Specificity of the Oxford knee status questionnaire

THE EFFECT OF DISEASE OF THE HIP OR LUMBAR SPINE ON PATIENTS’ PERCEPTION OF KNEE DISABILITY

W. G. V. Harcourt, S. H. White, P. Jones
From the Robert Jones and Agnes Hunt Orthopaedic Hospital, Oswestry, and the University of Keele, Stoke-on-Trent, England

There is a need for the accurate measurement of the outcome after knee surgery. The Oxford Knee Score is being increasingly used since it is reported to be short, simple, inexpensive and validated.

We sent the questionnaire to 346 patients awaiting surgery to the hip or lumbar spine. Only 11% of 141 patients with proximal pathology who denied knee problems gave a maximum score. Their mean score was substantially lower than expected at 28.7 (maximum 48), and was significantly lower than the score of 36.5 obtained from patients after total knee replacement.

We therefore suggest that the frequent coexistence of hip or spinal pathology will significantly alter both the absolute score and any improvement to be expected after knee surgery. Although sensitive to disability originating from the knee the Oxford Knee Score is not sufficiently specific since it is heavily influenced by more proximal pathology.

Received 26 April 2000; Accepted after revision 20 December 2000

It is important to establish a universally acceptable method of measuring the outcome of different surgical procedures used in the treatment of arthritis of the knee.1-4 Purchasers, providers and patients all require accurate information on the likely benefits of alternative procedures, in order to make an informed choice.5 Many instruments have been developed for the assessment of knee pain and function, such as the American Knee Society Score,6 the Bristol Knee Score,7 and the Hospital for Special Surgery Knee Score.8 They are derived from clinical and radiological data and depend on the judgement of the surgeon. The concerns and priorities of the patient and surgeon may differ.2,5

Research in many areas of medicine and surgery has shown that patients can provide reliable and valid judgements of their health status and of the benefits of treatment.9 Recently, a 12-point questionnaire, the Oxford Knee Score, designed specifically for patients undergoing knee replacement has been developed and validated.10 It is reported to be short, practical, reliable, valid and sensitive to clinically important changes over time, and is now being widely applied.11,12 Its specificity, however, has not been reported and it may prove to be unreliable in patients with other causes of pain and disability.

Patients and Methods

We contacted 346 patients who were on the waiting list for surgery to the hip or lumbar spine and asked them to complete the Oxford Knee Score questionnaire. A second questionnaire was included to determine the possibility of coexisting or previous knee pathology.

The Oxford Knee Score is a 12-item questionnaire with five possible responses to each question. Each item is scored from 0 to 4, and the items are summed, thus giving 0 for the worst possible status and 48 for a normal knee. It is designed to be used as a short and simple postal questionnaire and is not affected by the surgeon’s perception of the patient’s disability.

The patients were asked to respond with respect to the knee on the same side as their symptomatic hip, or the leg to which the referral of pain from the back was most prominent. The second questionnaire asked if they had ever attended their general practitioner or a hospital consultant with symptoms from their knees (right or left); and if they had, whether they had been given a diagnosis. They were also asked if they suffered from any systemic inflammatory joint disease such as rheumatoid arthritis or gout. They were also given space to comment further.

It was assumed that patients who had not presented to a doctor with knee symptoms, ipsilateral or contralateral to their more proximal pathology, did not suffer from significant knee pathology. We excluded those with systemic inflammatory joint disease who did not specifically deny any problems with their knees, but included patients who had seen a consultant orthopaedic surgeon, and had been
told specifically that the origin of their knee pain was their hip or back, and not their knee.

We compared our results with a group of patients who had undergone total knee replacement (TKR) and who had completed an Oxford Knee Score questionnaire two years after their procedure.

There were therefore three separate groups of patients: group A, patients awaiting hip surgery who denied having knee symptoms; group B, those awaiting spinal surgery who denied having knee symptoms; and group C, those awaiting either surgery who admitted having knee symptoms. A fourth group who had undergone TKR two years previously comprised group D.

Since the sample sizes are large, we used normal-based tests and scores were summarised by mean, standard deviation, and 95% confidence limits. Comparisons between pairs of independent groups were carried out by using $t$-tests, which were adjusted for unequal variances; 95% confidence intervals were given for the differences of means between groups. More than two groups were compared by using one-way analysis of variance (F-tests) with a Bonferroni adjustment for multiple testing.

**Results**

Of the 346 patients who had been sent questionnaires, 201 replied; only seven questionnaires (3.5%) were not adequately completed and these were excluded. A number of other forms had been annotated by explanatory notes, but these had a score for each item, and they were included. Of the remaining 194 patients, 141 fulfilled the criteria to establish that they had no knee symptoms; 104 were awaiting hip surgery (group A) and 37 spinal surgery (group B). The remaining 53 admitted to having knee symptoms (group C). The results are shown in Table I.

**Comparison of groups.** Table II shows the comparison between all patients awaiting hip or spinal surgery (groups A to C) with those in group D. It is clear that a patient awaiting hip or spinal surgery, whether or not they have had previous knee symptoms, had a significantly lower Oxford Knee Score than a patient two years after TKR.

Comparison of groups A to C, using one-way analysis of variance (F-tests), gave values of $p = 0.15$, $F_{2,191} = 1.9$ indicating that there was no significant difference between the means. Not only does this suggest that the presence of knee pathology does not greatly alter the score in the patient awaiting hip or spinal surgery, but also allows groups A and B to be combined for comparison with group D, once adjustment has been made for unequal variances (Table II).

**Other observations.** Seven of the Oxford Knee Score questions relate to function (questions 2, 3, 6, 7, 10 to 12) and the remaining five to pain (questions 1, 4, 5, 8 and 9). We divided the scores into these two groups (Tables III and IV), and compared them using analysis of variance and carrying out three tests on pairs and adjusting for multiple testing using the Bonferroni adjustment at the 5% level of significance. The only significant difference was between the function score in groups B and C; the remainder showed no significant difference.

**Discussion**

The value of a test lies not only in its sensitivity but also in its specificity. The Oxford Knee Score has been shown to be highly sensitive but not specific. Patients with normal knees but with disorders of other joints had very low scores. The effect of hip pathology, in the absence of knee pathology, was more profound than that of spinal pathology, but this was not statistically significant. Furthermore, both these groups of patients had lower scores than those who had undergone TKR.

The score was originally developed by recording changes within six months of TKR. In those circumstances it is highly likely that there would not have been much change in the status of the other joints. However, when comparing one group of patients with another or by measuring changes over a longer period of time such as after two, five or ten years, the variable status of other joints is
likely to be a highly confounding parameter. Primary osteoarthritis of the knee is often bilateral and there may be severe hip and knee disease in the same patient.

Our findings are consistent with those of Brinker, Lund and Barrack, who studied scoring instruments for the results of TKR and emphasised the importance of matching various clinically relevant factors since they may be as likely to represent differences in the patient populations as those due to operative technique or design of implant. A random population studied by any type of health questionnaire would be expected to show a spread of scores, and patients awaiting surgery to the hip or spine would be expected to show a higher than average rate of coexisting knee disease. Our aim, however, was to look specifically at those patients who denied, or had been deemed not to have, pathology in their knees. Despite this, patients with normal knees but with disorders of other joints showed remarkably low scores.

It is well known that patients with hip and spinal pathology may have pain referred to the region of the knee, and this may partially account for the low scores. If this is the case, it may be impossible to produce a questionnaire which is specific for the knee, and only reflects pathology of the knee. It is also possible, however, that the questions themselves may not be specific enough to detect knee pathology alone. Furthermore, from their occasional annotations, patients appeared to find the questions ambiguous.

Konig et al found that knee rating systems can be influenced by numerous factors linked to patients’ general health and condition. They advised that “adding up knee and functional ratings to an overall result should not be used”, and recommended that the results should be divided into separate ratings for pain and functional score. We applied their recommendation to the Oxford Knee Score, but low pain and function scores were still noted in the absence of knee disorders, suggesting that this approach would not eliminate the effects of symptoms from other joints.

The strengths of our study lie in the large sample of patients with clearly defined disease, a rigorous statistical analysis, and strongly significant results. The weaknesses are that a complete radiological and clinical evaluation of each patient was not possible.

We suggest that equivalent studies on other scoring systems for single joints may demonstrate similar results when applied to patients with more proximal pathology, for instance shoulder scores for patients with neck pathology. We recommend this approach, as a method of testing the specificity of a scoring system for a single joint. It may be that there is no substitute for expert clinical evaluation of a joint.

Dawson et al have demonstrated the internal consistency, reproducibility, and construct validity of the Oxford Knee Score. Its specificity to knee pathology alone, however, must be questioned.

No benefits in any form have been received or will be received from a commercial party related directly or indirectly to the subject of this article.

References