The Oxford hip scores for primary and revision hip replacement

R. E. Field,
M. D. Cronin,
P. J. Singh

*From Epsom and St
Helier University
Hospitals NHS Trust,
Carshalton, England*

We have used the Oxford hip score to monitor the progress of 1908 primary and 279 revision hip replacements undertaken since the start of 1995. Our review programme began in early 1999 and has generated 3900 assessments.

The mean pre-operative scores for primary and revision cases were 40.95 and 40.11, respectively. The mean annual score for primary replacement at between 12 and 84 months ranged between 20.60 and 22.57. A comparison of cross-sectional and longitudinal data showed no significant differences. All post-operative reviews showed a significant improvement ($p \leq 0.0001$). The 50- to 60-year-old group scored significantly better than the patients over 80 years of age up to 48 months ($p < 0.01$). A subgroup of 826 National Health Service (NHS) and 397 private patients, treated by the senior author (2292 Oxford assessments), had a higher (i.e. worse) mean pre-operative score for the NHS patients ($p \leq 0.001$). The private patients scored better than the NHS group up to 84 months ($p < 0.05$). Patients treated by a surgeon performing more than 100 replacements each year had a significantly better outcome up to five years than those operated on by surgeons performing fewer than 20 replacements each year. The age of the patients at the time of operation, and their pre-operative level of disability, have both been identified as affecting the long-term outcome. Awareness of the influence of these factors should assist surgeons to provide balanced advice.

The Oxford hip score was introduced in 1996 as an instrument to quantify a patient's disability associated with disorders of the hip. It is calculated from the responses to 12 questions on activities of daily living. For each question, the respondent chooses one of five alternative answers. That which corresponds to normal function scores one point and those indicating increasing levels of disability score two, three, four, or five points. The scores for the 12 questions are summated. The minimum total score of 12 points indicates normal function and the maximum score of 60 the most severe disability. The scoring system has been validated in several studies.¹⁻⁴ However, there is a paucity of published data on the normal ranges for pre- and post-operative Oxford scores^{1,4,5} and almost no information to indicate how the Oxford score may change over time.

We have established an outcome programme for all patients undergoing total hip replacement at St Helier Hospital, Carshalton, and a parallel programme for private patients under the care of the first author (REF). The programme was initiated at the beginning of 1999 and has been used to monitor the progress of

patients whose operations were undertaken since the beginning of 1995. Whenever possible, a pre-operative Oxford hip score was obtained during each patient's pre-admission examination. After surgery, postal questionnaires were sent to all patients during the month of each anniversary of their operation. In addition, they were invited to attend an outcome clinic appointment in the month of the second and fifth anniversaries of their operation, for clinical and radiological examination. The review programme was run by members of staff in the orthopaedic research and outcome unit and was independent of any continuing care provided by patients' designated consultants. Staff were available to answer queries regarding the questionnaire, with the exception of the pre-operative assessment, but each patient was asked to complete the form at home so that the influence of interviewer bias was minimised. The operating surgeons were not involved in the collection of the pre-operative baseline Oxford hip score or in the review process. If any problem was identified, an urgent appointment was arranged with the relevant consultant. We thus present the retrospective

■ R. E. Field, PhD, FRCS,
FRCS (Orth), Consultant
Orthopaedic Surgeon
■ M. D. Cronin, BSc, MBBS,
Orthopaedic Research
Fellow
■ P. J. Singh, MRCS,
Specialist Registrar
Orthopaedic Research Unit,
St Helier Hospital, 3th Floor,
Ferguson House, Wrythe
Lane, Carshalton, Surrey
SM5 1AA, UK.

Correspondence should be
sent to Dr M. D. Cronin at 42
Littleton Street, Earlsfield,
London SW18 3SY; e-mail:
michael.cronin@epsom-
sthelie.nhs.uk

©2005 British Editorial
Society of Bone and
Joint Surgery
doi:10.1302/0301-620X.87B5.
15390 \$2.00

J Bone Joint Surg [Br]
2005;87-B:618-22.
Received 10 February 2004;
Accepted after revision
9 June 2004

Fig. 1

Relationship between cross-sectional and longitudinal data over 48 months.

Fig. 3

The pre-operative elements of the Oxford hip score. The error bars show the SD.

Fig. 2

The mean Oxford hip score in the primary and revision groups.

Fig. 4

The post-operative change in the elements of the Oxford hip score.

audit of the Oxford hip scores for patients who have undergone total hip replacement during the past eight years.

Patients and Methods

Oxford hip assessment forms were completed by patients and their responses were transcribed by trained staff onto a computerised database. The original paperwork was archived for validation and audit. Our data set was derived from 1908 primary and 279 revision hip replacements. From these, 3458 Oxford assessments were completed from the primary hip replacement group and 442 from the revision group.

The data analysis was of cross-sectional origin on a heterogeneous group of patients. It was suggested that these data did not provide an accurate prediction of the longitudinal outcome for a specific patient. We therefore compared the cross-sectional data from a single surgeon with longitu-

dinal data from patients who underwent surgery by the same surgeon in 1999.

Analysis was performed on patient variables thought to affect the outcome of total hip replacement, such as gender, body mass index and age. In addition, we compared surgeons undertaking differing numbers of hip replacements each year.

A subset, comprising the patients treated by the first author, was analysed to compare the outcome of 826 primary hip replacements implanted under the National Health Service (NHS) (mean age 68.39 years, SD 12.48) with that of 397 patients who had the same operation performed privately (mean age 65.34 years, SD 13.30). From these, 1594 Oxford assessments were recorded in the NHS group and 698 in the private group.

Statistical analysis. Data were derived from queries built into the computerised database and graphs generated using

Fig. 5

The percentage of patients reaching incremental ranges in their Oxford hip score at one and five years.

Fig. 6

The Oxford hip scores of men and women.

Fig. 7

Change in the Oxford hip scores post-operatively for the different age groups.

Fig. 8

The percentage of patients reaching incremental Oxford hip score ranges at one year in different age groups.

Microsoft Excel (Microsoft Corp, Seattle, Washington). Analysis was undertaken using a two-sample *t*-test with unequal variances (Satterthwaite) or the calculation of a correlation co-efficient, as appropriate. For the former calculations a *p* value of < 0.05 was taken as the level of statistical significance.

Results

The cross-sectional and longitudinal data from the same surgeon's patients are shown in Figure 1. Over the four-year period, there was no significant difference in the Oxford scores at any time point.

The mean and SDs of the Oxford scores of all the patients in our review programme, are shown in Figure 2. The mean pre-operative scores were derived from a total of 724 primary and 96 revision hip replacements. The number of patients at other times was variable because of the longitudinal nature of the data collection. The primary and revision groups had similar pre-operative scores. However, the primary group showed a greater improvement after surgery. The post-operative difference between the primary and revision

groups was statistically significant and was maintained for up to five years. We analysed the 12 elements of the Oxford score obtained from patients, who had undergone primary hip replacement. The mean pre-operative value of each element with the mean and standard deviations are shown in Figure 3. The mean post-operative change, for each of the 12 elements, is represented graphically over time in Figure 4. Statistical analysis has been undertaken for each of the 84 data points. The results showed a significant improvement, at every data point, in comparison with the pre-operative score for that element ($p \leq 0.0001$).

We analysed the results at one and five years after primary hip replacement. The percentage of patients achieving incremental ranges of their Oxford scores at these time points was calculated (Fig. 5). At one year, 46% of patients achieved an excellent score (between 12 and 17). This range was sustained at five years with 44% of patients achieving an excellent score. At one year, the Oxford score for 5% of the patients remained higher than the mean pre-operative score. This number did not significantly increase at the five-year interval.

Fig. 9

The Oxford hip scores of patients operated on by surgeons who perform many (> 100 total hip replacements/yr) versus a few (< 20 total hip replacements/yr).

Fig. 10

The Oxford hip score of patients treated in the NHS or privately by a single surgeon.

The dataset for primary hip replacements was used to compare Oxford scores for men and women (Fig. 6). Statistical analysis showed a significantly higher score in women before operation, and this remained significantly higher up to two years after surgery.

We also analysed the effect of the body mass index (BMI) on patients both pre-operatively and at one year after operation. We found no significant correlation between the BMI and Oxford score at either time point. Comparison of the scores for men and women aged between 60 and 80 years has also been performed. The pre-operative score was no longer significantly different, although the post-operative score for years one and two remained significantly higher in women.

The dataset for primary hip replacements was subdivided according to the age of the patient at surgery (Fig. 7). In addition, we compared the incremental ranges of Oxford scores at one year of patients who were 50 to 60 years old at surgery with those who were 70 to 80 years of age at operation (Fig. 8).

A subset of results from two categories of surgeon was analysed. We compared the results of a single surgeon, who performs more than 100 primary hip replacements each year, with a group of surgeons who perform fewer than 20 each year (Fig. 9). Statistical analysis showed that the mean post-operative score of a patient operated on by the busier surgeon was significantly lower at all time points up to five years after surgery.

Finally, the subset of patients treated by the first author was analysed to compare the Oxford hip scores of patients who had a primary hip replacement under the NHS with those treated privately (Fig. 10). Statistical analysis was undertaken to compare the Oxford hip scores of the two subgroups. The mean age of the NHS group was three years more than that of the private group (68.39 *vs* 65.35 years). The possibility that this age difference may be a significant factor was investigated by comparing the post-operative Oxford scores of all patients aged 65 years and those aged 68 years in the whole study population. There was no difference between the two groups at any time point from before surgery up to eight years after operation.

Discussion

Our dataset does not represent sequential Oxford hip scores for individual patients. However, we have shown that large-volume cross-sectional data are statistically comparable with longitudinal data in our patient population. Furthermore, our data acquisition strategy has ensured that there are reasonable numbers of patients at the later time points, which has enhanced our ability to undertake longer term statistical analyses of the data. The mean pre-operative Oxford hip score of our patients (40.95 points) was slightly lower than that reported by Dawson et al¹ (44 points) and Fitzpatrick et al⁶ (44.5 points). Our score is derived from 724 patients. The score of Dawson et al¹ was obtained from 220 patients and that of Fitzpatrick et al⁶ from 7151 patients. The difference between our value and that of Fitzpatrick et al⁶ is statistically significant ($p < 0.001$). We are unable to provide any satisfactory explanation for this difference.

Comparison of patients with primary or revision hip replacement showed comparable levels of disability at the time that their surgery was undertaken. This finding is consistent with that of Robinson, Palmer and Villar.⁷ By contrast, our primary hip replacement group showed better outcome scores than the revision group. This is statistically significant up to five years ($p < 0.05$).

After primary hip replacement, the post-operative improvement was maintained for up to eight years. When a primary hip replacement fails, the Oxford hip score should increase by about 19 points. This corresponds to almost two standard deviations from the mean and should be detectable for individual patients. As yet, we have not analysed our data for such rises and do not yet know how rapidly such changes will occur. The case for long-term follow-up, using the Oxford hip score, will be greatly enhanced if

such changes can be shown to follow a consistent and identifiable course. Analysis of the 12 components of the Oxford hip score demonstrated considerable variation in the mean pre-operative scores and also the magnitude of the mean post-operative improvement for each element. Only 10% of our pre-operative scores exceeded 50 points. In theory, the sensitivity of the system would be improved if each element generated a higher and more consistent pre-operative score, with larger and more consistent post-operative improvements.

Analysis of the results for men and women patients showed that with age standardised, the pre-operative scores were similar, with men scoring significantly better at one and at two years. We currently do not have any explanation for this observed difference but both genders achieved consistent pre-operative scores despite obvious differences in their activities of daily living.

When a patient's age is taken into consideration, two observations may be made regarding the pre-operative Oxford hip score. First, patients under the age of 50 years scored higher than those in the 50 to 60-, 60 to 70- and 70 to 80-year age ranges. Idiopathic hip degeneration is rare below the age of 50 years. Most of these patients present with longstanding symptoms and frequently report that they were advised that hip replacement should be deferred for as long as possible. Secondly, for patients over 50 years of age, the pre-operative Oxford hip score increases with age. This may be explained by the earlier presentation of younger patients because of their aspiration to undertake more demanding physical activities, their greater awareness of the limitations imposed by a deteriorating hip and their higher expectation of the results of surgery.

While many patients are satisfied to know the mean score for their age group, a proportion wish to know what their own chances are for achieving this outcome. We have established that 46% of patients achieve a score of less than 17 points. This corresponds to normal function in more than half of the activities reviewed. Furthermore, the percentage falls by less than 3% at five years. When age is taken into account, the variation in outcome is noticeable. Almost 60% of patients in the 50- to 60-year range achieve an Oxford hip score of less than 17 points within a year. However, only 41% of the 70 to 80-year patients achieve the same outcome. A small proportion of patients consistently generate Oxford hip scores which exceed the expected normal range. An initial pilot study has indicated that many of these individuals are suffering from multiple musculoskeletal disabilities. Such comorbidity appears to compromise an accurate assessment of the degree to which the patients' activities are impaired by their hip symptoms. The investigation and interpretation of abnormally high Oxford hip scores require further study.

It has previously been suggested that the volume of operations which a surgeon performs may be associated with the long-term outcome of their patients. Analysis has shown that the Oxford score of patients operated on by a

busy surgeon (> 100 each year), is significantly better than that of a group of surgeons operating infrequently (< 20 each year). This difference is maintained for up to five years after operation despite a higher pre-operative score in the patients of the busier surgeon.

Comparison of public (NHS) with private patients revealed a significant difference in their pre-operative Oxford hip scores ($p < 0.0001$). This continued for seven years after surgery. The magnitude of the difference may be greater than we have observed because a proportion of the private group had spent time on an NHS waiting list before deciding to have their hip replacement privately. Although the mean age at surgery of the NHS patients was slightly greater (68.14 years *vs* 65.31 years), we have shown that this does not account for the variation between the groups. The mean pre-operative Oxford hip score of the private patients was lower than that of NHS patients and in general the private patients underwent hip replacement after a shorter period of disability.

Traditional orthopaedic teaching has advised that hip replacement should be deferred for as long as possible. This view originated at a time when the long-term outcome of hip replacement was unknown and was reinforced by reports of relatively poor long-term survivorship for hip replacement in younger patients.^{8,9} Our dataset has shown that younger patients achieve better outcomes and that delay in surgery may compromise a patient's prospects of achieving an optimal result.

Supplementary material

A further opinion by Professor Andrew Carr is available with the electronic version of this article on our website at www.jbjs.org.uk

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

1. Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. *J Bone Joint Surg [Br]* 1996;78-B:185-90.
2. Dawson J, Fitzpatrick R, Murray D, Carr A. Comparison of measures to assess outcomes in total hip replacement surgery. *Qual Health Care* 1996;5:81-8.
3. Dawson J, Fitzpatrick R, Murray D, Carr A. The problem of 'noise' in monitoring patient-based outcomes: generic, disease specific and site-specific instruments for total hip replacement. *J Health Serv Res Policy* 1996;1:224-31.
4. McMurray R, Heaton J, Sloper P, Nettleton S. Measurement of patient perceptions of pain and disability in relation to total hip replacement: the place of the Oxford hip score in mixed methods. *Qual Health Care* 1999;8:228-33.
5. Dawson J, Jameson-Shortall E, Emerton M, et al. Issues relating to long-term follow-up in hip arthroplasty surgery: a review of 598 cases at 7 years comparing 2 prostheses using revision rates, survival analysis, and patient-based measures. *J Arthroplasty* 2000;15:710-17.
6. Fitzpatrick R, Morris R, Hajat S, et al. The value of short and simple measures to assess outcomes for patients of total hip replacement surgery. *Qual Health Care* 2000;9:146-50.
7. Robinson AHN, Palmer CR, Villar RN. Is revision as good as primary hip replacement: a comparison of quality of life. *J Bone Joint Surg [Br]* 1995;81-B:42-5.
8. Hozack WJ, Rothman RH, Booth RE Jr, et al. Survivorship analysis of 1041 Charnley total hip arthroplasties. *J Arthroplasty* 1990;5:41-7.
9. Dorr LD, Kane TJ 3rd, Conaty JP. Long term results of cemented total hip arthroplasty in patients 45 years old or younger: a 16 year follow up. *J Arthroplasty* 1994;9:453-6.