LONG-TERM COMPARISON OF CHARNLEY AND STANMORE DESIGN TOTAL HIP REPLACEMENTS

A. R. BRITTON, D. W. MURRAY, C. J. BULSTRODE, K. McPHERSON, R. A. DENHAM
From the Nuffield Orthopaedic Centre, Oxford and the London School of Hygiene and Tropical Medicine, England

We reviewed the records of the long-term outcome of 208 Charnley and 982 Stanmore total hip replacements (THR) performed by or under the supervision of one surgeon from 1973 to 1987. The Stanmore implant had a better survival rate before revision at 14 years (86% to 79%, p = 0.004), but the difference only became apparent at ten years.

The later Stanmore implants did better than the early ones (97% to 92% at ten years, p = 0.005), the improvement coinciding with the introduction of a new cementing technique using a gun. Most of the Charnley implants were done before most of the Stanmore implants so that the difference between the results may in part be explained by improved methods, but this is not the complete explanation since a difference persisted for implants carried out during the same period of time.

We conclude that improved techniques have reduced failure rates substantially. This improvement was much greater than that observed between these two designs of implant. Proof of the difference would require a very large randomised controlled trial over a ten-year period.

Received 3 January 1996; Accepted after revision 20 March 1996

To choose an implant, surgeons need published information about the clinical results of different joint replacements. Long-term studies are necessary; some designs have good results at five years but develop high rates of failure by ten years (Malchau, Herberts and Ahnfelt 1993; Owen et al 1994; Savilahti et al 1995). A recent review of available total hip replacements in the UK by Murray, Carr and Bulstrode (1995) showed that only nine had published reports of results for over five years.

There are also few long-term comparisons of two or more prostheses. The Swedish multicentre study (Malchau et al 1993) provided a ten-year comparison of a number of implants of different design; this included data from many surgeons and operating centres, which may introduce confounding factors. Differences between the results may be due to the experience of different surgeons. Revision was the only measured outcome; this may be unreliable since surgeons differ in their indications for revision.

We aimed to achieve a long-term comparison between two implants used by the same surgeon over a period of 16 years.

PATIENTS AND METHODS

From 1973 to 1987, a total of 1190 patients had a THR (208 Charnley, 982 Stanmore) by or under the direct supervision of one of the authors (RAD), who was also personally responsible for most of the reviews. A previously designed standard protocol was completed at six months, one year and then subsequently every two years. The assessment included grading of pain and recording any complications or revisions.

Eighty-one were revised (7%); the major reasons for revision were aseptic loosening (47%) and septic loosening (27%). There were no significant differences in the cause of failure for different implants (chi-squared test = 0.14, p > 0.5).

Because of the differing lengths of follow-up we used survival analysis, defining the endpoint in five different ways using level of pain and revision (Table I). We used the life-table method (Dobbs 1980; Carr et al 1993), stopping when the number of surviving implants fell below ten (Murray, Carr and Bulstrode 1993). We calculated the 95% confidence limits (CL) (Peto et al 1977) and compared the curves by the log-rank method (Peto et al 1977).
Some Stanmore implants were used throughout the period of study, although most (72%) were implanted between 1977 and 1982. Of the Charnley design, 88% were employed between 1974 and 1976, before the first use of a cement gun in 1979 (Fig. 1). Because of the possible confounding effect of the change in technique associated with the use of a cement gun, we made a separate analysis of the first 560 Stanmore hips (1973 to 1979) which were performed without second-generation techniques and a cement gun, and then the remaining 422 (late 1979 to 1986) operations in which the newer procedures were used.

To test the influence of surgical experience on the outcome, we subdivided the groups into four cohorts: 280 implanted from 1973 to 1977, 280 from 1977 to 1979, 211 inserted from 1979 to 1981, and 211 from 1981 to 1986 (Fig. 1).

All the operations had been carried out through a posterior (Southern) approach with no trochanteric osteotomy.

### RESULTS

The median follow-up was for eight years, with 40 hips reviewed for the full 16 years. Of the 1190 patients 279 (24%) had died and 77 (6%) had been lost to follow-up. There was no significant difference in the proportion lost to follow-up between the Stanmore and the Charnley groups (Fig. 2).

Tables II and III show the life tables for each implant at 6, 10, 14 and 16 years for each of the five endpoints. The survival rates for endpoints 1 (revision) and 2 (severe pain) were identical, which indicate that all patients who developed severe pain reported it between their biannual reviews and had a revision operation between assessments.

The survival rate for the Charnley implant at ten years \((n = 107)\) was 84\% (CL ± 6.3) for endpoint 1 (revision only), but for endpoint 5 (the onset of slight pain) this was reduced to 44\% (± 8.7). The survival rate for the Stanmore implant for revision at ten years \((n = 332.5)\) was 93\% (± 2.6), and for onset of slight pain was 48\% (± 4.9).

The survivorship curves for the Charnley and Stanmore implants are similar up to eight years for all endpoints (Figs 3 to 6). For longer follow-up, the Charnley results become significantly worse for endpoints 1 and 2 (\(p = 0.004\)), 3 (\(p = 0.003\)) and 4 (\(p = 0.026\)). We found no significant difference between the implants for endpoint 5 (onset of slight pain).

Comparison of the four time cohorts of Stanmore implants (Table IV and Fig. 7) showed no significant difference between the first 280 and second 280 operations,
or between the third and fourth cohorts, but there was a difference between the second and third (p = 0.009), and between the first two (no gun, n = 560) and the later two (second-generation with gun, n = 422). The ten-year survival without revision was 91.6% compared with 97.4% (p = 0.005; Fig. 8).

Comparison between the Charnley implants and the first 560 Stanmore implants, all performed before the use of a cement gun, still showed a significant difference in survival without revision between the two implants (p = 0.015), but this was less than when all the Stanmores were included. A comparison between the first 280 Stanmore hips and all the Charnley joints also showed a difference (86.3% v 79.1%) but this failed to reach statistical significance (p = 0.07).

Using endpoint 4 (moderate pain) as the outcome rather than revision we found very similar results, with no difference between the first and second time cohorts, or the third and fourth, but a significant difference between the second and third (p = 0.0001). Although cohorts 3 and 4 of the Stanmore hips together gave significantly higher results
than the Charnley implants ($p = 0.037$), neither cohorts 1 and 2 combined nor cohort 1 alone showed a significant difference between the two types ($p = 0.315$ and $p = 0.489$, respectively).

**DISCUSSION**

Our long-term prospectively planned study has compared the outcome of large numbers of Stanmore and Charnley THRs. The overall analysis showed significantly better results for the Stanmore implant when the definition of failure was revision or the onset of severe or moderate pain. The study, however, was not randomised, and the influence of possible confounding factors needs to be considered before any firm conclusions can be drawn.

Probably the most important confounding factor is that most of the Stanmore implants were inserted later than the Charnley hips. This is important because the more recent Stanmore THRs had better outcomes than the early ones. Many factors may be involved including improved surgical techniques, increasing surgical experience, and altered patient selection. Although most implants, including the
Charnley, have been modified over the years, the Stanmore design had not been changed and design modification is therefore not a factor.

If increasing surgical experience had been responsible for the improvement in results then a steady improvement would have been expected with each of the time cohorts of Stanmore implants. Our analysis showed, however, that the outcome of the first 280 was similar to that of the next 280, as were the third and fourth cohorts. This finding reduces the likelihood that increased surgical experience was responsible. A highly significant difference ($p = 0.009$) was found between the second and third cohorts which coincided with the introduction of a cement gun. This suggests that the changes in cementing techniques associated with the introduction of the cement gun were responsible for the improvement (Paterson, Fulford and Denham 1986). This is substantial, amounting to a 70% decrease in the revision rate at ten years.

Our comparison of the two earlier cohorts of Stanmore implants (560) with all the Charnley hips (208), most carried out before the introduction of a cement gun, still showed a highly significant difference between the implants at ten years or more ($p = 0.001$). Restriction to the first 280 Stanmore operations and all the Charnley opera-
Table IV. Confidence limits for the four time cohorts of Stanmore operations

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Number of operations</th>
<th>Number at risk at ten years</th>
<th>10-year survival (%)</th>
<th>95% Confidence limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>280</td>
<td>122</td>
<td>91.6</td>
<td>86.8 to 96.3</td>
</tr>
<tr>
<td>2nd</td>
<td>280</td>
<td>121.5</td>
<td>91.6</td>
<td>86.9 to 96.4</td>
</tr>
<tr>
<td>3rd</td>
<td>211</td>
<td>31.5</td>
<td>98.5</td>
<td>94.2 to 100</td>
</tr>
<tr>
<td>4th</td>
<td>211</td>
<td>4.5</td>
<td>96.6</td>
<td>84.7 to 100</td>
</tr>
</tbody>
</table>

Fig. 7
The four cohorts of Stanmore implants.

Fig. 8
Stanmore THR pre- and postcement gun compared with Charnley pregun.
tions, most of which were performed in the same period, showed a difference which did not quite reach significance \((p = 0.07)\). Using endpoint 4 (moderate pain) it was shown that the later (postgun) Stanmore implants did significantly better than the Charnley implants but that pregun survival rates were similar. This whole analysis suggests that under comparable circumstances the Stanmore implant has a lower revision rate than the Charnley, and that the two types of implant function equally well as regards pain until they require revision.

We found no difference between the two prostheses until ten years of follow-up. This demonstrates that in this field ten-year observational studies are essential. To establish a real difference between implants a randomised controlled trial (RCT) is necessary, but the implementation of such a trial is difficult for many reasons including problems with surgical technique, blind assessment, and consent (Dorey, Grigoris and Amstutz 1994). Using modern cementing techniques the difference between the implants would be so small that such an RCT would require about 4000 patients and a ten-year follow up \((\alpha = 0.05, \beta = 0.1)\).

Our study is large and long-term, but has been weakened by loss to follow-up (Britton et al 1995). The 6% loss is relatively low and a worst-case curve for the Stanmore implant gives a survival without revision of 84.8% at ten years. It is more important that the loss to follow-up was not significantly different between the two implants \((p = 0.75)\), which suggests that this was unlikely to cause bias.

The survival rates for revision of both Charnley and Stanmore implants are comparable with other published studies (Malchau et al 1993; Schulte et al 1993; Alsema, Deutman and Mulder 1994), but our review of other recorded endpoints has shown that the result of a THR is perhaps not quite as good as such studies seem to indicate. The use of endpoint 5 (slight pain) shows a failure rate of about 50% within ten years. Even this may be an underestimate: the data were obtained at consulting interviews and it is known that patients do not always accurately report their pain to consultants (Morris 1993). There is clear room for improvement, since developments should aim at not only reducing failure rates but also the incidence of pain. Accurate reports of how an implant is functioning should preferably include, in addition to the revision rates, the development of various levels of pain, which can also be used in survival analysis. Information about pain should be obtained prospectively, using a well-validated patient-based questionnaire, such as that described by Dawson et al (1996).

We conclude that improved surgical techniques, associated with the use of the cement gun, have resulted in a 70% decrease in revision rate. We have also shown that in similar circumstances the Stanmore design appears to give better long-term results than the Charnley. The difference was less than that produced by an improvement in cementing technique and would require a large RCT to establish it. The difference between implants and between techniques only became clearly apparent after ten years of follow-up and reinforces the need for more long-term studies.

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


