R. Padua · G. Zanolì · E. Ceccarelli · E. Romanini
R. Bondì · A. Campi

The Italian version of the Oxford 12-item Knee Questionnaire – cross-cultural adaptation and validation

Accepted: 5 February 2003 / Published online: 2 April 2003
© Springer-Verlag 2003

Abstract Questionnaires must be translated and adapted to suit the cultural characteristics of different countries with different languages; they then need to be validated using a standard process. This was done for the Italian version of the Oxford Knee Questionnaire (OKQ), and its reliability and validity were assessed in a cross-sectional study of 50 consecutive patients referred for gonarthrosis and whose knees were replaced. A “subsample” of 20 patients was assessed prospectively to confirm the reliability of the Italian version, and these results were then compared with other clinical studies. The Italian OKQ scale correlated well with other clinical studies and also showed good results with regard to reproducibility, consistency, and validity when compared to the other published versions of the OKQ.

Résumé Les questionnaires doivent être traduits et doivent être adaptés pour convenir aux caractéristiques cul-

turelles de pays différents avec des langues différentes, et ils ont besoin d’être validés en utilisant un processus standard. Cela a été fait pour la version italienne de l’Oxford Genou Questionnaire (OKQ). Sa précision et sa validité ont été étudiées dans une étude croisée d’un groupe de 50 malades consécutifs opérés par prothèse pour gonarthrose. Un sous-échantillon de 20 malades a été étudié pour confirmer la fiabilité de la Version italienne. Ces résultats ont été comparés avec les autres études cliniques. L’échelle OKQ italienne a bien correspondu avec les autres études cliniques. Il a aussi montré de bons résultats quant à la reproductibilité, la valeur et la validité en comparaison à d’autres versions publiées de l’OKQ.

Introduction

Total knee replacement is one of the most widely used treatments for gonarthrosis, and there are published reports of different methods for assessing results. Many techniques have been developed for pain assessment and knee function [11], and usually depend on the surgeon’s judgment of clinical and radiological data. It is now universally accepted that function and symptoms, which are more relevant to patients [1, 3, 12, 14], are the most important assessment aspects, and in the last 10 years, the development of “validated patient-oriented measures” has added a new dimension to clinical outcome evaluation. Questionnaires designed to assess the outcome of knee arthroplasty are widely used, and large volumes of data are being published [4, 6, 10]. In order to be useful, a questionnaire must be properly validated, and this includes testing its reliability, sensitivity, and responsiveness [9, 13]. Questionnaires that have reached the appropriate standard will then function at least as well as – or some times even better – than an observer scoring system [8].

The use of generic, disease-specific and site-specific questionnaires has been recommended as a part of result assessment. However, another category, the procedure-

R. Padua · A. Campi
Orthopedic Department, San Giacomo Hospital,
Rome, Italy

R. Padua · G. Zanolì · E. Romanini
GLOBE, Evidence-based orthopedic working group
of the Italian Society of Orthopedics and Traumatology,

G. Zanolì
Institute of Orthopedics, University of Ferrara,
Ferrara, Italy

E. Ceccarelli
Institute of Orthopedics, Ospedale Bambin Gesù,
Rome, Italy

E. Romanini
IsPeO, Institute of Orthopedics Specialty,
Rome, Italy

R. Bondì
Institute of Orthopedics, Catholic University,
Rome, Italy

R. Padua (✉)
Via E. Gianturco 4, I-00196 Rome, Italy
e-mail: rpadua@mcLink.it
Tel.: +39-06-32-19092, Fax: +39-06-32-19514

specific questionnaire, recently has been widely accepted for practical reasons [2, 3]. These questionnaires are used to assess the outcome of a single surgical procedure such as a hip or a knee replacement. The Oxford Knee Questionnaire (OKQ) was developed to assess the outcome of knee replacement, and has been widely studied and used [4, 6, 10].

To be useful in culturally different areas using different languages, a questionnaire must not only be translated into the new language but also adapted to any different new cultural characteristics. It must then be validated against the original version. The cross-cultural adaptation guidelines, described by Guillemin et al. [12], are widely accepted and used for the translation and adaptation of questionnaires [9, 12] and we used these for our adaptation and validation of the Italian version of the OKQ.

Materials and methods

After translation, we followed the guidelines of Guillemin et al. [12] for validation.

Translation and cultural adaptation

Translations of the original English OKQ were made by a professional and two “medical” translators, and from these a draft was prepared that included Italian cultural characteristics. A “back-translation” was then made into English, which was checked for any inconsistencies when compared with the original English text. The Italian version was then tested on patients.

Patients

The study involved 50 patients (19 men and 31 women – mean age 68.1, range 50–80 years; SD 7.29) undergoing knee replacement for osteoarthritis. Patients with hip and lumbar spine problems were excluded, as they might have biased the OKQ score [10]. All patients were examined carefully and their diagnosis was confirmed by clinical and radiographic examination.

Each patient then completed both the draft OKQ Italian version and the SF-36 official Italian version [2, 15] under the supervision of medical students before they met their surgeons. [12, 15]. A note was made of the time required for completing both versions of the OKQ and for any difficulties encountered by each patient. The questionnaires were scored according to the accepted criteria.

Outcomes tools

The OKQ questionnaire consists of 12 questions that inquire about symptoms and functions concerning problems relating to knee replacement. Replies are converted to a single score, which is basically an arithmetic average of the responses on a one to five scale, then transformed to a scale of zero (no disability) to 100 (severe disability).

The questions test pain severity degree of difficulty in performing a variety of physical activities resulting from a knee problem. All items have five response choices ranging from “no difficulty” to “unable to perform activity” and from “no symptoms” to “severe symptoms”.

The SF-36 consists of 36 questions on the patient’s general health and thus provides eight separate scale scores: physical functioning (PF), role-physical (RP), body pain (BP), general health

(GH), vitality (VT), social functioning (SF), role-emotional (RE), and mental health (MH). These scores are then combined into two main scores: physical composite score (PCS) and mental composite score (MCS). Very low scores for PCS indicate severe physical disorder, distressful body pain, frequent tiredness, and poor health. Very low scores for MCS indicate frequent psychological distress and severe social and emotional problems. These scoring systems are now widely used in the relevant literature, and validated translations into many languages allows them to be used as a standard scale for comparison.

Testing

Reproducibility was tested by asking 20 randomly selected patients to complete the OKQ twice within an interval of 3–5 days, assuming that during this period the clinical situation had not altered. To minimize the risk of short-term clinical change, treatment of these patients remained constant during the interval. Content validity (item relevance and adequacy for intended use) was tested by health experts involved in the study. Score distribution and ceiling and floor effects were calculated by examining the item responses. We were unable to test the criterion validation due to the absence of a standard validated Italian language disability measurement scale for the lower limbs. Construct validity was tested by comparing OKQ and SF-36 results. We hypothesized that OKQ scores would correlate significantly with physical health scores on the SF-36.

Statistical analysis

Statistical analysis was performed by using the STAT-SOFT (Tulsa, OK, USA) and SPSS 8.0 (SPSS Inc., Chicago, IL, USA) packages. Kolmogorov-Smirnov and Liliefors probability tests were used to assess distribution. Interval measurements (SF-36 and OKQ scores) were normally distributed; therefore, correlation was performed by the parametric test (Pearson’s correlation), and comparison between population subgroups (test-retest groups versus whole population) by Student’s *t*-test to assess age, gender, and patient-oriented measures. Instrument test-retest reliability was assessed with the interclass correlation coefficient (ICC), and Cronbach’s alpha to assess internal consistency.

Results

Translation of the questionnaire was not particularly difficult, and the back translation corresponded very well to the original version. No patient had problems filling in the questionnaire, which took 7.4 min on average. Patients considered most of the OKQ was clear and relevant to their knee condition.

Items responses were well distributed for the OKQ (mean 45.09, SD 8.55; median 47, range 22–58), and no patients had either a maximum or a minimum score. Test-retest reliability showed good results. The ICC was 0.88 ($P < 0.001$), and the internal consistency reached a Cronbach’s alpha of 0.90. No significant differences were found between the whole population and the 20-patients test-retest group concerning gender, age, and disease distribution. Table 1 summarizes our data and the statistical analysis of the correlation between the OKQ and SF-36 scores using the eight domains and two composite scores.

Table 1 Oxford Knee Questionnaire (OKQ) and SF-36 scores and statistical analysis. *SD* standard deviation

	Mean scores (SD)	Correlation with OKQ
OKQ	45.09 (8.55)	
SF-36 (PF)	32.90 (23.02)	r 0.49 $P < 0.02$
SF-36 (RP)	18.60 (35.39)	r 0.53 $P < 0.02$
SF-36 (BP)	18.39 (14.12)	r 0.63 $P < 0.02$
SF-36 (GH)	58.28 (16.09)	r 0.41 $P < 0.05$
SF-36 (VT)	45.35 (17.94)	r 0.43 $P < 0.05$
SF-36 (SF)	55.84 (23.54)	r 0.42 $P < 0.05$
SF-36 (RE)	45.43 (45.54)	r 0.20 NS
SF-36 (MH)	58.41 (20.05)	r 0.26 NS
SF-36 (PCS)	28.48 (07.41)	r 0.69 $P < 0.02$
SF-36 (MCS)	46.42 (10.85)	r 0.21 NS

Discussion

Outcome questionnaires were developed to measure patients' perspectives with regard to symptoms and function. Most available questionnaires are in English and are tailored to the Anglo-Saxon culture. Many are now the de-facto standards for the world scientific community [12, 15]. However, appropriate international use of these tools depends on adapting them to different languages and cultures while maintaining cultural equivalence. For example, a culturally equivalent eating question might refer to forks in some cultures and chopsticks in others. A culturally equivalent mobility question might refer to cars in some cultures and buses, trams, or even carts in others. Yet, it might not be valid to substitute a question about driving a car with one about walking, as these are in different functional domains. Thus, in order to avoid potentially harmful distribution of new questionnaires not comparable with those existing in the literature, a rigorous adaptation process is necessary [14, 16]; mere translation is not enough.

Various recent studies have underlined the statistical scarcity of clinical trials due to sample size [7]. Thus, in order to increase the statistical value of clinical research, it is necessary to have more multicentric and international studies. The presence of culturally equivalent outcome measures allows such studies to be carried out reliably in different countries. In addition, the use of culturally equivalent, standardized questionnaires simplifies the problems of metaanalysis for clinical research [1, 12, 14].

The Italian OKQ questionnaire validation process performed in this study shows it preserves very well the characteristics of reliability and validity of the original

English version [4]. The strong correlation between OKQ and SF-36 scores supports this validity, and our test-retest assessment shows similar values to the Swedish validation paper [5]. Moreover, the lack of floor and ceiling effects reassures us of the validity of the Italian version. Dunbar's [6] recommendations about the use of the OKQ in any "large knee-disease affected population" now can be followed by Italian researchers.

Thus we can state that the OKQ Italian version has been proven to have equivalent assessment values to the original English version. We showed that the basic features of any measuring tool – such as reproducibility, consistency, and validity – have satisfied the appropriate statistical criteria, and we do not recommend the use of the two optional scales, at least for Italian populations.

References

1. Amadio PC (1993) Outcomes measurements. *J Bone Joint Surg [Am]* 75:1583–1584
2. Apolone G, Mosconi P (1998) The Italian health survey: translation, validation and norming. *J Clin Epidemiol* 51:1025–1036
3. Dawson J, Carr A (2001) Outcomes evaluation in orthopedics. *J Bone Joint Surg [Br]* 83:313–315
4. Dawson J, Fitzpatrick R, Murray D, Carr A (1998) Questionnaire on the perceptions of patients about total knee replacement. *J Bone Joint Surg [Br]* 80:63–69
5. Dunbar MJ, Robertsson O, Ryd L, Lidgren L (2000) Translation and validation of the Oxford 12-item Knee Score for use in Sweden. *Acta Orthop Scand* 71:268–274
6. Dunbar MJ, Robertsson O, Ryd L, Lidgren L (2001) Appropriate questionnaires for knee arthroplasty. Results of a survey of 3,600 patients from the Sweden knee arthroplasty registry. *J Bone Joint Surg [Br]* 83:339–344
7. Freedman KB, Back S, Bernstein J (2001) Sample size and statistical power of randomized, controlled trials in orthopedics. *J Bone Joint Surg [Br]* 83:397–402
8. Garellick G, Herberts P, Malchau H (1999) The value of clinical data scoring systems: are traditional hip scoring systems adequate to use in evaluation after total hip surgery? *J Arthroplasty* 14:1024–1029
9. Guillemin F, Bombardier C, Beaton D (1993) Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol* 46:1417–1432
10. Harcourt WG, White SH, Jones P (2001) Specificity of the Oxford knee status questionnaire. The effect of disease of the hip or lumbar spine on patients' perception of knee disability. *J Bone Joint Surg [Br]* 83:345–347
11. Insall JN, Dorr LD, Scott RD, Scott WN (1989) Rationale of the knee society clinical rating system. *Clin Orthop* 248:13–14
12. Padua R, Romanini E, Zanoli G (1998) Analisi dei risultati nelle affezioni dell'apparato muscolo-scheletrico. Guerini Ed, Milan
13. Padua R, Padua L, Ceccarelli E, Romanini E, Bondi R, Zanoli G, Campi A (2001) Cross-cultural adaptation of Lumbar North American Spine Society questionnaire for Italian-speaking patients with lumbar spinal disease. *Spine* 26: E344–347
14. Pynsent PB (2001) Choosing an outcome measure. *J Bone Joint Surg [Br]* 83:792–794
15. Ware JE, Sherbourne C (1992) The MOS 36-items short-form survey (SF-36): I. Conceptual framework and items selection. *Med Care* 30:473–483
16. Zanoli G, Stronqvist B, Padua R, Romanini E (2000) Lesson learned searching for HRQoL instruments to assess the results of treatment in person with lumbar disorder. *Spine* 25:3178–3185