A randomised, controlled trial of two mobile-bearing total ankle replacements

P. L. R. Wood, C. Sutton, V. Mishra, R. Suneja

From Wrightington Hospital, Wigan, England

We describe the results of a randomised, prospective study of 200 ankle replacements carried out between March 2000 and July 2003 at a single centre to compare the Buechel-Pappas (BP) and the Scandinavian Total Ankle Replacement (STAR) implant with a minimum follow-up of 36 months. The two prostheses were similar in design consisting of three components with a meniscal polyethylene bearing which was highly congruent on its planar tibial surface and on its curved talar surface. However, the designs were markedly different with respect to the geometry of the articular surface of the talus and its overall shape.

A total of 16 ankles (18%) was revised, of which 12 were from the BP group and four of the STAR group. The six-year survivorship of the BP design was 79% (95% confidence interval (CI) 63.4 to 88.5 and of the STAR 95% (95% CI 87.2 to 98.1). The difference did not reach statistical significance (p = 0.09). However, varus or valgus deformity before surgery did have a significant effect (p = 0.02) on survivorship in both groups, with the likelihood of revision being directly proportional to the size of the angular deformity. Our findings support previous studies which suggested that total ankle replacement should be undertaken with extreme caution in the presence of marked varus or valgus deformity.

The Buechel-Pappas ankle replacement (BP, Endotec, New Jersey and Wright Cremasco, Toulon-Cedex, France) (Fig. 1) and the Scandinavian Total Ankle Replacement (STAR, Waldemar Link, Hamburg, Germany) (Fig. 2) are two designs of mobile-bearing total ankle replacement (TAR). Descriptions of both designs and the results of their clinical use have been well reported in the literature, but a direct comparison of their outcome has not previously been undertaken. Therefore we carried out a randomised, prospective study to compare these implants with respect to their survival and to investigate the influence of pre-operative varus or valgus deformity on survival and on subluxation of the bearing, also called edge-loading. The third objective was to determine whether TAR restored the talus, which had been anteriorly subluxed before surgery, to an anatomical position.

Patients and Methods

The study was granted approval by the hospital local ethics committee. Between March 2000 and July 2003, all patients requiring primary TAR who had not had replacement of the opposite ankle were asked to participate. Those who had already had a TAR were excluded since it was anticipated that they might express a preference to receive the same type of implant for the second ankle. This also meant that patients in the study were not recruited for the second side if that also required replacement. Therefore the number of ankles was equal to the number of patients. Randomisation was performed using a closed envelope technique on the day prior to surgery so that 100 STAR prostheses and 100 BP implants were used.

During the study the composition of the tibial and talar components of the BP was changed from nitried titanium to cobalt chrome when production was moved from the United States to Europe. This occurred approximately half way through the study. We were able to negotiate that titanium components continued to be provided for the first 50 BP ankles. The second 50 were cobalt-chrome devices.

All the replacements were carried out by, or under the direct supervision of the senior author (PLRW). However, at the start of the study, he had greater experience with the STAR than with the BP replacement. Both designs had mobile-bearings in which the polyethylene insert had a flat upper bearing surface, allowing it to rotate about an axis in the sagittal plane and to slide freely in the transverse plane. The lower bearing surface was highly...
congruent and allowed rotation about a single axis which ran parallel to that of the upper bearing. Both designs were intended for uncemented implantation with a porous coating on the components to encourage osseo-integration. The STAR had an additional coating of calcium phosphate (20 μm thick, for this purpose.

Varus and valgus deformity were measured before surgery from standing radiographs as previously described.\(^5\) Post-operative varus and valgus could not be reliably measured because the bony landmarks on the talus used for this assessment had been excised. Post-operative anteroposterior and lateral radiographs were taken within 24 hours to confirm that the components had been satisfactorily positioned and that the bearing surfaces were articulating congruently. At review the relative alignment of the tibial and talar components was measured to identify any loss of congruence between the bearing surfaces, described in previous reports as ‘edge-loading’.\(^5\) In order to determine whether anterior subluxation of the talus, sometimes seen in osteoarthritic ankles, had been corrected by TAR, we measured the tibiotalar ratio before surgery and at review on the lateral radiographs. The tibiotalar ratio is the proportion into which the midlongitudinal axis of the tibial shaft divides the longitudinal length of the talus and is a radiological measure of the anteroposterior tibiotalar alignment (Fig. 3).\(^5\) Significant anterior subluxation is described as more than two SDs from the mean tibiotalar ratio of 40 which is found in normal ankles.\(^6\) Radiological evidence of aseptic loosening, defined as a radiolucent line more than 2 mm wide extending across the bone implant interface, was noted, as was the presence of osteolytic cavities in the tibia which were defined as areas of limited bone resorption found in conjunction with good bony incorporation elsewhere, implying that the implant was not loose. The shape of the talar component of the STAR prosthesis completely conceals the radiological view of the bone immediately beneath it and to a less extent this is also true in the BP design. Therefore, we did not report the incidence of cavities seen under the talar components since plain radiographs could not be relied upon to detect them.

Clinical assessment with a goniometer included the comparison of the range of dorsiflexion and plantar flexion at the ankle. The American Orthopaedic Foot and Ankle
Society (AOFAS) hindfoot score was completed to assess pain (maximum 40 points) and function (maximum 60 points) after TAR.

Between August 2006 and August 2007, all patients were traced to ensure that a minimum three-year assessment of outcome had been made. The primary outcome was failure of the replacement necessitating further surgery, either by fusion or revision.

**Statistical analysis.** The original specification of sample size for a follow-up at ten years was 41 patients per group. This choice was made to detect a hazard ratio of 2.0 (based on a ten-year failure rate of 22.5% in one group and of 5% in the other) with 80% power, using a 5% two-sided log-rank test for equality of survival curves. The sample size was increased to 50 patients per group to allow for a potential constant dropout rate of approximately 2% per year for the ten-year period. This was further increased twofold to 100 patients per group because of the change in design of the BP. This latter increase was to maintain the overall power in order to enable comparisons to be made between the BP subgroups while applying a significance level of 2.5%, so that a sample size of 50 in the BP titanium and cobalt-chrome subgroups would allow a hazard ratio of 2.0 to be detected with 80% power.

Survivorship was studied using Kaplan-Meier curves with calculation of the 95% confidence interval (CI). The effect of design on the survival of the implants was analysed using Cox proportional hazard modelling, adjusting for age, gender, primary diagnosis of either rheumatoid arthritis or osteoarthritis, pre-operative varus or valgus deformity with its severity measured in degrees. Effects were tested using the Wald test. Differences between the BP titanium and cobalt-chrome subgroups were investigated using multiple comparisons; a 10% significance level was used to limit the chance of a false-negative finding. Based on the result of this test, it was decided to combine the BP titanium and cobalt-chrome subgroups for the main comparison with the STAR. Additionally, the effect of varus or valgus deformity on outcome was investigated. For these analyses, a 10% level was used to assess statistical significance. The change in the tibiotalar ratio in osteoarthritic patients was compared between treatment groups using independent-samples t-testing and the proportion of patients whose tibiotalar ratio increased was compared using Fisher’s exact test. Differences in the proportions of prostheses between the groups which experienced edge-loading was analysed using Fisher’s exact test with sensitivity to potential confounding factors investigated using logistic regression modelling. Unless otherwise stated, tests were performed using a 5% significance level.

**Results**

The age, gender and diagnostic distributions were similar for the two types of ankle replacement (Table I). The extent of the arthritic change found pre-operatively is given in Table II.

**Clinical findings.** At the last review 21 patients had died at a mean of 18 months (1 to 50) after their TAR. The mean follow-up for the surviving patients was 54 months for both the BP and the STAR groups (36 to 75 for BP and 36 to 85 for STAR). The mean follow-up for the whole study group using date last seen for surviving ankles and date of any further surgery for failure was 49 months (1 to 85).

In all, 16 (8%) of 200 TARs failed (12 BP and 4 STAR) and of these 14 ankles underwent fusion and two had revision. In the BP group, fusion was carried out for aseptic loosening in five patients, for recurrent deformity in four and for a broken titanium tibial implant in one. Revision was carried out for recurrent deformity in two ankles. In the STAR group, removal of all components and fusion was carried out for early infection in one patient. Additionally, in this group fusion was required for aseptic loosening in one and replacement of a broken polyethylene insert in two patients. A second operation subsequent to the TAR which was neither a fusion nor revision was carried out in a further four (2%) of the 200 ankles. One patient (BP) had autologous bone grafting at 63 months for osteolytic cavities beneath the tibial and talar components and another (BP) required re-impaction of the tibial component three days after surgery as a result of a technical error at the time of the initial procedure. One ankle (STAR) had a debridement at 20 months of the medial side of the joint for persistent pain and one (BP) had a washout at three weeks post-operatively for infection. At the last review all replacements had a satisfactory radiological appearance and the patients were all satisfied with the outcome. At the time of fusion we found that the bone of the talus was better preserved after

---

**Table I. Clinical details of the two groups**

<table>
<thead>
<tr>
<th></th>
<th>BP*</th>
<th>STAR†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male:female</td>
<td>58:42</td>
<td>48:52</td>
</tr>
<tr>
<td>Mean (range) age in yrs</td>
<td>64 (29 to 84)</td>
<td>65 (23 to 83)</td>
</tr>
<tr>
<td>Osteoarthritis: rheumatoid arthritis</td>
<td>69:31</td>
<td>69:31</td>
</tr>
<tr>
<td>Previous fracture of tibia or ankle</td>
<td>15</td>
<td>13</td>
</tr>
</tbody>
</table>

* BP, Buechel-Pappas
† STAR, Scandinavian total ankle replacement

**Table II. Nature and degree of severity of the pre-operative deformity in both groups (number of failures of total ankle replacements in parentheses)**

<table>
<thead>
<tr>
<th></th>
<th>&lt; 10°</th>
<th>10° to 15°</th>
<th>16° to 25°</th>
<th>&gt; 25°</th>
</tr>
</thead>
<tbody>
<tr>
<td>BP*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Varus</td>
<td>68 (4)</td>
<td>5 (1)</td>
<td>7 (2)</td>
<td>4 (2)</td>
</tr>
<tr>
<td>Valgus</td>
<td>6 (0)</td>
<td>10 (3)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>STAR†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Varus</td>
<td>66 (2)</td>
<td>9 (0)</td>
<td>1 (0)</td>
<td></td>
</tr>
<tr>
<td>Valgus</td>
<td>8 (1)</td>
<td>6 (1)</td>
<td>1 (0)</td>
<td></td>
</tr>
</tbody>
</table>

* BP, Buechel-Pappas
† STAR, Scandinavian total ankle replacement
a BP than after a STAR replacement which reflected the difference in the designs of the talar components. In the BP implant the talar fixation fins are short and the cortical sidewalls are left intact whereas in the STAR the talar fin is long and the cortical sidewalls are resected and resurfaced. However, on the tibial side, rather more bone needs to be removed at the time of revision of a BP replacement than of a STAR. The BP tibial component has a stem and this necessitates a window being cut in the anterior cortex of the tibia. Successful removal of the tibial component, however, required re-opening only its distal one-third and this window invariably united.

There were 163 surviving replacements at the time of review (75 BP and 88 STAR). The mean AOFAS score for pain and function improved to a similar extent for the two groups. All patients had severe pain as an indication for surgery with a pre-operative pain score of 0. The mean post-operative pain score in the BP group was 37 (30 to 40) and in the STAR group it was 34 (20 to 40). The mean score for function improved in the BP group from 33 (12 to 51) before surgery to 46 (29 to 58) at follow-up and in the STAR group from 31 (10 to 44) to 45 (21 to 58). The range of ankle movement did not deteriorate after surgery in any patient, but comparatively few gained any marked increase. The range of dorsiflexion improved by 10° or more in 14 of 75 (19%) BP and in 16 of 88 (18%) STAR ankles. Plantar flexion improved by 10° or more in six of 75 (8%) BP and in 15 of 88 (17%) STAR ankles.

Survival. The survivorship curves for the titanium (BPT) and cobalt-chrome (BPC) implants with fusion or revision as the endpoint.

STAR although this did not reach statistical significance (Wald test; p = 0.09; hazard ratio (HR) 2.71; 95% CI 0.86 to 8.55). However, the presence of varus or valgus deformity before surgery had a significant effect on survivorship (Wald test; p = 0.02) and this was found to be very similar irrespective of whether the deformity was in a valgus or varus direction (Wald test; p = 0.63). Nevertheless, there were relatively more failures of varus than valgus ankles at low angles of deformity. The influence of pre-operative varus or valgus deformity on survivorship was similar for both designs (Wald test; p = 0.83) and for both was more pronounced in those ankles with greater degrees of deformity (Wald test; p < 0.001). This equated to a hazard ratio of 1.64 (95% CI 1.26 to 2.14) for each 5° increment in deformity. It meant that for a patient presenting with varus or valgus deformity of the ankle of 15°, the predicted six-year survival for the STAR would be 89.9% and for the BP 73.6% whereas for an ankle neutrally aligned the predicted survival would be 97.1% for the STAR and 91.9% for the BP (Table IV).

Change in tibiotalar ratio in osteoarthritis ankles with pre-operative anterior subluxation. It was found that 46 (33%) of the 138 patients with osteoarthritic ankles had a pre-operative tibiotalar ratio of 34 or less. The mean tibiotalar ratio pre-operatively in the STAR group from 31 (10 to 44) to 45 (21 to 58). The range of ankle movement did not deteriorate after surgery in any patient, but comparatively few gained any marked increase. The range of dorsiflexion improved by 10° or more in 14 of 75 (19%) BP and in 16 of 88 (18%) STAR ankles. Plantar flexion improved by 10° or more in six of 75 (8%) BP and in 15 of 88 (17%) STAR ankles.

Survival. The survivorship curves for the titanium and cobalt-chrome BP prostheses were similar with no statistically significant difference (Wald test; p = 0.60; Fig. 4).

Because of this similarity in performance we combined the data from the two subgroups to compare the outcomes of the 100 BP ankles with the 100 STAR ankles (Fig. 5; Table III). We found a trend towards a higher failure rate for BP than
in the BP group and 7.4 (4.7) in the STAR group and although the observed increase was greater for the STAR group, this difference was also not statistically significant (Independent samples t-test, p = 0.25, 95% CI -1.3 to +5.0).

**Edge-loading.** This was observed in 12 of BP and in six of STAR prostheses. Seven of these 18 edge-loaded components were revised, of which six were BP and one was a STAR.

A total of 41 (20 STAR and 21 BP) patients had presented with a varus or valgus deformity exceeding 15°. Of these, four of the STAR and seven of the BP patients developed edge-loading. There was no significant difference between the BP and STAR designs (odds ratio (OR) BP vs STAR; 2.14, 95% CI 0.77 to 5.94, Fisher’s exact test, p = 0.22). On adjusting for the direction and extent of varus or valgus using logistic regression no statistically significant difference in survival was likely to emerge by the end of follow-up. The observed difference in survivorship could have been attributable to the greater experience which the surgeon had with the STAR than with the BP design. However, this seemed to be unlikely because the failures were not confined to those ankles operated on early in the series. In addition, the outcome was similar for the BP titanium replacement and the BP cobalt-chrome design even although the cobalt-chrome implants were used in the second half of the study. An alternative reason for the better results for the STAR prosthesis may have been that the BP design has very much shorter talar fins than those of the STAR and these may not have adequately prevented talar subsidence.

**Discussion**

There were more failures with the BP than the STAR prosthesis although this was not statistically significant. However, this was a report of medium-term follow-up powered to detect a twofold difference in hazard rates at follow-up at ten years. The observed hazard ratio was 2.71, indicating that if the pattern of failure remained consistent a statistically significant difference in survival was likely to emerge by the end of follow-up. The observed difference in survivorship could have been attributable to the greater experience which the surgeon had with the STAR than with the BP design. However, this seemed to be unlikely because the failures were not confined to those ankles operated on early in the series. In addition, the outcome was similar for the BP titanium replacement and the BP cobalt-chrome design even although the cobalt-chrome implants were used in the second half of the study. An alternative reason for the better results for the STAR prosthesis may have been that the BP design has very much shorter talar fins than those of the STAR and these may not have adequately prevented talar subsidence.
There was a significantly greater incidence of failure in ankles which had a pre-operative varus or valgus deformity, particularly when the deformity was more than 15°, leading to a predicted failure rate at six years exceeding 10% for STAR and 25% for BP. When these more deformed ankles did not fail, there was a higher incidence of edge-loading than in those with minimal pre-operative varus or valgus deformity. Edge-loading reduces the area of contact between the polyethylene and metal-bearing surfaces and may lead to loosening or osteolysis.

The six-year survivorship of 86.5% (95.0% for STAR and 79.0% for BP) for the whole study group was similar to that of previous reports. However, when the pre-operative deformity was 15° or less the estimated six-year survivorship was 90.6% for the whole group, 86.7% for the BP and 95.6% for the STAR. Doets et al. found a similar effect of varus and valgus deformity on the survivorship of ankle replacement in rheumatoid patients. We had observed this in an earlier study with a different group of patients. The difficulties encountered in undertaking successful ankle replacement in the presence of varus or valgus deformity have been previously described. We recommend that replacement be advised with caution in the presence of varus or valgus deformity and that it should not be performed when the deformity exceeds 15°.

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