Can pre-operative Oxford hip and knee scores be used to predict patient satisfaction following joint replacement surgery and to guide patient selection?

Assessing patients for joint replacement

We obtained pre-operative and six-month post-operative Oxford hip (OHS) and knee scores (OKS) for 1523 patients who underwent total hip replacement and 1784 patients who underwent total knee replacement. They all also completed a six-month satisfaction questionnaire.

Scatter plots showed no relationship between pre-operative Oxford scores and six-month satisfaction scores. Spearman’s rank correlation coefficients were -0.04 (95% confidence interval (CI) -0.09 to 0.01) between OHS and satisfaction and 0.04 (95% CI -0.01 to 0.08) between OKS and satisfaction. A receiver operating characteristic (ROC) curve analysis was used to identify a cut-off point for the pre-operative OHS/OKS that identifies whether or not a patient is satisfied with surgery. We obtained an area under the ROC curve of 0.51 (95% CI 0.45 to 0.56) for hip replacement and 0.56 (95% CI 0.51 to 0.60) for knee replacement, indicating that pre-operative Oxford scores have no predictive accuracy in distinguishing satisfied from dissatisfied patients.

In the NHS widespread attempts are being made to use patient-reported outcome measures (PROMs) data for the purpose of prioritising patients for surgery. Oxford hip and knee scores have no predictive accuracy in relation to post-operative patient satisfaction. This evidence does not support their current use in prioritising access to care.

Methods

We obtained prospectively gathered information from the database at the South West London Elective Orthopaedic Centre, which is a purpose-built orthopaedic treatment centre opened in 2004 that performs hip and knee replacement surgery for four acute NHS Trusts, covering a population of 1.5 million people in southwest London. Ethical approval was not required. Patients routinely complete an OHS or OKS questionnaire pre-operatively and six months after surgery. For the purposes of this study patients were also asked about their overall satisfaction with the outcome of surgery at six months, as measured on a visual analogue scale (VAS) from 0 to 100 using the following question: ‘We would like you to indicate on this scale your overall satisfaction with the outcome of your operation. Please do this by circling whichever point on the scale best indicates your satisfaction. 0 means not satisfied and 100 means very satisfied.’

The OHS and OKS each consist of 12 questions asking patients to describe their hip or knee replacement experience, and their scores are calculated on a 0 to 100 scale. A mean score of 90 means very satisfied and 10 means not satisfied.
ASSESSING PATIENTS FOR JOINT REPLACEMENT

Table I. South West London Effective Commissioning Initiative (SWLECI) criteria that must be met for the PCT to agree to fund primary hip and knee replacement surgery

<table>
<thead>
<tr>
<th>Primary hip replacement surgery</th>
<th>Primary knee replacement surgery</th>
</tr>
</thead>
<tbody>
<tr>
<td>The patient’s Oxford hip score is ≤ 26 on the 0 to 48 system OR where the patient falls into one of the following groups: 1. Patient complains of severe joint pain; AND Has severe functional limitation; OR Patient has minor to moderate functional limitation, despite the use of non-surgical treatments such as adequate doses of NSAID analgesia, weight control treatments and physical therapies.</td>
<td>The patient’s Oxford knee score is ≤ 20 on the 0 to 48 system OR where he/she complains of intense or severe symptomatology and falls into one of the following two groups: 1. Has radiological features of severe disease; AND Has demonstrated disease within all three compartments of the knee (tri-compartmental) or localised to one compartment.</td>
</tr>
</tbody>
</table>

Knee pain and function during the previous four weeks. Each question is scored on a Likert scale\(^\text{13}\) from 0 to 4, and the overall score is created by adding up the responses. The total score can range from 0 to 48, where 0 is the worst possible score indicating severe symptoms and poor joint function, and 48 is the best score suggesting no adverse symptoms and excellent joint function.

**Statistical analysis.** Spearman’s rank correlation coefficients were used to measure the association between the pre-operative Oxford scores and satisfaction at six months. The median and interquartile range (IQR) was calculated for satisfaction scores at six months, stratified by baseline OHS/OKS (tertiles), and the cut-off points chosen by health planners (≤ 26 for OHS, ≤ 20 for OKS). In order to test whether there was a difference in median satisfaction scores by categories of baseline OHS/OKS, we used a Kruskal-Wallis test (one-way analysis of variance (ANOVA) by ranks), the non-parametric equivalent to ANOVA.

The satisfaction score at six months was simplified into a binary variable of whether or not the patient was satisfied with surgery. A cut-off of ≥ 50 identified 89% as being satisfied after their TKR and 93% after THR. The proportions identified as satisfied using this cut-off were chosen to be consistent with those reported elsewhere.\(^\text{14,16}\) We estimated the percentage of patients who were satisfied with the outcome of their THR or TKR at six months, stratified by pre-operative OHS or OKS. A receiver operating characteristic (ROC) curve analysis was used to determine a cut-off point for the pre-operative OHS or OKS that identified whether or not a patient was satisfied with their operation. The gold standard was whether or not a patient was satisfied at six months, and we identified the cut-off point on the pre-operative OHS/OKS which maximised sensitivity and specificity. The area under the ROC curve (AUC) was interpreted as the probability of correctly identifying whether or not patients were satisfied at six months, based on their pre-operative OHS/OKS. The area ranges from 0.5 (a useless test with no accuracy in distinguishing satisfied from not satisfied) to 1.0 (perfect accuracy).

**Results**

Data were available for 1523 patients after THR and 1784 after TKR who had completed baseline and six-month Oxford scores and a six-month satisfaction questionnaire. Demographic details are described in Table II. There were small differences between patients who did, and did not, respond to the six-month follow-up questionnaire. For TKR those who responded were older, whereas after THR the pre-operative OHS was better. Scatter plots were used to describe the association between a patient’s pre-operative Oxford score and their six-month satisfaction score (Fig. 1). The plots show no relationship: Spearman’s rank correlation coefficients were -0.04 (95% confidence interval (CI) -0.09 to 0.01) between OHS and satisfaction and 0.04 (95% CI 0.01 to 0.08) between OKS and satisfaction. There was no evidence of a difference in median satisfaction scores by categories of baseline OHS or OKS using the Kruskal-Wallis test (Table III).

Patients with a lower OKS (worse pre-operative pain and function) were less likely to be satisfied with their TKR (Table III). No association was observed for patients undergoing THR, with equivalent satisfaction for those above and below the new proposed thresholds.

A ROC analysis was used to determine a cut-off point for the pre-operative OHS/OKS associated with satisfaction after surgery. We obtained an area under the ROC curve of 0.51 (95% CI 0.45 to 0.56) for hip replacement and 0.56 (95% CI 0.51 to 0.60) for knee replacement. If we repeated the analysis using higher and lower cut-offs of ≥ 60 and ≥ 40 to classify whether or not patients were satisfied, our findings remained.

**Discussion**

For a joint replacement to be considered successful, the patient must experience pain relief, functional recovery and satisfaction.\(^\text{17}\) Our data demonstrate that the pre-operative OHS and OKS do not predict which patients will be satisfied or dissatisfied with their total joint replacement six months after surgery. There were some differences between hip and knee replacement: following THR the pre-operative scores were not predictive, whereas after TKR those with the worst pre-operative scores were more dissatisfied. Currently those with a higher OKS would not be
prioritised for surgery according to SWLECI criteria, even though they would have better satisfaction outcomes.

The fact that a proportion of patients are not satisfied with their joint replacement has been consistently reported in several large cohort studies, including the National Joint Registry, with levels of dissatisfaction ranging from 10% to 20%.14-16,18-21 Satisfaction with surgery remains high beyond six months,14,15,18,22,23 with a recent study observing no association between baseline OHS and satisfaction at 12 and 24 months after operation.16 These data challenge the recent trend for using the Oxford PROMs as a pre-operative assessment tool to stratify patients in their requirement for surgery, restricting surgery to those with lower scores. The Oxford scoring system was designed to assess the post-operative outcome following arthroplasty. It cannot be used on its own pre-operatively to predict patient satisfaction as an outcome. With this in mind, it should not be used in isolation as a tool to select patients for hip or knee replacement against an arbitrary threshold. The data suggest that this would unfairly discriminate against patients with higher pre-operative scores, who would do well post-operatively with significant relief of pain and improved mobility. In fact, our data after TKR support surgery in those with higher scores so as to avoid poor outcomes in patients with very low pre-operative scores. Other pre-operative scoring methods have been used to prioritise patients for surgery, such as the Western Ontario and McMaster Universities (WOMAC) osteoarthritis index25 and the New Zealand score.26 Further work is required to determine whether these scores can reliably predict satisfaction.

Satisfaction is an important criterion for assessing outcome. Although most previous studies have found that pre-operative pain and function are not associated with satisfaction after surgery,20,27 some suggest that patients with less pre-operative pain are more likely to be satisfied with the outcome of TKR.21,28 This is consistent with our findings and goes against current thinking, where those with a higher OKS would not be prioritised for surgery even though they would be more satisfied with the outcome. Other factors have been linked to satisfaction, particularly after TKR, where for example it has been shown to be related to fulfil-
ASSESSING PATIENTS FOR JOINT REPLACEMENT

Table III. Satisfaction with surgery six months after joint replacement, stratified by pre-operative Oxford hip (OHS) and knee scores (OKS) (IQR, interquartile range; TKR, total knee replacement; THR, total hip replacement)

<table>
<thead>
<tr>
<th></th>
<th>Number</th>
<th>Satisfied at six months (median [IQR])</th>
<th>Kruskal-Wallis p-value</th>
<th>Satisfied at six months (n, %)</th>
<th>Chi-squared p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TKR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1784</td>
<td>90 (80 to 100)</td>
<td>0.079</td>
<td>1591 (89.2)</td>
<td>0.01</td>
</tr>
<tr>
<td>Pre-operative OKS ≤ 20</td>
<td>954</td>
<td>90 (75 to 100)</td>
<td>0.36</td>
<td>540 (86.7)</td>
<td>0.037</td>
</tr>
<tr>
<td>Pre-operative OKS &gt; 20</td>
<td>830</td>
<td>90 (80 to 100)</td>
<td>0.97</td>
<td>511 (90.0)</td>
<td></td>
</tr>
<tr>
<td>Pre-operative OKS (low) (0 to 16)</td>
<td>623</td>
<td>90 (75 to 100)</td>
<td>0.97</td>
<td>496 (93.1)</td>
<td>0.97</td>
</tr>
<tr>
<td>Pre-operative (medium) (17 to 23)</td>
<td>568</td>
<td>90 (80 to 100)</td>
<td>0.97</td>
<td>506 (92.7)</td>
<td></td>
</tr>
<tr>
<td>Pre-operative (high) (24 to 47)</td>
<td>593</td>
<td>90 (80 to 100)</td>
<td>0.97</td>
<td>540 (91.1)</td>
<td></td>
</tr>
<tr>
<td>THR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1523</td>
<td>100 (90 to 100)</td>
<td>0.45</td>
<td>1415 (92.9)</td>
<td>0.63</td>
</tr>
<tr>
<td>Pre-operative OHS ≤ 26</td>
<td>1170</td>
<td>100 (90 to 100)</td>
<td>0.97</td>
<td>1085 (92.7)</td>
<td></td>
</tr>
<tr>
<td>Pre-operative OHS &gt; 26</td>
<td>353</td>
<td>100 (90 to 100)</td>
<td>0.97</td>
<td>330 (93.5)</td>
<td></td>
</tr>
<tr>
<td>Pre-operative OHS (low) (0 to 15)</td>
<td>533</td>
<td>100 (90 to 100)</td>
<td>0.97</td>
<td>496 (93.1)</td>
<td></td>
</tr>
<tr>
<td>Pre-operative OHS (medium) (16 to 24)</td>
<td>546</td>
<td>100 (90 to 100)</td>
<td>0.97</td>
<td>506 (92.7)</td>
<td></td>
</tr>
<tr>
<td>Pre-operative OHS (high) (25 to 46)</td>
<td>444</td>
<td>100 (90 to 100)</td>
<td>0.97</td>
<td>413 (93.0)</td>
<td></td>
</tr>
</tbody>
</table>

In summary, our data show that pre-operative Oxford hip and knee scores have no predictive accuracy in relation to post-operative patient satisfaction. This evidence does not support their current use in prioritising access to care.

This study received financial support from the Oxford Musculoskeletal NIHR BRU, University of Oxford.

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


