ABSTRACT

Objectives: To assess changes in quality of life and costs of patients undergoing primary total hip replacement using the Exeter prosthesis compared with a hypothetical ‘no surgery’ group.

Design: The incremental quality of life, quality-adjusted life years (QALYs) and cost of Exeter Primary Outcomes Study patients was compared with hypothetical ‘no surgery’ group over 5 years. Scores from annual SF-36 assessments were converted into utility scores using an established algorithm and the QALY gains calculated from pre-operative baseline scores. Costs included implant costs and length of stay.

Setting: Secondary care hospitals.

Participants: Patients receiving a primary Exeter implant enrolled in five of seven Exeter Primary Outcomes Study centres.

Results: On average, patients gained around 0.8 QALYs over 5 years. Younger and male patients or those with lower body mass index and poorer Oxford Hip Scores were significantly associated with increased QALYs. Treatment costs for a primary episode of care were just over £5000 (95% CI £4588 to £5612) per patient. Compared with ‘no surgery’, the cost per QALY was £7182 (95% CI £6470 to £7678), and this remained stable when key cost parameters were varied. The most likely cost per QALY was between £7058 and £7220. Older patients (age 75+) cost more, mainly due to longer average hospital stays and had a higher cost per QALY, although this remained below £10 000.

Conclusions: 85% of cases had a cost of <£20 000 per QALY (with 70% having a cost per QALY under £10 000) compared with no surgery. Cases would be considered cost-effective under currently accepted thresholds (£25 000—£30 000) compared with ‘no surgery’. However, depending on age and severity, younger patients and more severe patients had below average cost per QALYs. These results help to confirm the long-term benefits and cost-effectiveness of total hip replacement in a wide variety of patients using well-established implant models such as the Exeter. However, further and ongoing economic appraisal of this and other models is required for comparative purposes.

ARTICLE SUMMARY

Article focus
- The cost-effectiveness of the Exeter THR compared with no treatment.
- The quality of life gain and incremental number of QALYs gained and cost.
- The cost per QALY by age, sex, OHS and BMI.

Key messages
- There have been few good prospective economic evaluations of THR that measure quality of life, preoperative severity of disease and control for prosthesis type.
- THR in EPOS patients was found to be cost-effective (compared with no treatment). Cost per QALY was below the accepted NICE threshold in all groups and under all sensitivity assumptions.

Strengths and limitations of this study
- Longer term follow-up of patients is advantageous in assessing the economic benefits of THR and this study was exceptional for the length of time in which this was possible.
- The hypothetical control group could provide only an indirect comparison with other interventions and prostheses but a sound new estimate of the absolute cost-effectiveness of THR.

INTRODUCTION

Hip replacement is one of a few cardinal and successful operations in the NHS and yet it has mainly gone unchallenged from a cost-effective perspective since its inception. However, with the current financial environment in the NHS, it is important to reassess the costs and benefits and to dispel any uncertainties surrounding its cost-effectiveness for the majority of patients.

Long-term studies of hip replacements have not provided good conclusive economic proof, which is surprising, given the longevity of the procedure. However, implant models come and go and many get modified, so the
field is not static for economic assessment. It is therefore important to assess those that are stable and have endured in practice over many years. The orthopaedic outcome literature is mainly based on studies of the implant’s survival and few have incorporated a full economic analysis. Most economic studies nowadays rely on modelling, given the limitations of the randomised controlled trial design and well-designed studies compare alternative models.\(^1\) Good economic studies of total hip replacement (THR) also need an appropriate measure of patient outcome (preferably in terms of changes in health-related utility) as well as a robust costing.\(^4\) This study has addressed these issues but further work is now ongoing to assess the longer-term cost-effectiveness.

**PATIENTS AND METHODS**

The Exeter Primary Outcomes Study (EPOS) is one of the largest longitudinal studies of a single prosthesis undertaken in the UK and continues to follow-up recipients for 10 years post-operatively. In this seven-centre study, a case series of 1589 patients underwent hip replacement with the Exeter implant between March 1999 and February 2002. This retrospective economic study was undertaken at the 5-year follow-up stage in patients who had the required outcome data available. These patients were compared with a hypothetical ‘no surgery’ group in terms of their additional costs and outcomes (quality of life (QoL)) from baseline.

The cost of an episode of care (including the operation) was estimated using the average cost of the implant (in the year the operation was performed) and the actual number of bed days used per patient at the Department of Health reference cost per day (for the year in question). The SF-36 patient outcome score was collected annually in the main study. Provided patients did not die or were not lost to follow-up, the change in their QoL between pre-operative assessment and 5-year follow-up was assessed.

Due to a lack of a control group, the quality-adjusted life year (QALY) gain could only be compared hypothetically with the QoL estimates that might have prevailed without surgery. For this, we decided to use a patient’s pre-operative QoL (as measured by the SF-36 score) as the counterfactual scenario. We recognised that in reality, without surgery, QoL might have improved or possibly even deteriorated.

The SF-36 is a multipurpose short-form health survey of 36 questions that yields an eight-scale profile of functional health and well-being as well as psychometrically based physical and mental health summary measures.\(^6\) However, in the EPOS, completion of this was optional and only 938 patients had sufficiently complete scores for the economic study. We calculated the overall SF-36 score following the user’s manual guidance. However, the SF-36 does not directly provide a utility score, so we used the Brazier algorithm to convert SF-36 questions to respective utility scores.\(^7\) This utility score captures a patient’s value for being in different health states on a scale from 0 (death) to 1 (perfect health). Where some patients did not have an SF-36 completed every year over the 5 years, we used their previous year’s SF-36 score to derive the utility score for the missing year. This approach seemed reasonable, as after the first post-operative assessment, most utility scores for the group as a whole did not change greatly over remaining years.

We calculated QALY gains made each year compared with the pre-operative baseline using an ‘area under the curve’ approach (see figure 1). In reality, other treatments might have improved this baseline score and reduced the net potential utility gain found in this study. On the other hand, we might have assumed the condition worsened. The direction of change could not be
known with any certainty and therefore we assumed that no change occurred in QoL from baseline over the 5-year period. The QALY gain per patient was calculated up to the 5-year follow-up period or until the last annual review before death or revision (whichever occurred first). As the EPOS excluded patients who had revisions, we could only assume a zero QoL gain after revision had taken place (although this would be likely to be higher).

As resource use data were not collected in the main EPOS, we retrospectively constructed a proxy cost per case. This was based on each episode of care and was a composite of the national variable cost per day (for the year in which the surgery was performed) adjusted for the patient’s own length of stay plus the estimated total cost of the Exeter implant and other components added as the fixed cost element. An uplifted value of £7500 was used for the cost of a revision.3

Finally, the average cost per QALY was calculated for patients, assuming a zero cost for no surgery. Again this would have been unlikely but we felt it was a conservative assumption (as other treatment costs would be incurred in reality) reducing the cost difference between the two options.

Statistical methods
Analyses were carried out in Stata V.10.8 Bootstrapped, bias-corrected methods were used to calculate 95% CIs for costs per QALY.9 Multiple linear regression was used to model QALY gains, with standard t-test and F-test used to evaluate the significance of β coefficients and model fit.

SF-36 dimension scores were calculated using recommended procedures. Missing values were replaced by scale means if valid responses were available for at least half of the scale items. For the items used in the utility scores, we used chained equations (ICE) to estimate the missing values based on the values of all the other variables in the data set.10 This was carried out when fewer than half of the values for the items used in the calculation of a utility score (from the SF-36) were missing in any individual questionnaire. However, where more than half of the values needed were missing, questionnaires were excluded from the analysis.

RESULTS
The EPOS data set contained complete information on 1589 patients who received the Exeter implant or partial component. However, only five of the seven centres collected QoL data on 1087 patients. Of these, 938 (86%) had sufficient data to be included in the economic study. During the 5-year follow-up, there were 4598 potentially useable questionnaires from the 1087 patients. One thousand and eight of these had completed a baseline score and at least one other follow-up questionnaire completed in sufficient detail for utility

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Utility score and change from baseline utility by year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>938</td>
</tr>
<tr>
<td>Year 1</td>
<td>835</td>
</tr>
<tr>
<td>Year 2</td>
<td>728</td>
</tr>
<tr>
<td>Year 3</td>
<td>550</td>
</tr>
<tr>
<td>Year 4</td>
<td>389</td>
</tr>
<tr>
<td>Year 5</td>
<td>720</td>
</tr>
</tbody>
</table>

Figure 2 Utility score by baseline and subsequent year.

Figure 3 Quality-adjusted life years (QALYs) gained in Exeter Primary Outcomes Study patients up to 5 years.
estimation. Use of the multiple imputation method (see above) was made to adjust 344 of these surveys. Seventy people with important missing prognostic indicators were excluded, leaving 938 subjects for the main economic analysis. There was good SF-36 completion at the 5-year follow-up date (77%) with 720 patients having this maximum follow-up score. There were 69 deaths and only 17 revisions in the study population within the 5-year follow-up period.

Table 1 shows the characteristics of patients included in the economic study.

Nearly two-thirds of patients were women. The mean age of all cases was 62 years, but there was a large SD and upper and lower ranges. The average patient was not obese before surgery, although there was considerable variation in body mass index (BMI). The average patient fell into the third quintile of severity on the Oxford Hip Score (OHS).

Changes in utility scores from baseline to Year 5 are shown in table 2 and figure 2. The change in utility scores varied little after the initial large gain in the first post-operative year. The largest component of the increase in overall utility (around 0.18) was seen in the first year after operation. Much smaller changes were found in subsequent years. Both the overall SF-36 score and the individual dimensions that comprise it showed similar changes during this period. The largest changes on the SF-36 occurred in physical functioning (a 47 point increase), physical role functioning (+50 points) and pain (+48 points).

The QALYs gained was calculated over this 5-year period shown in figure 3. The majority of patients (90.7%) gained positive QALYs compared with no surgery. These gains were approximately normal distributed around a mean value of 0.8 QALYs (95% CI 0.76 to 0.84). However, a small group of patients (9.3%, n=87) lost QALYs (in a theoretical sense, they would have been better without surgery).

In terms of estimating cost per episode, average length of stay was 10.8 days (SD 7.3) and the median estimated cost per patient was £5084 (IQR: £4588—£5812). The distribution of costs is shown in figure 4.

Table 3 shows the average QALYs gained and combined with the average cost to derive a cost per QALY. In order to take account of variation and uncertainty in these estimates, we calculated the associated CI using bootstrapping simulation methods. The average cost per QALY for all 938 subjects was £7182 (95% CI £6740 to £7678).

![Figure 4](https://via.placeholder.com/150)

**Figure 4** Cost of primary hip replacement in Exeter Primary Outcomes Study patients.

<table>
<thead>
<tr>
<th>N</th>
<th>Length of stay/day</th>
<th>Estimated cost/£</th>
<th>QALY gain</th>
<th>Cost per QALY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Median</td>
<td>IQR</td>
</tr>
<tr>
<td>All</td>
<td>938</td>
<td>10.8</td>
<td>7.3</td>
<td>5084</td>
</tr>
</tbody>
</table>

*Bootstrapped, bias-corrected CI, based on 10 000 bootstrap replications.

QALY, quality-adjusted life year.

<table>
<thead>
<tr>
<th>N</th>
<th>Length of stay/day</th>
<th>Estimated cost/£</th>
<th>QALY gain</th>
<th>Cost/QALY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Median</td>
<td>IQR</td>
</tr>
<tr>
<td>23–49 years</td>
<td>82</td>
<td>9.5</td>
<td>3.0</td>
<td>4776</td>
</tr>
<tr>
<td>50–59 years</td>
<td>149</td>
<td>9.4</td>
<td>5.4</td>
<td>4720</td>
</tr>
<tr>
<td>60–64 years</td>
<td>126</td>
<td>10.3</td>
<td>5.1</td>
<td>4981</td>
</tr>
<tr>
<td>65–69 years</td>
<td>190</td>
<td>9.6</td>
<td>2.8</td>
<td>4981</td>
</tr>
<tr>
<td>70–74 years</td>
<td>162</td>
<td>10.5</td>
<td>4.8</td>
<td>5084</td>
</tr>
<tr>
<td>75–90 years</td>
<td>229</td>
<td>13.7</td>
<td>12.1</td>
<td>5654</td>
</tr>
</tbody>
</table>

*Bootstrapped, bias-corrected CI, based on 10 000 bootstrap replications.

QALY, quality-adjusted life year.
We also analysed the QALYs and cost per QALY by age group as shown in Table 4. As might be expected, QALY gains were significantly lower in older patients, with the largest gains found in younger patients. Conversely, cost per QALY increased in older age groups because of the increased length of stay combined with a lower QALY gain.

With regard to the baseline OHS, we found that pre-operative severity was a good predictor of cost-effectiveness. The poorer the initial score on the OHS, the greater the QALY gain found and similarly the lower the cost per QALY. There were significant differences in the cost per QALY between quintile 1 (least severe) and quintiles 3, 4 and 5 (most severe) and also between quintiles 3 and 5 (see Table 5).

The QALY gain was approximately normally distributed and therefore linear regression could be carried out to determine patient and treatment characteristics associated with total QALYs. Age, BMI and OHS were significantly associated with QALYs gained (see Table 6).

Given the limitations of the cost data on which the study was based, a sensitivity analysis was undertaken to determine robustness of the cost and cost per QALY results (see Table 7).

Using various cost assumptions, mean estimated cost per case varied from £4950 to £5516. Given a small variation in cost, the cost per QALY remained relatively stable in the range of £7058 to £7220. This confirmed that the cost-effectiveness results were robust and insensitive to some relatively large changes in cost assumptions. This is also reassuring in terms of potential variability in costs between treatment centres and/or surgical practice that occurs in practice.

A cost per QALY threshold analysis is shown in Figure 5. Over 85% of cases had a cost per QALY of £20 000 or less with 70% of these having a cost per QALY under £10 000 thus making it very cost-effective when compared hypothetically with no surgery. However, 40 cases had a cost per QALY over £50 000. These patients were largely those where the QoL gain was very small rather than due to their cost being above average.

**DISCUSSION**

We have shown that from the perspective of the absence of surgery, the majority of EPOS subjects were treated cost-effectively. The value to patients in terms of their health utility and QALY gains has been demonstrated.

### Table 5  Length of stay, estimated cost, QALY gain and cost/QALY by Oxford Hip Score (OHS) quintile

<table>
<thead>
<tr>
<th>OHS (score band)</th>
<th>N</th>
<th>Length stay/ day</th>
<th>Estimated cost/£</th>
<th>QALY gain</th>
<th>Cost/QALY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean SD</td>
<td>Median IQR</td>
<td>Mean 95% CI</td>
<td>Mean 95% CI*</td>
</tr>
<tr>
<td>Quintile 1 (13–37)</td>
<td>191</td>
<td>10.2 4.1</td>
<td>5084 4588–6075</td>
<td>0.61 0.52 to 0.70</td>
<td>9188 7893 to 10915</td>
</tr>
<tr>
<td>Quintile 2 (38–43)</td>
<td>206</td>
<td>11.3 10.2</td>
<td>4981 4588–5767</td>
<td>0.83 0.74 to 0.91</td>
<td>7102 6126 to 8186</td>
</tr>
<tr>
<td>Quintile 3 (44–47)</td>
<td>174</td>
<td>10.4 5.7</td>
<td>4981 4391–5812</td>
<td>0.70 0.61 to 0.80</td>
<td>7907 6814 to 9253</td>
</tr>
<tr>
<td>Quintile 4 (48–51)</td>
<td>191</td>
<td>11.0 7.0</td>
<td>5233 4720–6160</td>
<td>0.89 0.79 to 0.99</td>
<td>6628 5830 to 7577</td>
</tr>
<tr>
<td>Quintile 5 (52–60)</td>
<td>176</td>
<td>11.2 7.5</td>
<td>5084 4720–5812</td>
<td>0.98 0.87 to 1.08</td>
<td>5924 5189 to 6826</td>
</tr>
</tbody>
</table>

*Bootstrapped, bias-corrected CI, based on 10 000 bootstrap replications.

QALY, quality-adjusted life year.

### Table 6  Linear regression modelling of quality-adjusted life years (n=938)

<table>
<thead>
<tr>
<th></th>
<th>β Coefficient</th>
<th>SE (β)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>–0.009</td>
<td>0.002</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Centre age centred/C0</td>
<td>–0.027</td>
<td>0.011</td>
<td>0.012</td>
</tr>
<tr>
<td>Sex</td>
<td>–0.086</td>
<td>0.045</td>
<td>0.053</td>
</tr>
<tr>
<td>Female</td>
<td>0.086</td>
<td>0.045</td>
<td>0.053</td>
</tr>
<tr>
<td>Body mass index</td>
<td>–0.015</td>
<td>0.005</td>
<td>0.001</td>
</tr>
<tr>
<td>Oxford Hip Score</td>
<td>0.017</td>
<td>0.003</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Model F-statistic (5 df) = 12.74, p<0.0001, R²=0.084.

### Table 7  Sensitivity analysis for main results

<table>
<thead>
<tr>
<th></th>
<th>Estimated cost/£</th>
<th>Cost/QALY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median IQR</td>
<td>Mean 95% CI*</td>
</tr>
<tr>
<td>Modelling cost entirely due to bed days (no fixed cost element)</td>
<td>4950 4240–5940</td>
<td>7220 6723 to 7748</td>
</tr>
<tr>
<td>Fixed cost element reduced by 50%</td>
<td>5012 4410–5870</td>
<td>7197 6725 to 7714</td>
</tr>
<tr>
<td>Fixed cost element reduced by 25%</td>
<td>5053 4499–5847</td>
<td>7193 6728 to 7687</td>
</tr>
<tr>
<td>Fixed cost element as used</td>
<td>5084 4588–5812</td>
<td>7182 6740 to 7678</td>
</tr>
<tr>
<td>Fixed cost element increased by 25%</td>
<td>5115 4669–5777</td>
<td>7167 6728 to 7653</td>
</tr>
<tr>
<td>Fixed cost element increased by 50%</td>
<td>5146 4758–5742</td>
<td>7160 6732 to 7641</td>
</tr>
<tr>
<td>Totally fixed cost (at 5516)</td>
<td>5516 5516–5516</td>
<td>7058 6679 to 7476</td>
</tr>
</tbody>
</table>

*Bootstrapped, bias-corrected 95% CI, based on 10 000 bootstrap replications.

QALY, quality-adjusted life year.
Based on reasonably conservative assumptions, the mean QALYs gained was 0.8 (95% CI 0.76 to 0.84), while the mean cost per hospital stay was just over £5000 per patient. Although these costs would have been more accurate if the study had been undertaken prospectively, comprehensive data allowed us to build a reasonably accurate cost profile for each patient. Most of this cost could be attributed to length of stay, although the study could not directly account for variability in the price of implant costs. Uncertainty surrounding the fixed cost data was examined using sensitivity analysis and its cost was shown to possibly increase to £5500 per case. Bootstrapping techniques increased the robustness of these findings by reducing bias by multiple replications of the primary study results.

In terms of cost per QALY, we have shown that THR may be more sensitive to optimal treatment and care in the most appropriate patient groups than to local variations in cost. But such results should be treated with caution. An actual alternative comparator implant rather than no surgery would have been more realistic, but no such data existed in this study. However, we deliberately made highly conservative assumptions both about cost and any likely net QALY gain. Furthermore, a probabilistic sensitivity analysis demonstrated that THR is likely to be good value for money even when willingness-to-pay thresholds are set quite low. This study confirms what is perhaps implicitly assumed in every day orthopaedic practice that hip replacement (using a reliable implant) is worth doing for the majority of patients. The EPOS patients had their pain, function and ultimately their QoL improved by the Exeter hip, even those with above average age, disability and BMI.

RECOMMENDATIONS

Further modelling studies are still needed to establish the longer-term cost-effectiveness of THR. The most cost-effective implants will be those with the best survival rates (and hence the fewest revisions), with the best patient outcomes and the least cost. More studies of a comparative nature incorporating economic evaluation would immensely improve the still imperfect knowledge of the cost-effectiveness of different THR implants in today’s NHS.

Acknowledgements We are grateful to Stryker UK Ltd. and in particular, Mr David Forsythe for sponsoring this study and for his comments on earlier drafts of the paper and his support throughout. We are also grateful to Professor David Murray, Oxford University, Mr John Timperley, Exeter and the members of the EPOS Group for allowing us access to the data. The following are principal investigators of the EPOS group: Prof DW Murray, Mr G Andrew, Mr J Nolae, Prof DJ Beard, Mr P Gibson, Mr A Hamer, Mr M Fordyce and Mr K Tuson. The following are or have been study coordinators for the EPOS group: A Potter, A McGovern, K Reilly, C Jenkins, K Barker, A. Cooper, C. Darrah, L. Cawton, P Inaparthy and C Pitchfork. We would also like to thank Professor Alastair Gray, Director of the Health Economics Research Centre, Division of Public Health and Primary Care, University of Oxford and Helen Dakin, Researcher, Health Economics Research Centre, University of Oxford for their critical advice on later drafts.

Contributors RF designed the conception, economic evaluation, led the analysis and wrote the first and subsequent drafts. JS undertook all the statistical processing and analysis and worked with RF and JN on the interpretation of results. XW helped revise the paper for publication. JN provided clinical input to the study throughout. JN as representative of Exeter Primary Outcomes Study (EPOS) provided approval of this paper along with other members of the EPOS Team.

Funding Stryker UK Ltd.

Competing interests RF and JS received consultancy payments for the original economic analysis. RF is currently receiving a 2-year study grant from Stryker and any likely net QALY gain. Furthermore, a probabilistic sensitivity analysis demonstrated that THR is likely to be good value for money even when willingness-to-pay thresholds are set quite low. This study confirms what is perhaps implicitly assumed in every day orthopaedic practice that hip replacement (using a reliable implant) is worth doing for the majority of patients. The EPOS patients had their pain, function and ultimately their QoL improved by the Exeter hip, even those with above average age, disability and BMI.

REFERENCES


The economic benefit of hip replacement: a 5-year follow-up of costs and outcomes in the Exeter Primary Outcomes Study

Richard Fordham, Jane Skinner, Xia Wang, John Nolan and the Exeter Primary Outcome Study Group

BMJ Open 2012 2:
doi: 10.1136/bmjopen-2011-000752

Updated information and services can be found at:
http://bmjopen.bmj.com/content/2/3/e000752

These include:

Supplementary Material
Supplementary material can be found at:
http://bmjopen.bmj.com/content/suppl/2012/05/25/bmjopen-2011-000752.DC1
http://bmjopen.bmj.com/content/suppl/2012/06/01/bmjopen-2011-000752.DC2

References
This article cites 6 articles, 1 of which you can access for free at:
http://bmjopen.bmj.com/content/2/3/e000752#BIBL

Open Access
This is an open-access article distributed under the terms of the Creative Commons Attribution Non-commercial License, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited, the use is non commercial and is otherwise in compliance with the license. See: http://creativecommons.org/licenses/by-nc/2.0/ and http://creativecommons.org/licenses/by-nc/2.0/legalcode.

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections
Health economics (325)
Surgery (356)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/