Cost-outcome analysis of joint replacement: evidence from a Spanish public hospital

José Luis Navarro Espigares/ Elisa Hernández Torres
Subdirección Económica de Control de Gestión, Hospital Universitario Virgen de Las Nieves, Departamento de Economía Internacional y de España, Facultad de Ciencias Económicas y Empresariales, Campus La Cartuja, Granada, España.

Abstract

Background and objectives: Efficiency-based healthcare decision-making has been widely accepted for some time, with cost per quality-adjusted life year (QALY) as the main outcome measure. Nevertheless, for numerous medical procedures, little data are available on the cost per QALY gained. The aim of the present study was to calculate the cost per QALY gained with primary hip and knee replacement and to compare the result with the cost per QALY for other medical procedures, as well as with the maximum threshold cost considered acceptable in Spain.

Methods: We performed a prospective cohort pre-test/post-test study of patients undergoing primary hip or knee arthroplasty. Age, sex, and clinical variables were recorded. Functional status and quality of life were measured by means of the WOMAC and EuroQol instruments, respectively, before the intervention and 6 months later. The direct costs of the intervention were calculated, with length of hospital stay and the type of surgery considered as the main cost drivers.

Results: A total of 80 patients, 40 from each intervention, were included in this study. Both functional and perceived health status improved after the intervention. The number of QALYs gained in the knee cohort was 4.64, while that in the hip cohort was 0.86. The total cost of knee replacement was lower (6,865.52 €) than that of hip replacement (7,891.21 €). The cost per QALY gained was 1,275.84 € for knee and hip interventions, respectively. The calculations performed included a 6% discount rate for health outcomes, a 3% inflation rate for costs, and a success rate of 95% at 15 years.

Conclusions: The costs of both knee and hip replacement were lower than the threshold of 30,000 € per QALY considered acceptable in Spain, and compared favorably with other medical and surgical procedures.

Key words: Cost-utility analysis. Hip arthroplasty. Knee arthroplasty. Quality of life.

Introduction

During the last 3 decades, total hip and knee replacement (arthroplasty) have become very common throughout the world, and are procedures that present low mortality rates. This type of surgery is considered an effective treatment, even
one of the most successful surgical treatments. In fact, it is generally accepted that in these processes functional benefits outweigh clinical risks and costs. However, since this is a major surgical procedure, its application is generally restricted to patients who, while showing different pathologies, suffer severe pain and functional disability, as well as in cases where other treatments have already failed.

The objective of these medical interventions is to improve the condition of functional disabilities and relieve pain caused by the deterioration of joints. A further aim is to restore the necessary mobility for patients, so they are able to maintain their functional independence and appropriate performance of daily-living activities, thereby improving their quality of life.

The efficacy of hip and knee arthroplasty has been assessed from several points of view. The most widely extended analysis, from a clinical perspective, has indicated that arthroplasty of both hip and knee prosthetic surgery is comparable to that of other medical and surgical interventions commonly implemented, e.g. bypass surgery and renal dialysis.

During the last 15 years, the number of published investigations on economic evaluation of arthroplasties, mainly those performed as cost-effectiveness analyses (CEA), has noticeably increased. Nevertheless, since the typical measurements of outcomes in CEA usually have a limited scope that is different for each type of intervention evaluated, the results of cost-effectiveness analysis, when used as an instrument for decision-making of healthcare resources allocation, present certain limitations.

In the current scenario with increasing pressures regarding healthcare costs, decision-making about healthcare resources allocation based on explicit and objective criteria has become critically important.

In accordance with previous statements, methodological guidelines for performance of pharmacoeconomic evaluations, aligned with the well-known information necessities of decision-makers, show preferences more and more oriented towards the realisation of cost-utility assessments. Nonetheless, this technique has been, up to now, the least applied and published technique of economic assessments in Spain and at the international level as well. Although the absolute number of studies using this approach has risen, the relative share of economic evaluations using QALYs or life-years gained fell from 1986 to 1996. Consequently, at present, the availability of cost-effectiveness data expressed in terms of quality adjusted life years (QALY) is still scarce.

The objective of this paper is to estimate the cost-utility of primary hip and knee arthroplasty and to assess its monetary value with general criteria of acceptability indicating an efficient cost-utility relationship.

Methods

Design

This is a partially stochastic cost-outcome description, where effectiveness data were collected by means of a prospective cohort study of patients undergoing primary, total or partial, hip or knee arthroplasty, and costs data are deterministic.

Patients and methods

A sample of 80 patients was selected by a stratified random sampling method, 40 patients of each intervention under study, at the University Hospital Virgen de las Nieves of Granada during the year 2005. Patients were selected from the waiting list for primary hip and knee replacement surgery. Exclusion criteria were revision surgery, emergencies, as well as patients lacking mental capability to give informed consent for the surgical intervention or to participate in the study plus also lacking relatives to give such legal consent. Before admission to the hospital for the intervention, selected patients were seen in consultations in which they were informed about the project and the procedures included in it. Informed consent by the patients for voluntary participation in the study was requested.

Sociodemographic (age, sex), clinical and some healthcare characteristics (duration of stay, diagnosis indicating arthroplasty according to the International Classification of Diseases, IDC-9, and medical processes carried out) were obtained from hospital records (clinical and administrative).

Functional status related to the joints under study was assessed. We measured specific characteristics of the joint operated on (pain, disability and stiffness of joint) by means of the WOMAC Index, the most commonly used tool for assessing the osteoarticular health related quality of life, especially in reference to patients with hip or knee arthrosis.

Health related quality of life was analysed with the Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36).

Finally, QALYs gained were estimated by using the EuroQol instrument. In this study we used the social Spanish Euroqol tariff associated with the time trade-off (TTO) method.
In accordance with surgeons' practice, patients were called again for consultations six months after surgery. We took advantage of these ordinary revisions to re-assess functional and health related quality of life status by using the same questionnaires.

The final impact of the intervention was assessed as pre-test and post-test changes in the specific variables under study and by means of QALYs gained after interventions.

Costs were analysed from the perspective of the hospital. Thus we have considered the hospital's direct costs incurred from interventions for hip and knee arthroplasties (stays, prostheses, drugs).

Costs data source was the hospital's Analytical Accounting System. In order to calculate the cost per patient, we considered the 2 main short-term cost drivers in hip and knee arthroplasties: stays and prostheses (table 1). Costs of prosthesis are very similar between hip and knee patients groups, but stays show high variability. Then, we calculated the average cost of each procedure considering the average stay, and a range of maximum and minimum cost with regard to the highest and lowest in stays.

Data analysis

A standard statistical analysis was carried out, by means of both univariate (descriptive analysis of variables) and bivariate (analysis of statistical significant differences between pre-test and post-test scores in the evaluated characteristics) analysis. The statistical significance level was 1% in all cases. Analyses were performed using the statistical software SPSS 12.

In addition to the statistical significance, we evaluated the changes in scale scores estimating the «effect size». The effect size standardises the measurement units so that all the various changes observed can be compared17. By using the effect size measurement it is possible to know not only whether an experiment has a statistically significant effect, but also the size of any observed effects. Within the context of a pre-test and post-test study, effect size (hypothetical effect sizes for the population) was calculated as the absolute value of the mean difference divided by the standard deviation of the differences. In this framework, .2, .5, and .8 are small, medium, and large effect sizes, respectively18.

Cost-utility analysis

Cost-utility ratios were calculated for knee and hip replacement. Since this is a cost-outcome description of hip and knee replacements, there is no comparative analysis, so incremental cost-utility ratios have not been calculated.

It must be pointed out that this work is not an economic evaluation, as it does not fulfil one of the defining characteristics of such studies: i.e. to have an implicit election11. Actually, this paper offers a description of the cost-outcome of 2 health interventions in an independent way. Although we compare their results in terms of cost-effectiveness ratios (CER), we are not performing an incremental analysis between both interventions, nor is an incremental cost-effectiveness ratio (ICER) calculated.

Discount

Since the scope of the results is considered from the moment of the replacement to 15 years later, discount rates were applied for both costs and effectiveness. Following the recommendations of the working group for standardisation of economic evaluations of health technologies in Spain19, a 6% discount rate for effectiveness and costs has been utilised. Also a 3% annual inflation rate was applied in estimating future costs.

Sensitivity analysis

We performed a two-dimensional deterministic sensitivity analysis by considering the worst (lowest effectiveness and highest costs) and the best (highest effectiveness and lowest costs) scenarios.

Results

Forty patients who received surgery for knee arthroplasty, as well as 40 patients for hip arthroplasty, were included in the study.

Demographic and clinical characteristics show statistically significant differences between hip and knee groups in relation to age, sex, stays and diagnosis. Regarding afflictions in other joints and previous surgery of other joints, there are no such differences (table 2).
The WOMAC Index of arthrosis in knee patients reported a statistically significant functional improvement after intervention. Taking each of the 3 dimensions of the WOMAC Index separately, statistically significant improvements in all of them were found, except in the area of stiffness. The SF-36 questionnaire reveals improvements in both physical and mental dimensions, statistically significant only in the former. Regarding the effect size, changes in WOMAC are "high" or "medium" in all cases. Changes in SF-36 are "medium" for the 2 aspects considered, although much higher for the physical than for the psychological dimension (table 3).

Similar results were obtained for the hip cohort, although in this case improvements were generally more moderate. The WOMAC Index for this group shows differences not considered statistically significant. Regarding the three WOMAC dimensions, as in the case of the knee cohort, statistically significant improvements were found for pain and disability, while the improvement in stiffness of joints was not statistically significant. Again, as with the knee cohort, in the assessment of the health-related quality of life, a statistically significant improvement was reported in the physical dimension. For the mental aspect, a mild decrease, while not statistically significant, was observed. The effect size in the WOMAC questionnaire ranked "high" in all cases, for dimensions and total score as well. The SF-36 showed a "medium" change in the physical aspect and a "small" one in the psychological dimension (table 3).

Costs of interventions are calculated considering the two main determinants of arthroplasties: stays and prosthesis. Knee arthroplasty costs 6,865.52 € (range, 6,426.60-7,309.43) while the total cost of hip arthroplasty is 7,891.21 € (range, 7,407.40-8,380.00) (table 1).

Assuming a survival rate of the prosthesis of 15 years in 95% of the cases, the total gains of QAL Ys were 4.64 (95% CI, 2.23-6.72) in the knee cohort and 0.86 (95% CI, 0.15-1.51) in the hip cohort, both statistically significant (p < 0.05) (table 4).

Regarding the number of QAL Ys gained and the discounted total costs for each intervention, the cost per QAL Y gained in knee arthroplasties is 1,275.84 €, and the cost per QAL Y gained in hip arthroplasty is 7,936.12 € (table 4).

Sensitivity analysis shows that changes in the effectiveness and costs considered have an important effect on the cost-utility ratio for each intervention. In the case of knee arthroplasty, the cost per QAL Y in the worst scenario is more than twice the baseline cost-utility ratio, and the cost per QAL Y in the best scenario is almost half of the baseline value (table 4).

**Discussion**

Total hip and knee arthroplasty are not new surgical procedures, but their widespread application is relatively recent. In the last decades both indications and effective age range were extended, with a reduced or

---

**Table 2. Characteristics of patients who underwent hip and knee arthroplasty**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Knee (n = 40)</th>
<th>Hip (n = 40)</th>
<th>p^a</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean SD</td>
<td>Mean SD</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>71.70 7.76</td>
<td>63.20 12.03</td>
<td>p &lt; 0.01</td>
</tr>
<tr>
<td>Stays</td>
<td>9.33 8.95</td>
<td>11.21 9.03</td>
<td>p &lt; 0.01</td>
</tr>
<tr>
<td>Sex</td>
<td>p &lt; 0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>31 77.50</td>
<td>24 60.00</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>9 22.50</td>
<td>16 40.00</td>
<td></td>
</tr>
<tr>
<td>Diagnosis</td>
<td>p &lt; 0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Located osteoarthrosis</td>
<td>31 77.50</td>
<td>16 40.00</td>
<td></td>
</tr>
<tr>
<td>Closed femur neck fracture</td>
<td>0 0.00</td>
<td>10 25.00</td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>9 22.50</td>
<td>14 35.00</td>
<td></td>
</tr>
<tr>
<td>Other joints affected (No)</td>
<td>34 85.00</td>
<td>39 97.50</td>
<td>0.28</td>
</tr>
<tr>
<td>Previous surgery of other joints (No)</td>
<td>30 75.00</td>
<td>35 87.50</td>
<td>0.28</td>
</tr>
</tbody>
</table>

^aTest t-Student.

**Table 3. Functional and health status**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Preoperative Mean SD</th>
<th>Postoperative Mean SD</th>
<th>p^a</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knee (n = 40)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WOMAC Pain</td>
<td>9.45 3.47 3.00 4.03 0.05</td>
<td>3.00 4.03 &lt; 0.05 1.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stiffness</td>
<td>3.75 1.77 3.55 2.14 0.72 0.74</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Functional capability</td>
<td>37.15 14.28 24.20 12.58 &lt; 0.05 1.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total score</td>
<td>50.35 17.79 30.75 17.12 &lt; 0.05 1.18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF 36 Physical D</td>
<td>31.57 7.01 37.13 9.32 &lt; 0.05 0.79</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental D</td>
<td>46.08 13.75 49.78 12.66 0.16 0.27</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hip (n = 40)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WOMAC Pain</td>
<td>9.36 2.73 4.50 4.50 0.05</td>
<td>4.50 &lt; 0.05 1.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stiffness</td>
<td>4.72 1.35 3.70 2.10 0.14 1.62</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Functional capability</td>
<td>40.45 7.60 24.22 19.96 &lt; 0.05 1.49</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total score</td>
<td>54.54 8.73 32.50 22.54 0.71 1.79</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF 36 Physical D</td>
<td>32.76 10.05 36.10 10.64 &lt; 0.05 0.33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental D</td>
<td>46.72 14.40 45.81 16.33 0.10 0.06</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^aTest t-Student.

SD: standard deviation.

The WOMAC Index of arthrosis in knee patients reported a statistically significant functional improvement after intervention. Taking each of the 3 dimensions of the WOMAC Index separately, statistically significant improvements in all of them were found, except in the area of stiffness. The SF-36 questionnaire reveals improvements in both physical and mental dimensions, statistically significant only in the former. Regarding the effect size, changes in WOMAC are "high" or "medium" in all cases. Changes in SF-36 are "medium" for the 2 aspects considered, although much higher for the physical than for the psychological dimension (table 3).
null mortality rate and good clinical outcomes. The results realised in this work describe positive outcomes for both interventions, from an objective clinical point of view as well as from the health perceived quality of life expressed by the patients. In other words, the two perspectives were similar, although those aspects did not always converge. These results bear out those obtained in previous studies. Nevertheless, due to the variety of instruments used and the different periods of time under study in each investigation, the results can hardly be considered comparable regarding effect size. At any rate, the effect size measurement allows us to observe the direction of certain trends.

Cost-utility analysis, although regarded as the most suitable economic evaluation method for healthcare resources allocation, still represents a very small percentage of the economic evaluation published in healthcare technologies in general and in particular in the field of orthopaedic surgery. Reviews including economic evaluations of both knee and hip arthroplasty corroborate this observation. Thus, a review of economic evaluations of knee arthroplasty published between 1966 and 1996 found that only 2.5% of the reviewed papers were cost utility analyses, the majority being cost-minimisation analyses. On the other hand, reviews including economic evaluations of hip arthroplasty carried out between 1966-1996 and 1966-2002 as well, found that 5.88% of the former and 16% of the latter were cost-utility analyses, the majority being cost-effectiveness analyses. Furthermore, even though the developers of the EuroQol recommend it as a complementary instrument, there are only a few studies on the outcome of arthroplasty in which the EuroQol was employed.

Direct costs and cost per QALY of hip and knee arthroplasties are lower than previously reported costs estimated for other countries. But our present costs, compared with results of a prior Spanish study, in the case of total hip arthroplasty are higher than before. This difference is mainly due more to the difference in QALYs gained between both studies (45% higher in the former) than to the difference in costs (16% higher in our study), although differences in the periods of time considered make direct comparisons of outcomes inadvisable.

A great difference of cost per QALY gained is observed between hip and knee interventions. As in other studies, patients suffering knee dysfunctions report worse pre-operative results in quality of life than patients with hip dysfunctions, but 6 months later, knee patients report better results than do patients of hip surgery. Besides the higher cost per process of hip arthroplasty, we obtained a more favourable cost per QALY relationship for knee arthroplasties than for hip. This result is the opposite of that obtained by Rääsänen et al (2007), where total hip arthroplasty emerged as more cost-effective than total knee arthroplasty with a cost per QALY of 4,517 and 9,421 €, respectively. Alternatively, Chang, Pellissier, and Hazen (1996) report a cost per QALY of 4,637 € and Segal et al (2004) suggest a cost per QALY of 3,639 € for knee replacement.

In all the previous analyses, the time horizon considered up to the post-test assessment was 1 year, against 6 months considered in the present work. This is the main methodological difference that we have found between our study and those previously published. In our opinion this circumstance could explain the higher cost per QALY calculated for hip arthroplasty.

---

**Table 4. Costs per QALY gained of hip and knee arthroplasties and sensitivity analysis**

<table>
<thead>
<tr>
<th>Joint</th>
<th>Scenario</th>
<th>Stays</th>
<th>Social Tariff</th>
<th>Euroqol</th>
<th>QALYs gained</th>
<th>Cost per process (€)</th>
<th>Cost per QALY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knee</td>
<td>Average</td>
<td>9.33</td>
<td>0.20</td>
<td>0.64</td>
<td>4.64</td>
<td>5,921,71</td>
<td>1,275,84</td>
</tr>
<tr>
<td></td>
<td>Best</td>
<td>8.44</td>
<td></td>
<td>6.72</td>
<td>5,543,13</td>
<td>824,87</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Worst</td>
<td>10.21</td>
<td>0.35</td>
<td>2.23</td>
<td>6,304,59</td>
<td>2,827,17</td>
<td></td>
</tr>
<tr>
<td>Hip</td>
<td>Average</td>
<td>11.21</td>
<td>0.47</td>
<td>0.55</td>
<td>0.86</td>
<td>6,806,40</td>
<td>7,936,12</td>
</tr>
<tr>
<td></td>
<td>Best</td>
<td>10.24</td>
<td></td>
<td>1.51</td>
<td>6,389,10</td>
<td>4,231,19</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Worst</td>
<td>12.19</td>
<td>0.15</td>
<td>0.15</td>
<td>7,228,00</td>
<td>48,186,64</td>
<td></td>
</tr>
</tbody>
</table>

SD: standard deviation.

\(^{a}\)QALYs gained assuming a survival rate of the prosthesis of 15 years in 95% of the cases and a 6% discount rate: \( \frac{\sum QALYs \text{ gained}}{(1 + r)^n} \times 95\% \)

\(^{b}\)Total cost assuming an inflation rate of 3%: \( \frac{\sum \text{Cost}/n}{(1 + i)^n} \)
Table 5. Cost-utility ratios obtained in different context

<table>
<thead>
<tr>
<th>Disease</th>
<th>Cost (€, 2007)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CER Knee arthroplasty (Min)</td>
<td>824.87</td>
</tr>
<tr>
<td>CER Knee arthroplasty (Av)</td>
<td>1,275.87</td>
</tr>
<tr>
<td>CER Knee arthroplasty (Max)</td>
<td>2,827.17</td>
</tr>
<tr>
<td>CER Hip arthroplasty (Min)</td>
<td>4,231.19</td>
</tr>
<tr>
<td>Higher recommended Spain (hepatitis treatment)</td>
<td>6,783.07</td>
</tr>
<tr>
<td>CER hip arthroplasty (Av)</td>
<td>7,396.12</td>
</tr>
<tr>
<td>Critical care</td>
<td>19,756.55</td>
</tr>
<tr>
<td>Congenital anomalies</td>
<td>25,379.13</td>
</tr>
<tr>
<td>Genito-urinary diseases</td>
<td>28,525.71</td>
</tr>
<tr>
<td>Spanish threshold</td>
<td>30,000.00</td>
</tr>
<tr>
<td>CER hip arthroplasty (Max)</td>
<td>48,186.64</td>
</tr>
<tr>
<td>International threshold</td>
<td>50,000.00</td>
</tr>
<tr>
<td>Injuries/exposures</td>
<td>66,265.79</td>
</tr>
<tr>
<td>Digestive diseases</td>
<td>89,348.43</td>
</tr>
<tr>
<td>Cardiovascular diseases</td>
<td>92,629.31</td>
</tr>
<tr>
<td>Malignant neoplasms</td>
<td>152,652.84</td>
</tr>
<tr>
<td>Anemias</td>
<td>153,988.48</td>
</tr>
<tr>
<td>Allergy/immunology</td>
<td>214,824.95</td>
</tr>
<tr>
<td>Infectious diseases</td>
<td>649,038.17</td>
</tr>
<tr>
<td>Hematology-non cancer</td>
<td>3,621,573.48</td>
</tr>
</tbody>
</table>

Source: Sacristán et al.12


From an extended point of view, comparing the analysed interventions with «doing nothing» (zero outcomes and costs), we can consider the cost-effectiveness ratios as incremental cost effectiveness ratios (ICER). With this assumption, the processes under investigation are well positioned in cost-utility acceptability values regarding the generally accepted cost-effectiveness threshold (30,000 € per QALY gained in Spain32 and around 50,000 € per QALY gained accepted in other contexts33,34). If we take as comparative environment other technologies considered to be cost-effective in Spain as point of reference, the range of acceptability is between 57.76 € and 6,783.07 €32 (2007 euros) and thus knee arthroplasty can be considered cost-effective whereas hip arthroplasty exceeds this range. The sensitivity analysis on extreme scenarios shows that, in the worst possible case, the cost-utility ratio for hip arthroplasties would exceed the Spanish efficiency threshold, but remain within the boundaries of international acceptability. From the global perspective, the cost-utility ratios published from 1976 to 200135 in general report higher cost per QALY gained than the ones obtained in our study for several diseases (table 5).

Previous results should be considered with certain limitations taken into account. First of all, the small sample size together with the analysis of only one hospital greatly reduce the possibility of generalising the results obtained. Secondly, we do not consider the long term costs of hip and knee arthroplasties, e.g. revision surgery costs. This fact can lead to an undervaluation of cost per QALY gained with the analysed interventions.

It is worth mentioning that comparison by means of league tables, at national and international levels, has a limited scope when the structure of costs is not known. Still, in order to facilitate comparisons between cross-national and cross-care interventions, we do agree with the advice of guidelines encouraging researchers to report the results of economic evaluations of healthcare technologies in QALYs.

References


