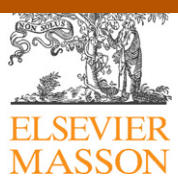
Orthopaedics & Traumatology: Surgery & Research (2012) **98**, 409—412



Available online at

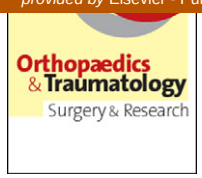
SciVerse ScienceDirect

www.sciencedirect.com

Elsevier Masson France

EM|consulte

www.em-consulte.com/en



ORIGINAL ARTICLE

The Oxford Knee Score: Compared performance before and after knee replacement

J.-Y. Jenny*, Y. Diesinger

Hand and Orthopaedics Surgical Center, Strasbourd Teaching Hospital Centers, 10, avenue Baumann, 67400 Illkirch, France

Accepted: 5 March 2012

KEYWORDS
Knee;
Knee replacement arthroplasty;
Outcomes assessment;
Self-assessment;
Patient reported outcome measures

Summary
Background: Self-administered quality-of-life questionnaires are now crucial to the evaluation of orthopaedic surgical patient-reported outcomes, as they reflect patient satisfaction. The Oxford Knee Score (OKS) is a validated instrument that is widely used to assess outcomes of knee osteoarthritis surgery.
Hypothesis: The relevance of the OKS (comprehension and relevance of the items and responses, and internal and external validity) and its discriminating performance measured based on the ceiling and floor threshold effects are better before than after knee replacement surgery.
Materials and methods: We included 200 patients (100 scheduled for knee replacement and 100 having had knee replacement more than 1 year earlier). The OKS questionnaire was handed to each patient during the first surgeon visit or during a follow-up visit. The American Knee Society (AKS) score was determined simultaneously.
Results: The mean OKS was 43.7 (range, 21—56; SD, 6.9) before surgery and 20.5 (range, 12—45; SD, 5.6) after surgery. The floor effect was absent (0%) before surgery and substantial (33%) after surgery; a weak ceiling effect (7%) was noted before surgery and no ceiling effect after surgery. Internal consistency of the OKS was excellent. The OKS correlated negatively with the AKS knee and functional scores, both before and after surgery.
Discussion: The OKS is well-suited to the evaluation of knee function both before and after knee replacement surgery. Before surgery, the absence of substantial floor and ceiling effects lead to excellent discrimination. After surgery, the substantial floor effect limits the ability to discriminate among the best results. Efforts should be made to develop more demanding scoring systems.
Level of evidence: Level 2. Exploratory cohort study with universally applied reference standards.

* Corresponding author. Tel.: +33 03 88 55 21 45; fax: +33 03 88 55 23 57.
E-mail address: jean-yves.jenny@chru-strasbourg.fr (J.-Y. Jenny).

1877-0568/$ - see front matter © 2012 Published by Elsevier Masson SAS.
doi:10.1016/j.otsr.2012.03.004

Introduction

Measurements of clinical and functional outcomes are crucial to the assessment of healthcare quality. In knee osteoarthritis surgery, the most widely used and accepted instrument is probably the American Knee Society (AKS) score [1]. However, quality-of-life questionnaires that are completed by the patient with no input from the healthcare personnel may be more objective [2], because they more faithfully reflect the experience of the patient. In knee osteoarthritis surgery, the self-assessed Oxford Knee Score (OKS) is widely used in its original English version [3]. The French version of the OKS was validated recently [4].

Studies done before surgery have provided ample evidence that the OKS is relevant (comprehension and relevance of the items and responses, and internal and external validity) [5—8]. However, we are not aware of similar studies assessing the postoperative relevance of the OKS.

Our working hypothesis here was that the relevance (comprehension and relevance of the items and responses, and internal and external validity) and the discriminating performance measured based on the floor and ceiling effects of the OKS differed before and after knee replacement surgery. We included different patients before and after surgery, as our objective was to evaluate the relevance of the OKS before and after surgery, as opposed to longitudinally comparing preoperative and postoperative scores in the same patients.

Material and methods

The Oxford Knee Score (OKS)

The French version of the OKS has been described elsewhere [4]. It has 12 items on daily activities, which the patient must answer without help from healthcare personnel. Each item is scored from 1 (normal function) to 5 (extreme difficulty). The global score is the sum of the 12 item scores; therefore, the best possible score is 12 and the worst possible score is 60.

The American Knee Society (AKS) Score [1]

The AKS score is completed by the healthcare personnel after a patient interview and physical examination. It has two components, a knee score and a function score, each of which is scored from 0 to 100, with higher values indicating better knee condition. The knee component has four items that add points (pain, flexion, mediolateral stability, and anteroposterior stability) and three items that subtract points (malignment, flexion contracture and extension lag). In the functional component, two items add points (walking and stairs) and one item removes points (use of a walking aid).

Study population

We included 200 consecutive patients in 2008 and 2009. Among them, 100 were awaiting total or unicompartmental knee replacement surgery for advanced knee osteoarthritis and 100 had had one or the other of these two procedures more than 1 year before the visit date. For each patient, we collected the following data: age, sex, side affected, body weight, height, body mass index, presence of other orthopaedic disorders likely to affect scoring, and type of surgery (total or unicompartmental knee replacement).

Methods

The French version of the OKS was handed to the patient either during the surgeon visit at which the need for surgery was determined or during a postoperative follow-up visit. The patient was instructed to complete the questionnaire alone or, if needed, with help from family members. On the same day, the surgical team completed the AKS score. All data were entered into an Excel spread sheet then transferred to a statistics computer program (Statview 9.0, SAS Institute France, Grégy-sur-Yerre, France). Preoperative and postoperative data were analysed separately.

The basic data were described using the mean, standard deviation (SD), and range. The Shapiro-Wilk test was used to determine whether the data for each variable were normally distributed. Feasibility of the OKS was assessed by determining the percentage of patients who were unable to complete the questionnaire. A floor effect (affecting the best results) was looked for by determining the percentage of responses below the lowest possible score increased by 1 SD. Similarly, to look for a ceiling effect (affecting the poorest results), we determined the percentage of responses above the highest possible score decreased by 1 SD. Fisher's exact test was used to compare the floor and ceiling effects before and after surgery. Internal consistency was assessed by computing the Cronbach alpha coefficient.

We assessed associations between demographic data and OKS values using Student's *t* test and the Mann-Whitney U test for qualitative variables and the linear correlation coefficient and Spearman's correlation coefficient for quantitative variables.

To assess external validity of the OKS, the linear correlation coefficient and Spearman's correlation coefficient were used to evaluate associations between the global OKS and the AKS scores (global score, knee score, and function score). With the OKS, lower scores indicate better knee condition, whereas the opposite is true of the AKS scores. Consequently, a statistically significant negative correlation would indicate a good clinical correlation.

For all statistical tests, *p* values smaller than 0.05 were considered significant.

Results

We included 76 men and 124 women, with a mean age of 71 years (range, 48—91years; SD, 9 years). The right knee was involved in 103 patients. Mean body weight was 82 Kg (range, 42—144 Kg; SD, 17 Kg). Mean height was 166 cm (range,149—193 cm; SD, 9 cm). Mean body mass index was 29.3 Kg/m^2 (range,17.3—49.3 Kg/m^2; SD, 5.2Kg/m^2). In the postoperative group, mean time since surgery was 18 months (range, 12—24 months; SD, 4 months).

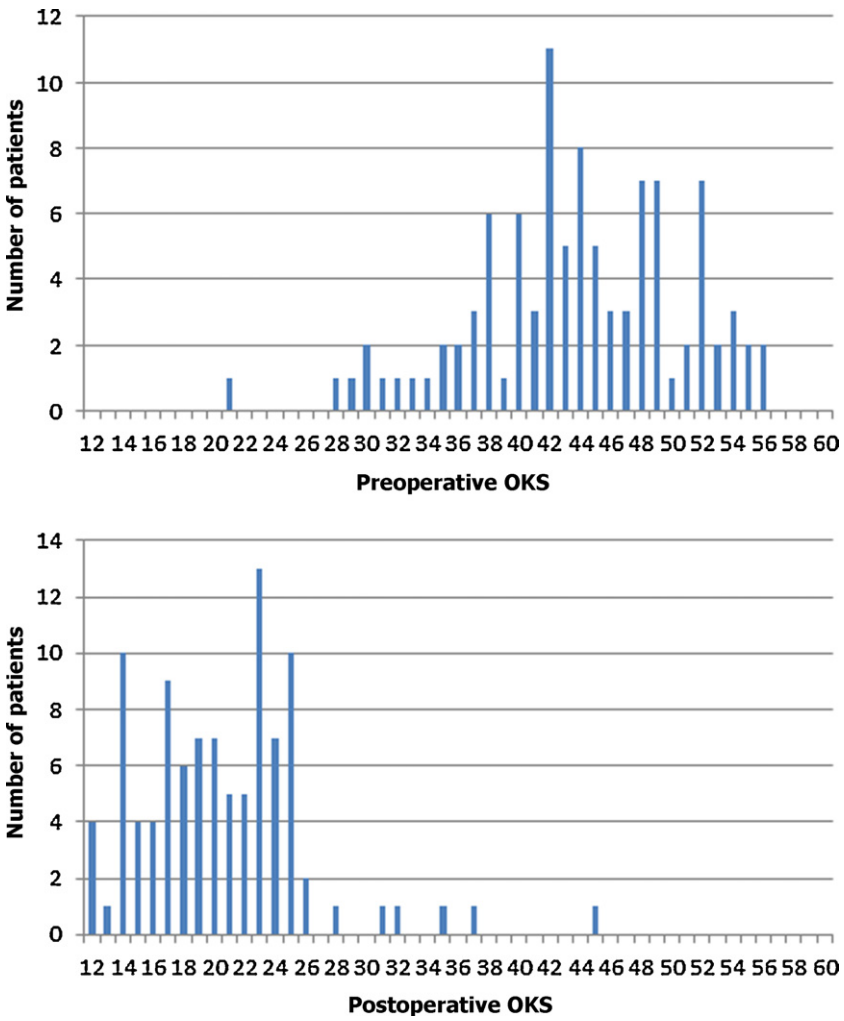


Figure 1 Oxford Knee Score (OKS) in the preoperative and postoperative groups. The dotted lines are the limits of the ceiling and floor effects.

None of the patients had any difficulty understanding any of the OKS items. All patients were able to answer all items, in some cases with help from their family.

Mean OKS was 43.7 (range, 21—56; SD, 6.9) before surgery and 20.5 (range, 12—45; SD, 5.6) after surgery. The values were normally distributed in both groups. In the preoperative group, there was no floor effect (no scores were lower than 19) and the ceiling effect was small (seven scores greater than 53). In contrast, in the postoperative group, there was a substantial floor effect (33 scores lower than 18) with no ceiling effect (no scores greater than 54) (Fig. 1). We found significant differences between the two groups for both the floor effect ($p < 0.001$) and the ceiling effect ($p = 0.007$).

Internal consistency of the OKS was satisfactory, with a Cronbach alpha coefficient of 0.88 in the preoperative group and 0.66 in the postoperative group. No significant associations were demonstrated in either group between the OKS value and age, sex, height, body weight, body mass index, presence of another orthopaedic disorder, or type of surgery planned or performed.

The mean AKS knee score was 40.3 (range, 0—79; SD, 11.9) in the preoperative group and 87.3 (range, 32—100; SD, 11.6) in the postoperative group. The mean AKS function score was 61.4 (range, 0—100; SD, 15.9) in the preoperative group and 92.9 (range,45—100; SD, 10.3) in the postoperative group. Before surgery, a strong negative correlation was demonstrated between the OKS and both the AKS knee score ($r = -0.33$; $p = 0.004$) and the AKS function score($r = -0.47$; $p < 0.001$). After surgery, there was a weak negative correlation between the OKS and the AKS knee score ($r = -0.19$; $p = 0.06$) and a strong negative correlation between the OKS and the AKS function score ($r = -0.49$; $p < 0.001$).

Discussion

Our working hypothesis was partially confirmed: although relevance of the OKS (comprehension and relevance of the items and responses, and internal and external validity) were not significantly different between the preoperative and postoperative groups, discriminating performance was significantly poorer after surgery than before surgery. As expected, we found no difference in item comprehension or in item and response relevance between the preoperative and postoperative groups. Neither did surgery significantly

affect internal validity, as assessed based on the Cronbach alpha coefficient. External validity measured by the correlation between the OKS and the AKS scores was satisfactory, despite slight weakening of the correlation with the AKS knee score after surgery that did not call into question the usefulness of the OKS.

A ceiling effect before surgery indicates that a substantial number of patients have very high preoperative scores indicating very serious knee abnormalities. In this situation, further worsening after surgery might go undetected. Fortunately, a ceiling effect of the OKS is very rare in clinical practice and is therefore not a meaningful disadvantage of this evaluation tool. On the other hand, absence of a preoperative floor effect indicates that even small clinical improvements after surgery will be detected. After surgery, the situation was very different. The OKS values were substantially lower, indicating that the outcomes were usually good. The result was a substantial floor effect (33%), making it difficult or perhaps impossible to detect subtle differences between two patients. Thus, our results indicate that the OKS exhibits better discriminative performance in patients awaiting knee replacement surgery for osteoarthritis than in patients who have already had the procedure. Consequently, the OKS may be less than optimal for some studies, most notably those comparing outcomes after knee replacement surgery. The AKS scores also showed weak performance in our study, with ceiling effects of 43% for the knee component and 83% for the function component. As the clinical and functional outcomes are usually fairly satisfactory after knee replacement surgery, significant differences are probably difficult to detect in populations of reasonable size using currently accepted tools, including self-assessment questionnaires. More specifically, detecting subtle differences is difficult in patients whose outcomes are considered satisfactory based on the usual scores. Clearly, there is a need for developing and using more demanding scores, although the results would then be further from the ideal outcome, which might have a negative psychological impact on the patients and surgeons. Several new instruments have been developed [9—11], but they are still too rarely used, at least in French publications. The French version of the Knee injury and Osteoarthritis Outcome Score (KOOS) has been validated [12] and French versions of the Activity Scale for Arthroplasty Patients (ASAP) and High-activity Arthroplasty Score (HAS) are being validated in parallel with the present study. Future studies should measure the floor and ceiling effects and compare them to those obtained with the OKS in order to determine whether these new tools deserve to be used more extensively.

Conclusion

Our validated French version of the OKS is reliable for evaluating overall knee function in patients awaiting knee replacement surgery. The lower discriminating performance of the OKS after surgery may preclude the detection of subtle outcome differences, most notably in patients whose outcome is considered favourable. The use of more demanding scores would help to conduct a detailed analysis of outcomes after knee replacement surgery.

Disclosure of interest

JYJ has received royalties from AESCULAP.

References

[1] Insall JN, Dorr LD, Scorr RD, Scott WN. Rationale of the Knee Society clinical rating system. Clin Orthop 1989;248:13—4.
[2] Davies AP. Rating systems for total knee replacement. Knee 2002;9:261—6.
[3] Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. J Bone Joint Surg Br 1998;80:63—9.
[4] Jenny JY, Diesinger Y. Validation of a French version of the Oxford knee questionnaire. Orthop Traumatol Surg Res 2011;97:267—71.
[5] Charoencholvanich K, Pongcharoen B. Oxford Knee Score and SF-36: translation and reliability for use with total knee arthroplasty patients in Thailand. J Med Assoc Thai 2005;88:1194—202.
[6] Haverkamp D, BreugemSJ, 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 2005;76:347—52.
[7] Naal FD, Impellizzeri FM, Sieverding M, Loibl M, von Knoch F, Mannion AF, et al. The 12-item Oxford Knee Score: cross-cultural adaptation into German and assessment of its psychometric properties in patients with osteoarthritis of the knee. Osteoarthritis Cartilage 2009;17:49—52.
[8] Xie F, Li SC, Yeo SJ, Yang KY, Yeo W, Chong HC, et al. Cross-cultural adaptation and validation of Singapore English and Chinese version of the Oxford Knee Score (OKS) in knee osteoarthritis patients undergoing total knee replacement. Osteoarthritis Cartilage 2007;15:1019—24.
[9] RoosEM, Roos PH, LohmanderLS, Ekdahl C, Beynnon BD. Knee injury and Osteoarthritis Outcome Score (KOOS): Development of a self-administered outcome measure. J Orthop Sports Phys Ther 1998;78:88—96.
[10] Domzalski T, Cook C, Attarian DE, Kelley SS, Bolognesi MP, Vail TP. Activity scale for arthroplasty patients after total hip arthroplasty. J Arthroplasty 2010;25:152—7.
[11] Talbot S, Hooper G, Stokes A, Zordan R. Use of a new high-activity arthroplasty score to assess function of young patients with total hip or knee arthroplasty. J Arthroplasty 2010;25:268—73.
[12] Ornetti P, Parratte S, Gossec L, Tavernier C, Argenson JN, Roos EM, et al. Cross-cultural adaptation and validation of the French version of the Knee injury and Osteoarthritis Outcome Score (KOOS) in knee osteoarthritis patients. Osteoarthritis Cartilage 2008;16:423—8.