-item knee questionnaire − crosscultural adaptation and validation. Int Orthop 2003;27:214−6.
8. Xie F, Li SC, Lo NN, Yeo SJ, Yang KY, Yeo W, et al. Cross-cultural adaptation and validation of Singapore English and Chinese versions of the Oxford knee score (OKS) in knee osteoarthritis patients undergoing total knee replacement. Osteoarthritis Cartilage 2007;15:1019−24.
9. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine 2000;25:3186−91.
10. Conaghan PG, Emerton M, Tennant A. Internal construct validity of the Oxford knee scale: evidence from Rasch measurement. Arthritis Rheum 2007;57:1363−7.
11. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med 1998;26:217−38.
12. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1(8476):307−10.
13. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. 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.
14. Insall JN, Dorr LD, Scott RD, Scott WN. Rationale of the Knee Society clinical rating system. Clin Orthop Relat Res 1989;248:13−4.
15. Bizzini M, Gorelick M. Development of a German version of the knee outcome survey for daily activities. Arch Orthop Trauma Surg 2007;127:781−9.
16. Ware J Jr, Kosinski M, Keller SD. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996;34:220−33.
17. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika 1951;16:297−334.
18. Lingard EA, Berven S, Katz JN, Kinemax Outcomes Group. Management and care of patients undergoing total knee arthroplasty: variations across different health care settings. Arthritis Care Res 2000;13:129−36.
19. Kapstad H, Rustoen T, Hanestad BR, Moum T, Langeland N, Stavem K. Changes in pain, stiffness and physical function in patients with osteoarthritis waiting for hip or knee joint replacement surgery. Osteoarthritis Cartilage 2007;15:837−43.
20. Lingard EA, Sledge CB, Learmonth ID, Kinemax Outcomes Group. Patient expectations regarding total knee arthroplasty: differences among the United States, United Kingdom, and Australia. J Bone Joint Surg Am 2006;88:1201−7.