

## The variability in Oxford hip and knee scores in the preoperative period: is there an ideal time to score?

C Quah<sup>1</sup>, D Holmes<sup>1</sup>, T Khan<sup>2</sup>, S Cockshott<sup>1</sup>, J Lewis<sup>1</sup>, A Stephen<sup>1</sup>

<sup>1</sup>Department of Trauma and Orthopaedic Surgery, Royal Derby Hospital, Derby, UK

<sup>2</sup>Department of Academic Orthopaedics and Trauma, University of Nottingham, Nottingham, UK

### ABSTRACT

**BACKGROUND** All NHS-funded providers are required to collect and report patient-reported outcome measures for hip and knee arthroplasty. Although there are established guidelines for timing such measures following arthroplasty, there are no specific time-points for collection in the preoperative period. The primary aim of this study was to identify whether there was a significant amount of variability in the Oxford hip and knee scores prior to surgical intervention when completed in the outpatient clinic at the time of listing for arthroplasty or when completed at the preoperative assessment clinic.

**METHODS** A prospective cohort study of patients listed for primary hip or knee arthroplasty was conducted. Patients were asked to fill in a preoperative Oxford score in the outpatient clinic at the time of listing. They were then invited to fill in the official outcome measures questionnaire at the preoperative assessment clinic. The postoperative Oxford score was then completed when the patient was seen again at their postoperative follow up in clinic.

**RESULTS** Of the total of 109 patients included in this study period, there were 18 (17%) who had a worse score of 4 or more points difference and 43 (39.4%) who had an improvement of 4 or more points difference when the scores were compared between time of listing at the outpatient and at the preoperative assessment clinic. There was a statistically significant difference ( $P = 0.0054$ ) in the mean Oxford scores.

**CONCLUSIONS** The results of our study suggest that there should be standardisation of timing for completing the preoperative patient-reported outcome measures.

### KEYWORDS

Total hip replacement – Total knee replacement – Patient reported outcome measures – Preoperative – Oxford score

Accepted 23 April 2017

### CORRESPONDENCE TO

Conal Quah, E: conalquah@yahoo.com

## Introduction

The demand for hip and knee arthroplasty continues to increase, together with rising quality expectations from both stakeholders and patients. Projected estimates made for the UK show that by 2030 the demand for primary total knee replacement and hip replacements will grow by 117% and 134%, respectively.<sup>1</sup> The delivery of high-quality care to patients is expected alongside the need for managing resources efficiently in this time of economic austerity. These factors have contributed to the increased focus on assessing patient outcomes to measure cost-benefit ratio, quality of care, health gains and health-related quality of life.

Since 1 April 2009, providers of unilateral hip and knee replacements in the UK have been required to collect and report patient-reported outcome measures (PROMs), under the terms of the Standard NHS Contract for Acute Services.<sup>2</sup> In practice, this means that trusts are required to invite patients undergoing these procedures to complete a

preoperative questionnaire in accordance with the relevant guidance.<sup>2</sup> Our institution has been involved in the collection of PROMs since their introduction in 2009. We perform over 1,460 primary joint replacements a year. Patients complete outcome measures in the preoperative assessment clinic in the form of Oxford hip and knee scores and a Euro-Qol five-dimensions questionnaire with a visual analogue score. These are collected and processed by Quality Health, an independent data collection company which provides questionnaires and reports the data to the NHS. A postoperative questionnaire is then posted to patients approximately 6 months after surgery and the results are published monthly by the Health and Social Care Information Centre (HSCIC).<sup>3</sup>

The Oxford hip and knee scores are patient-reported outcome instruments containing 12 questions on the activities of daily living that assess function and residual pain in patients undergoing total knee or hip replacement.<sup>3</sup> Each domain is scored as 0–4 (0 indicating the worst pain or function and 4 indicating no pain or no functional impairment). The overall score is therefore out of a total of 48, ranging

from 0 (most severe symptoms/problems) to 48 (least severe).<sup>4</sup> Orthopaedic specialists can thus use the grading system as a useful aid to ascertain and monitor the severity of patients' symptoms and outcome following surgery. Published PROMs from the HSCIC in November 2016 showed that the national average health gain for Oxford scores post total knee and hip replacement were 16.4 and 21.6, respectively, after case-mix adjustment.<sup>5</sup>

PROMs relating to arthroplasty have been the subject of significant controversy, particularly when used to compare the performance of different trusts in terms of health gains for their patients undergoing these surgical procedures. These results are made available to the general public on the NHS Choices website, together with the expected range for the national average.<sup>5</sup> The main purpose of this publication is to enable patients to make an informed decision about where they would like to be treated. HSCIC has provided some guidance on the timings for the administration of the pre- and postoperative questionnaires. They state that preoperative questionnaires may be administered on the day on which the patient is admitted for treatment or at a pre-assessment clinic beforehand. If the preoperative questionnaire links to a hospital episode, the postoperative questionnaire should be sent out to their home address 6 months after the procedure date.<sup>6</sup> Given the less-defined guidance on when the preoperative questionnaire should be completed, this assumes that the preoperative score is not subject to variability.

The primary aim of this study was to identify whether there was a significant amount of variability in the Oxford hip and knee scores prior to surgical intervention when completed in clinic at the time of listing or at the preoperative assessment clinic.

## Methods

The study population consisted of patients receiving single unilateral total hip or knee replacement under the care of two consultants (JL, SC) at the Royal Derby Hospital between September 2013 and January 2016. These patients were asked to complete a preoperative Oxford score questionnaire in the outpatient clinic at the time of listing. They were then invited to fill in the PROMs questionnaire at the preoperative assessment clinic, which is usually held within 2–4 weeks of the scheduled operation date. The postoperative Oxford score was then completed when the patient was seen again at their postoperative follow-up in clinic. Patient demographics were recorded, including age, sex, American Society of Anesthesiologists (ASA) grade and duration from listing to preoperative assessment clinic and to surgery.

Oxford hip and knee scores collected on the day of listing for arthroplasty were compared with those collected in the preoperative assessment clinic using the paired t-test. For each patient, the difference between postoperative scores and each of the preoperative scores was calculated and paired t-tests were performed for the change in score.

The differences in Oxford scores between the preoperative assessment clinic and the outpatient clinic (time of listing for surgery) were categorised into two groups. The

'small difference' group showed a difference of 3 points or less and the 'large difference' group had a difference of 4 points or more. To assess whether the magnitude of score difference was influenced by the duration between outpatient and preoperative assessment clinics (in days), a paired t-test was used to compare the two groups.

## Results

There was a total of 109 patients who had Oxford scores documented at the time of listing, at the preoperative clinic and at postoperative review. The overall mean age of the patients was 70 years (range 48–88 years) of whom 61% ( $n = 66$ ; mean age 69.8 years, SD 9.69) underwent a total knee replacement and 39% ( $n = 43$ ; mean age 70.7 years, SD 8.9) underwent a total hip replacement. There were 46 male and 63 female patients. Male to female ratios were 31 : 35 for hips and 15 : 28 for knees. The mean number of days from preoperative assessment to surgery was 46.7 days (SD 35.1) for total hip replacement and 38.9 days (SD 49.5) for total knee replacement. The mean outpatient clinic (time of listing for surgery) and preoperative assessment clinic Oxford scores were 16.75 (SD 7.28) and 18.74 (SD 8.12), respectively. This difference was statistically significant ( $P = 0.0054$ ). The mean difference in Oxford score between the post- and preoperative assessment clinics was 17.04 (SD 1.05). Similarly, the mean difference between postoperative and outpatient clinic scores was 16.75 (SD 7.28).

The range of differences in the two scores taken was -17 to +28 points (Fig 1). Forty-six patients (42%) showed a smaller change (0–3 points) in their scores from the time of being listed for surgery to the preoperative assessment clinic. These scores were either in a negative direction (patient slightly worse) or in a positive direction (patient slightly improved). Only six patients (5.5%) had exactly the same score. Sixty-three (57.8%) patients had a score

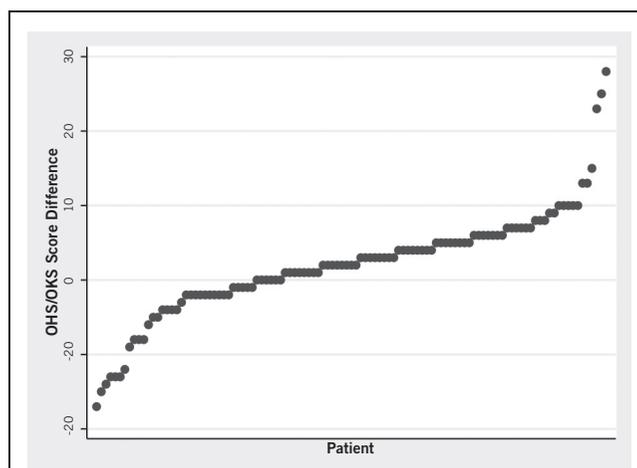


Figure 1 The score difference of the Oxford questionnaire taken at the time of listing for surgery versus the Oxford score taken at preoperative clinic.

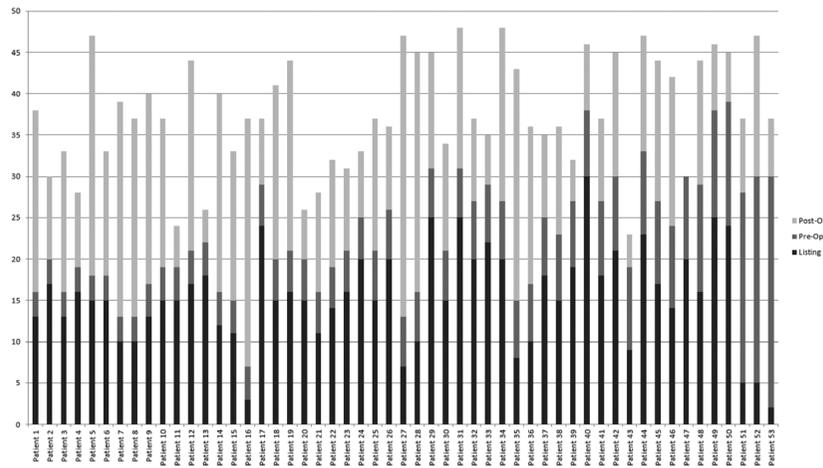


Figure 2 The dark gray bars indicate the difference in both of the Oxford scores taken at the time of listing and at the preoperative assessment for patients that had an improved score of 3 or more points.

difference of 4 or more between the two clinics prior to surgery. Of these, 18 patients (17%) had a worse score of 4 points or greater and 45 (41.3%) had an improved score of 4 points or greater when the scores were compared between time of listing and at preoperative assessment. There was no significant difference in the mean number of days from outpatient clinic (time of listing) to the preoperative assessment clinic for patients with a score difference of less than 4 points (66.04, SD 19.45) compared with those with a difference of 4 points or greater (69.60, SD 37.06;  $P = 0.55$ ).

Figure 2 demonstrates how much difference in potential health gain would have been achieved if the score at the time of listing was used as the ‘official’ score instead of the preoperative assessment score. Using patient 52 as an example, the Oxford score at the time of listing was 5. This improved to 30 at the time of the pre-assessment clinic, which was 7 weeks later. The score 1-year post surgery was 47 and therefore this only resulted in a 17-point improvement following surgery, compared with a 42-point improvement if the score at the time of listing was used as the ‘official’ score instead.

### Discussion

The results of our study have shown that there is significant variability in the Oxford scores in the preoperative period. Nearly 40% of our patients improved by 4 or more points from the time of listing to the preoperative assessment clinic. A study by Beard *et al.* demonstrated that the meaningful changes for the Oxford hip and knee scores after joint replacement surgery was 4 and 5, respectively, hence the difference of 4 or more was chosen for our study.<sup>7</sup>

As far as we are aware, this is the first study that has looked into the variability of Oxford scores in the preoperative period. The explanation for this variability is likely multifactorial. It is possible that there are psychological

factors that may have contributed to the improvement in their Oxford scores. Patients seen in the clinic at the time of listing may have been highly anxious or depressed, as they had been struggling with their symptoms for a period of time and were attending clinic desperate for treatment. Furthermore, patients at the time of listing may have believed that their scores determined the surgeon’s decision on whether or not to proceed with surgery may have scored lower once surgery was confirmed.

A qualitative study of patients listed for total knee replacement revealed that patients underwent a very difficult emotional journey in the period before they decided to have surgery.<sup>8</sup> Many admitted delaying surgery for as long as possible, often years, in the hope of finding less invasive treatments. Thus, it may be possible that our patients were less worried at the time of the preoperative assessment clinic compared with before surgery was confirmed, owing to the resolution of their own indecision about whether or not to have surgery, relief that the surgery had been approved and hope that their problem will be addressed by the impending surgery.

The relative improvement of such a high number of PROM scores preoperatively is a concern because of the ‘ceiling effect’ of the scoring systems.<sup>9</sup> Patients have a smaller capacity for improvement following their surgery, which will be reflected in nationally reported scores by NHS England, pushing the overall rating down. Submission of the most accurate preoperative scores would be beneficial to trusts, which are judged on delta scores for each patient (that is, the greatest possible improvement between the pre- and postoperative scores reflecting the relative health improvements for the patient).<sup>5</sup> This also creates a potential for biased reporting. As we have demonstrated a significant variability in the scores depending on when they are completed, there should be clearer guidance, standardisation and implementation of the preoperative data collection

across providers, particularly with regards to timing of PROM questionnaires. This should allow for more accurate comparison between providers. Certainly within our region, there appears to be some variation as to when the official preoperative PROM questionnaires are collected.

Owing to the pressure of the 18-week target for referral to treatment, hospital trusts are not afforded long waiting times for surgical intervention.<sup>10</sup> A patient may wait approximately 8–12 weeks from the decision to operate being made, to them attending the preoperative clinic where the official NHS PROMs questionnaire is completed. In terms of the management of patients with chronic joint problems, this is a relatively short wait, considering that many patients have suffered degenerative joint pains for many years and have managed their symptoms adequately up to the point of needing surgical intervention. It is therefore arguable that we should not expect to see such sudden vast changes in functional and pain scores. In our opinion, this raises questions as to how these scores should be interpreted.

There are limitations in our study. There is a growing body of evidence to suggest that some patients are prone to 'survey overload' as they are asked to complete questionnaires about their condition on a regular basis.<sup>11</sup> This leads to the completion of surveys in a haphazard manner without any detailed thought. Further investigation would be needed to ascertain if some of our patients fall into this category. There may also be a proportion of patients who do not understand what they are completing, for a variety of reasons, be it poor eyesight, poor literacy skills, learning difficulties or language barriers. We should therefore not underestimate the numbers of patients who may need a degree of assistance to complete the forms accurately. NHS Scotland conducted a 2-year project looking into patient limitations with regards to questionnaire completion.<sup>12</sup> They have made considerable improvements for patient inclusion taking into account that one in five patients has low literacy skills alone and may be too embarrassed to request assistance.<sup>12</sup>

Patients have been shown to have their own interpretation of the questions and variable scores depending on the adequacy of pain relief or pain from other joints on any given assessment day.<sup>15,14</sup> This again raises the question of a potential need for patient guidance when scores are completed. The use of modern technology, via a tablet or smartphone for completion of questionnaires, provides an instant score that could be instantaneously compared with a previous score.<sup>15,16</sup> This would allow the comparison of scores during the consultation; moreover, the clinician would be able to gauge the patients' interpretation of the questions and contextualise the individual.

This service evaluation might suggest that the preoperative clinic (which is usually held a few weeks before surgery) is not the optimal time or place to complete the NHS England PROMs questionnaire because of the high numbers of patients demonstrating a score improvement compared with the time of surgical listing. However, the problem with PROMs collection may be infinitely more complex than it first appears. We do know that staff engagement in the process is crucial and that management must seek to remove the

barriers to appreciating the importance of the data collected.<sup>16</sup> A more patient-centred spotlight by the trust as a whole on the PROMs process may be needed, regardless of wherever we ultimately decide to complete the official preoperative PROMs questionnaire.

## Conclusions

There is a significant disparity between the Oxford scores collected for a high percentage of our patients, which remains unexplained. PROMs are an important tool for measuring the effectiveness of surgical interventions from the patient's perspective. They are intended to represent the patient's judgement of their own health, so will remain a part of the NHS monitoring system for the foreseeable future. They have also started to become part of the financial plans of some clinical commissioning groups when decisions are made for the funding of hip and knee replacement surgery. It is therefore important that the data used are meaningful and accurate. We have highlighted the need to have a full understanding of the data collection process and its pitfalls. We propose that there should be a standardisation of the timing of completion of the preoperative PROMs to allow for more accurate comparison against healthcare providers. What this study effectively shows is the fragility of a scoring system that can be affected by the timing of gathering the scores. It is therefore open to huge scrutiny when comparing patients, surgeons and units.

## Recommendations

- > Educate and engage all staff involved in the process of PROMs collection and interpretation.
- > Explore the available technology for electronic data collection.
- > Limit the collection of outcome scoring to the most meaningful episodes in patient care to prevent 'survey overload'.
- > Incorporate the analysis of PROMs scores into the consultation.
- > Explore the use of designated staff to provide support and assistance for patients in clinical areas where PROMs are routinely completed.
- > Standardise the timing of completion of official preoperative PROMs to ensure more accurate comparison between providers.
- > Exercise caution in the interpretation of the variability of PROMs results between different units and surgeons.

## Acknowledgements

The authors would like to acknowledge Dr Andrea Quah (DPsy), Clinical Psychologist, Monash Health, Victoria, Australia, for her professional opinion on the psychological factors that may have contributed to the improvement in the Oxford scores demonstrated in our study.

## References

1. Patel A, Pavlou G, Mújica-Mota RE, *et al*. The epidemiology of revision total knee and hip arthroplasty in England and Wales: a comparative analysis with projections for the United States. A study using the National Joint Registry dataset. *Bone Joint J* 2015; **97B(8)**: 1,076–1,081.
2. Devlin NJ, Appleby J. *Getting the most out of proms: putting health outcomes at the heart of NHS decision-making*. London: King's Fund; 2010.
3. NHS Digital. Patient Reported Outcome Measures (PROMs). <http://content.digital.nhs.uk/proms> (accessed 27 April 2017).
4. Murray D, Fitzpatrick R, Rogers K *et al*. The use of the Oxford hip and knee scores. *J Bone Joint SurgBr* 2007; **89(8)**: 1010e4.
5. NHS Choices. <http://www.nhs.uk/pages/home.aspx> (accessed 26 April 2017).
6. Health and Social Care Information Centre. *Provisional Quarterly Patient Reported. Outcome Measures (PROMs) in England – Data Quality Note April 2015 to March 2016, November 2016 release*. Leeds: NHS Digital; 2016.
7. Beard DJ, Harris K, Dawson D *et al*. Meaningful changes for the Oxford hip and knee scores after joint replacement surgery. *J Clin Epidemiol* 2015; **68(1)**: 73–79.
8. Jacobson AF, Myerscough RP, Delambo K *et al*. Patients' perspectives on total knee replacement. *Am J Nurs* 2008; **108(5)**: 54–63.
9. Department of Health. *Patient Reported Outcome Measures (PROMs) in England: A Methodology for Identifying Outliers*. London: DoH; 2011.
10. Department of Health. *Referral to Treatment Consultant-Led Waiting Times*. London: DoH; 2015.
11. Health Foundation. *Measuring Patient Experience*. (Evidence Scan no. 18). London: Health Foundation 2013.
12. Healthcare Improvement Scotland. PROMs Knowledge Transfer Partnership Project. [http://www.healthcareimprovementscotland.org/our\\_work/person-centred\\_care/proms\\_questionnaire\\_project.aspx](http://www.healthcareimprovementscotland.org/our_work/person-centred_care/proms_questionnaire_project.aspx) (accessed 26 April 2017).
13. Wylde V, Learmonth ID, Cavendish VJ. The Oxford hip score: The patient's perspective. *Health Qual Life Outcomes* 2005; **3**: 66.
14. Whitehouse SL, Blom AW, Taylor AH *et al*. The Oxford knee score; problems and pitfall. *Knee* 2005; **12(4)**: 287–291.
15. Roberts N, Bradley B, Williams D. Use of SMS and tablet computer improves the electronic collection of elective orthopaedic patient reported outcome measures. *Ann R Coll Surg Engl* 2014; **96(5)**: 348–351.
16. Boyce MB, Browne JP, Greenhalgh J. The experiences of professionals with using information from patient-reported outcome measures to improve the quality of healthcare: a systematic review of qualitative research. *BMJ Qual Saf* 2014; **23(6)**: 508–518.