The continued value of clinical and radiological surveillance

THE CHARNLEY ELITE PLUS HIP REPLACEMENT SYSTEM AT 12 YEARS


From Norfolk & Norwich University Hospital, Norwich, England

We report the follow-up at 12 years of the use of the Elite Plus total hip replacement (THR). We have previously reported the results at a mean of 6.4 years. Of the 217 patients (234 THRs), 83 had died and nine had been lost to follow-up. The patients were reviewed radiologically and clinically using the Oxford hip score.

Of the 234 THRs, 19 (8.1%) had required a revision by the final follow-up in all but one for aseptic loosening. Survivorship analysis for revision showed a survival of 93.9% (95% confidence interval (CI) 89.2 to 96.5) at ten years, and of 88.0% (95% CI 81.8 to 92.3) at 12 years. At the final follow-up survival analysis showed that 37% (95% CI 37.3 to 44.7) of the prostheses had either failed radiologically or had been revised. Patients with a radiologically loose femoral component had a significantly poorer Oxford hip score than those with a well-fixed component (p = 0.03). Radiological loosening at 6.4 years was predictive of failure at 12 years.

Medium-term radiographs and clinical scores should be included in the surveillance of THR to give an early indication of the performance of specific implants.

The Elite Plus total hip replacement (THR) system (DePuy, Warsaw, Indiana) was introduced in 1994 as the second modular evolution of the original Charnley low-friction arthroplasty.1 It is a cemented prosthesis with a modular, stainless-steel (Ortron 90, DePuy) femoral component and an ultra-high molecular-weight polyethylene (UHMWPE) acetabular component. The standard head is also made of stainless steel. Several changes were made to the shape and the dimensions of the femoral component. These were based on finite-element analysis with the aims of improved proximal load transfer and to decreasing contact stresses2 in an attempt to reduce subsidence and improve rotational stability. The design also incorporated an undercutting of the neck flange and a vaquasheen finish, intended to improve compression of cement at implantation and optimise the proximal cement-prosthesis interface.

Up until now, outcome data on the Elite Plus implant have been mixed and the survival rate of this prosthesis has yet to be determined conclusively. Mid-term outcome studies have shown differing radiological outcomes,1,3 as have roentgen stereophotogrammetric analysis studies.4,5 Long-term results are also inconsistent, with survival at ten years ranging from 99% (95% confidence interval (CI) 95 to 100)6 to 83% (95% CI 73 to 93).7 A survivorship of 85.2% to 91.5% at ten years has been reported in data from the Swedish Hip Registry.5,8

Our aim was to evaluate the results of the Elite Plus prosthesis at 12 years using revision as a primary endpoint, and to determine secondary outcome measures of clinical and radiological failure. The predictive value of mid-term radiographs was also assessed. This series was previously reported at 6.4 years and the need for radiological surveillance of a new prosthesis to evaluate performance was highlighted.1

Patients and Methods

A series of 234 consecutive, non-selected primary THRs was performed in 1995 on 217 patients in a single centre. All the patients received a cemented Elite Plus THR with a 28 mm Ortron 90 head. The patients were recruited at the time of surgery and prospectively followed up. The final review took place at a minimum of 12 years (12 to 13) after surgery. All the data regarding the surgical teams involved, the grade of the surgeon, the surgical approach and the type of cement used were evaluated. Follow-up of the series was reviewed by the Local Research Ethics Committee and ethical approval was obtained.

Of the 217 patients (234 THRs), 83 had died with surviving prostheses before the final
review and nine were lost to follow-up with no radiographs available. A total of 19 patients had been revised in 142 surviving hips (133 patients) available at the time of the final follow-up. All the patients were included in the survivorship analyses. Local efforts to contact patients previously lost to follow-up and the use of the national tracing service resulted in a higher rate of follow-up than that in our previously published report.1

Of the 234 THRs, 137 had been performed by consultants and 97 by registrars. All but three procedures had been undertaken through a modified anterolateral approach. A standardised second-generation cementing technique was used by all the surgical teams.

All the patients had undergone standardised postoperative surveillance including a radiological and a clinical review immediately after the operation and at one, five and ten years post-operatively. Anteroposterior pelvic radiographs were obtained in the standardised manner with the patients positioned supine. The anterior superior iliac spines were equidistant from the table top, the femora were internally rotated and the heels separated. The film and beam were centred 2.5 cm above the symphysis pubis with a focal film distance of 115 cm and the beam collimated to include the proximal femora. The radiographs were assessed by two independent observers (BO, CD) according to the criteria proposed by Johnston et al.9 They were reported in conjunction with the post-operative and mid-term (mean 6.4 years) radiographs. Agreement between both observers was necessary for data to be included and, when disagreement occurred, a consensus view was obtained with a third observer. A validation exercise was performed to determine the intra- and inter-observer reliability.

The radiographs were classified as well-fixed, possibly loose or definitely loose according to previously published criteria. Loosening of the femoral component was defined according to the criteria of Harris, McCarthy and O’Neill.10 The refined criteria of Kobayashi et al11 for loosening of the femoral component were separately analysed. Loosening of the acetabular component was defined according to the criteria of Hodgkinson, Shelley and Wroblewski.12 The survival of the prosthesis was calculated using these definitions and assessed at six and 12 years. We reviewed the results of the mid-term radiograph obtained at 6.4 years to determine the predictive value of early loosening for revision at 12 years.

The patients were reviewed clinically using the Oxford hip score (OHS).13 A poor clinical result was defined as patients outside the 95% confidence limits for their expected post-operative OHS as described by Field, Cronin and Singh.14 The number of patients in each group with a poor outcome was compared with the whole series.

Statistical analysis. This was performed using Fisher’s exact test for categorical data and a two tailed t-test for continuous data. Outcome data were presented as actual survival figures and survival analysis. Survivorship analysis was calculated using the Kaplan-Meier method with the 95% confidence interval. Survival curves were tested for difference using the Mantel-Cox (log-rank) test. A p-value ≤ 0.05 was defined as significant in all cases.

Results

Revision. Of the 234 THRs, 19 (8.1%) required revision during the 12-year follow-up period. One patient was revised for early deep infection and the remaining 18 for aseptic loosening. There were no revisions for dislocation and both components were revised in all but one case. Kaplan-Meier survivorship analysis showed survival of 93.9% (95% CI 89.2 to 96.5) at ten years and of 88.0% (95% CI 81.8 to 92.3) at 12 years for all causes (Fig. 1).

A subset analysis excluding all patients with identifiable risk factors for poor prognosis (low viscosity cement, poor
cement mantle or unacceptable component position) as previously reported\(^1\) gave a survival of 93.3\% (95\% CI 88.1 to 96.2) at ten years (Fig. 2). There was no statistically significant difference between this subset and the whole group (p = 0.98).

**Radiological.** A retrospective review of the immediate post-operative radiographs confirmed that 158 (99.4\%) THRs had a grading of A or B according to the criteria of Barrack, Mulroy and Harris\(^{15}\) and one (0.6\%) of C or D. The position of the femoral component was neutral in 111 (69.8\%), while 30 (18.9\%) femoral components were in > 4° of varus and 18 (11.3\%) in > 4° of valgus.\(^1\)

Complete radiological analysis was possible in 142 THRs at 12 years, and in a further 19 which had been revised. Definite evidence of femoral loosening was present in 41 of 142 femora (28.9\%) and in 66 of 142 acetabula (46.5\%); 82 of 142 THRs (57.7\%) had definite loosening of at least one component. At the time of the final review, 19 patients had undergone a revision giving a combined failure rate of 63\% (101 of 161) (Table I). The validation exercise demonstrated that the intra-observer reliability was excellent (κ = 0.94) and interobserver reliability good (κ = 0.89).

**Clinical findings.** Oxford hip scores\(^{13}\) were available for 111 of the 133 surviving patients in the series. The mean OHS was 24 (SD 11.5) at 12 years (Table II). Patients with a radiologically loose femoral component had a significantly poorer outcome (mean OHS 25.7 (SD 11.1)) compared with those with a well-fixed stem (p = 0.03). Despite a poorer mean score there was, however, no significant correlation between a loose acetabular component and a poor clinical outcome.

**Prediction of failure.** The predictive values and sensitivity and specificity for the mid-term radiographs taken at 6.4 years are given in Table III. Definite evidence of loosening on a mid-term radiograph had a positive predictive value of 0.33 and a negative predictive value of 0.89. The sensitivity was 0.15 and the specificity 0.96.

**Discussion**

Our results show that the Elite Plus THR did not perform as well as the original Charnley version. We have previously reported the mid-term results of this series, with 52 of 168 (31\%) Elite Plus hips either being revised or with evidence of definite loosening at six years.\(^1\) The concerns surrounding the longevity of this prosthesis suggested by our initial report appear to have been confirmed by a survivorship of 93\% at ten years, with all causes of revision as the endpoint. Although this is above the standard of 90\% defined by the National Institute for Clinical Excellence performance is below that of other generations of Charnley
stem,\textsuperscript{16} and other well-recognised prostheses\textsuperscript{17,18} when taking revision of either component as an endpoint.

Only the Swedish Hip Register has reported results for the Elite Plus prosthesis at ten years, giving a survival of 85.2\% to 91.5\% with various combinations of acetabular component and using revision for any cause as an endpoint. This poorer survival might be accounted for by the low rates of infection and revision for dislocation in our series.

In a similar series Hauptfleisch et al\textsuperscript{7} reported a radiological survival rate of 59\% at ten years and of 83\% when taking revision as the endpoint. However, only 44 of 118 femoral components (37\%) were surviving and available for radiological follow-up at a mean of ten years. By contrast, Kim et al\textsuperscript{6} initially reported a survival of 99\% at ten years, although this was later revised downwards to 97.4\% in a corrigendum.\textsuperscript{19} It is conceivable that the use of a 22.25 mm modular head in their series,\textsuperscript{6,19} with a lower rate of polyethylene wear, may explain this disparity.

It has previously been argued that good surgical technique\textsuperscript{2} and the use of high-viscosity cement\textsuperscript{6} may account for the better results with this femoral component reported by some groups. In order to address these issues, we performed a subgroup analysis excluding all ten patients in whom low-viscosity cement had been used or who had a Barrack grade C or D\textsuperscript{15} cement mantle or other radiologically identifiable poor prognostic factors. Surprisingly, after these exclusions the analysis showed a poorer survival at ten years than that of the whole series (Fig. 2), although the difference was not statistically significant (p = 0.98). Objective evaluation of our cement mantles and component positions\textsuperscript{1} using the post-operative radiographs was similar to that reported in other series.\textsuperscript{20} We have previously reported no association between Barrack grade, side, distal cement leakage, position of the stem, surgical team or grade of surgeon and outcome at six years\textsuperscript{1} and our results at 12 years support this finding. Additionally, our series had no revisions for dislocation, and only one case of deep infection which gave rates of complications well below the accepted level. It is generally agreed that the use of more modern cementation techniques increases the likelihood of a good cement mantle and thereby outcome.\textsuperscript{21} However, it is unlikely that surgical technique alone explains the early failure in this series.

It is more likely that the disparity in the reported survivorship is reflective of local policy on revision which is traditionally used as an endpoint for reporting the success of a surgical implant with 95\% CIs given around this point. Revision is a firm, but coarse endpoint and the decision to revise a joint is based on many factors including the patient’s wishes and health status. As patients become older their physical demands are probably lower and it is also probable that they are less likely to request or present for revision surgery. This makes objective reporting of results essential, particularly in long-term follow-up studies. When definite radiological loosening and revision were defined as failures, the failure rate at 12 years remained alarmingly high with a survival of 59 of 161 (37\%; 95\% CI 31 to 44). The true number of failed prostheses realistically falls between the observed radiological failures and the number of revisions performed. This may render traditional CIs of less value. Since the true number of failures lies somewhere in between, it may be of more value to plot survival analysis of the radiological failures as the lower confidence limit because the true number of loose femoral components lies between this line and those revised (Fig. 3).

In our series the 6.4 year (mid-term) radiograph showed a high negative predictive value and high specificity making it an extremely good screening test. However, with a moderate positive predictive value and a low sensitivity it would be inappropriate for patients always to be followed up by radiological means alone, and ongoing clinical scoring or review appears to be essential for the detection of failure.

We have established a significant link between a poor clinical outcome and radiological loosening of the femoral component (p = 0.03). Loose components are usually painful, and increasing pain in the light of radiological loosening is a widely recognised indication for revision. Of note was the fact that there was no significant link between radiological loosening of the acetabular component and a poor clinical outcome (p > 0.05). This may be accounted for by either poor sensitivity of the OHS for loose acetabular components or a poor specificity of the criteria of Hodgkinson et al.\textsuperscript{12}

The finding of poorer functional outcome in radiologically loose femoral components adds some weight to
the argument that radiological changes represent symptomatic loosening of the stem. These findings are in contrast to our previously published results at 6.4 years. This disparity may be accounted for by the longer follow-up period, assuming that as a femoral component becomes progressively looser, it will become more painful and functional scores will deteriorate accordingly. In the follow-up of patients with a THR, patient-reported outcome measures such as the OHS are increasingly used. These scores are designed to monitor the functional outcome after THR. However, no score has previously been demonstrated to be sensitive for loosening of a component. We have shown that the OHS is sensitive to such loosening. While our findings support the inclusion of a patient-reported outcome measure in the follow-up of THR we were unable to show a link between a poor score and radiological acetabular loosening. We would therefore advise caution in the use of patient-reported outcomes in isolation.

The reporting of a successful outcome in THR is fraught with difficulty. The outcome measures selected may have as much of an influence on the results as the data itself. A realistic picture of the performance of an implant cannot be given by the single outcome measure of revision surgery. This supports the argument that, for a local or national joint registry to be accurate in recording the performance of prostheses, revision, clinical data and radiological data should all be collected.

The use of national and local joint registries is becoming the method of choice in the surveillance of joint replacements and in the evaluation of their performance. All current national joint registries use the endpoint of revision, which does not allow for the detection of early failure in new or modified designs of prosthesis. We have shown that radiological and clinical data at the mid-term are predictive of failure. There certainly remains a place for reporting independent series including early clinical and radiological data to aid in the evaluation of prostheses which do not have long-term survival data.

Our results show that the Elite Plus THR is, by some measure, successful at 12 years. However, the high rate of radiological loosening and correlation with a poor clinical outcome continue to be a cause for concern, and therefore caution in its use must be advocated.

References