Assessing the dog as a model for human total hip replacement

ANALYSIS OF 38 CANINE CEMENTED FEMORAL COMPONENTS RETRIEVED AT POST-MORTEM

C. P. Skurla, G. E. Pluhar, D. J. Frankel, E. L. Egger, S. P. James

From Baylor University, Waco, USA

Post-mortem retrieval of canine, cemented femoral components was analysed to assess the performance of these implants in the dog as a model for human total hip replacement (THR). Mechanical testing and radiological analysis were performed to determine the stability of the implant and the quality of the cement. Thirty-eight implants from 29 dogs were retrieved after time intervals ranging from 0.67 to 11.67 years. The incidence of aseptic loosening was 63.2%, much higher than in human patients (6% in post-mortem studies). Failure of the femoral implants began with debonding at the cement-metal interface, similar to that in implants in man. The incidence of aseptic loosening was much lower in bilateral than in unilateral implants. Significant differences were observed for three different designs of implant. While the dog remains the animal model of choice for THR, results from this study provide insight into interspecies differences in the performance of implants. For example, the performance of THR in dogs should be compared with that in young rather than in elderly human patients.

Despite many improvements in the design of prostheses and surgical technique since the introduction of cemented total hip replacement (THR) in humans, the rate of failure because of aseptic loosening remains high. Rates of aseptic loosening of between 5% and 8% have been reported at follow-up of 15 years1,2 and 16% at 20 years in primary cemented femoral components implanted using second-generation cementing techniques.2 In revised, cemented femoral components, it was found to be 33% at a follow-up of ten years,3 much higher than for primary surgery.

In addition to clinical studies, examination of successful cemented THRs retrieved at post-mortem in humans4,6 has provided valuable insight into their performance and failure, despite the difficulties in controlling statistical power. These studies have given information which cannot be obtained from failed implants retrieved at revision or from review of clinical outcomes. They have shown that 6% of clinically successful femoral components in humans were loose and have identified the initiating event for failure as being mechanical in nature.4,6 Debonding of the polymethylmethacrylate (PMMA) cement-implant interface began proximally and distally, progressing towards the middle of the stem.

Experimental studies in animal models7-15 have also provided valuable knowledge for the development of new materials and designs for implants. The dog is the model of choice16 for preclinical trials of new implants for THR in man. The short duration, three weeks to three years, of experimental studies7-15 in the dog contrasts sharply with the duration of human clinical use and of post-mortem retrievals4,6 at two weeks to 21 years.

Our aim was to gain a better understanding of the performance of the cemented THR in the most common animal model16 by analysing canine, cemented femoral components which had been retrieved at post-mortem and comparing the results with similar studies in human patients.

While not as common as in man, cemented THRs are performed in dogs for much the same reasons but before the present study, there had been no post-mortem retrieval assessment. The long-term performance of these implants was unknown except by retrospective studies of clinical cases carried out using office visits, evaluation of clinical radiographs, and subjective responses to phone and mailed surveys from owners.17-24 In a review of seven retrospective studies of canine THR,17-23 aseptic loosening of the femoral component requiring revision was reported in 19 (3.0%) of 625 cases with follow-up from 0.1 to 6.2 years. However, there had been no objective engineering analysis to verify the clinical performance of THR in dogs.
In our study the specimens retrieved comprised metal femoral stems and non-metal-backed acetabular components made from ultra-high-molecular-weight polyethylene (UHMWPE) and implanted using similar cementing methods with PMMA.

We tested five hypotheses. The first was that the initiating event for clinical failure of canine, cemented femoral components would be mechanical debonding of the PMMA from the metal implant, similar to that observed in man, as shown by mechanical testing and analysis of contact radiographs. If true, the dog would prove to be a good model for testing implants designed to address these mechanisms of failure. Any significant differences in events initiating failure would have to be taken into account when planning to use the dog as a model for testing of human implants in vivo.

The second hypothesis was that the incidence of aseptic loosening of the cemented femoral component would be lower in dogs than in man. This was based on several studies which have shown that dogs carry between 53% and 65% of their weight on their front legs, make subtle adjustments to their gait, and have lower joint reaction forces as a percentage of body-weight than man. If true, the dog may not be a good animal model for studying human implants designed to reduce rates of loosening.

The third hypothesis was that there would be no difference in the incidence of loosening between unilateral and bilateral implants. If true, then studies on both unilateral and bilateral implantation should produce similar results. Anecdotal feedback from veterinary surgeons and their clients before this study indicated that dogs with unilateral THR, despite bilateral degenerative joint disease, experienced surgical outcomes comparable with animals with bilateral THR. Experimental studies using healthy dogs to test human implants most often use unilateral implantation. However, bilateral implantation has occasionally been used to decrease the number of animals involved or to compare the performance of different implants in the same animal.

The fourth hypothesis was that implants of short duration (less than three years) would be more stable than those of longer duration. If true, then most experimental studies, which have durations significantly shorter than three years, may not give an accurate indication of the long-term performance of implants.

The fifth hypothesis was that different designs of implant would show differences in initiating failure and/or in the incidence of aseptic loosening as shown by mechanical testing. If true, this would reinforce the validity of the dog as a model for studies of human implants since most experimental studies are designed to look at differences in design.

**Materials and Methods**

Implants with hemipelvses and femora were donated by owners (Table I), after signing a standard form for donation at the time of their dog’s death by natural causes or euthanasia.

<table>
<thead>
<tr>
<th>Implant design</th>
<th>Unilateral</th>
<th>Bilateral</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Richards</td>
<td>3</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>BioMedtrix 1</td>
<td>14</td>
<td>8</td>
<td>22</td>
</tr>
<tr>
<td>BioMedtrix 2</td>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
<td>18</td>
<td>38</td>
</tr>
</tbody>
</table>

There were 29 dogs, nine of which had bilateral implants giving a total of 38 prostheses. The duration of implantation ranged from 0.67 to 11.67 years. The weight of the dogs ranged from 18.2 kg (Brittany spaniel) to 66.4 kg (Newfoundland). Approval from the committee for institutional animal use and care was not required for our study since the animals were donated after death. Clinical details including the weight of the dog, the surgeon, and the date of surgery were collected from the clinical records. The specimens were stored at -20°C until testing began. After removal of the soft tissues, cranio-caudal and mediolateral (ML) contact radiographs of the intact femora were made using a Faxitron x-ray cabinet (Faxitron X-Ray Corporation, Wheeling, Illinois).

Implants of three designs were retrieved. All cemented femoral components had collars and articulated with cemented, non-metal-backed acetabular components made of conventional UHMWPE. The first implant used was the Richards (Smith & Nephew Richards, Memphis, Tennessee) and consisted of a monolithic Co-Cr femoral component. The shape of the proximal cross-section was rectangular, and the distal tip was tapered. The second design, the BioMedtrix 1 (BioMedtrix Inc, Allendale, New Jersey), had a modular femoral component with a Ti-6Al-4V stem and Co-Cr head. This implant has been reported to have a high incidence of aseptic loosening. The femoral component had rounded corners on the lateral aspect of the proximal stem. The cross-section of the distal one-third of the femoral stem was round with grooves on the cranial and caudal sides. The third design, the BioMedtrix 2 (BioMedtrix Inc.), was modular and made entirely of Co-Cr. The proximal cross-section was oval in shape with more rounded corners and was thinner in the cranio-caudal direction. The shape of the distal one-third of the femoral stem was the same as that of the BioMedtrix 1.

**Mechanical testing to assess stability.** Existing techniques for measuring the stability of the femoral component were adapted from those used in human post-mortem retrieval studies. On manual testing, if the femoral implant subsided within the femur or moved in a medial-to-lateral direction (ML translation) by more than 3 mm, it was classified as being grossly loose and was not subjected to further mechanical testing. In previous human studies, ML translation has been described as rotation of the implant within the femur. The term, rotation, will be used henceforth, for simplicity.
The grossly firm implants were subjected to quantitative testing of movement at the bone-PMMA interface, and then between the bone and the implant. The movement at the PMMA-implant interface was calculated as the difference between the two measurements.

A special jig was developed to hold the femur in the correct position on the MTS biaxial mechanical testing system (MTS Systems Corporation, Eden Prairie, Minnesota) in order to apply a load to the femoral head in the direction of the peak joint reaction force as determined by Page et al.28 An extensometer (MTS Systems Corporation), attached to a relative movement device (RMD) for detection as described by O’Connor et al30 was used to quantify relative movement (Fig. 1).

From the contact radiographs of the intact femora (Fig. 2), the thickness of the bone and PMMA in the ML view as well as the position of the implant within the femur in the craniocaudal view were measured in order to determine the placement of a hole 6.4 mm in diameter to be drilled through the bone to the bone-PMMA interface. The hole was placed on the cranial side of the femur at least 1 cm distal to the lower edge of the collar of the implant in the widest and flattest part of the femur and implant. The RMD was cemented to the bone according to the method of O’Connor et al.30 When the extensometer blades were placed perpendicular to the long axis of the femur (Fig. 1), the measurements reflected subsidence of the implant within the femur. When the blades were placed parallel to the long axis of the femur, the measurements reflected rotation of the implant within the femur. Two tests, one for subsidence and one for rotation, were performed to measure relative movement at the bone-PMMA interface. The hole was extended through the PMMA to the PMMA-prosthesis interface. The procedure was repeated to measure movement of the prosthesis relative to the bone in subsidence and rotation. The order of the tests (subsidence first or rotation first) at each interface was randomised.

In retrieval studies on human specimens4–6 femora have been tested for stability using a peak load of 1112 N (2.5 times the body-weight of a 45.5 kg person) regardless of the body-weight of the donor. Since the aim of this study was to be non-destructive, and since the retrieved implants came from dogs which varied considerably in size and weight, the peak load used in the testing of stability was normalised to the body-weight of the individual dog from which the implant had been retrieved. A load of 75% of the peak joint reaction force (1.5 times body-weight)26,28 was selected. The test was repeated four times for each test condition. The maximum relative movement achieved during each of the four test conditions was used to decide whether the implant was mechanically loose or not.

Control femora were used to determine the criteria for classification of mechanically loose vs firmly implanted and to verify that the test method did not damage a firmly implanted specimen. The firmly implanted control consisted of a cadaver canine femur with a Biomedtrix 2 femoral component which had been introduced by a veterinary orthopaedic surgeon using a standard surgical procedure.
The mechanically loose control, similarly prepared, had a loose PMMA-prosthesis interface produced by moving the femoral implant in subsidence while the PMMA cured. The control femora were tested using the same method as that used for the retrievals. The firmly implanted control femur was tested on multiple occasions to ensure that the test method did not cause loosening.

The measurements of relative movement of the firmly implanted control was 2 µm in both subsidence and rotation at the bone-PMMA interface and 10 µm in subsidence and 6 µm in rotation between the bone and implant. For the mechanically loose control the movement was 10 µm in subsidence and 4 µm in rotation at the bone-PMMA interface and 57 µm in subsidence and 40 µm in rotation between the bone and implant. Since the PMMA-implant interface was the only disrupted interface and was disrupted in subsidence only, the measurements of subsidence between bone and implant was used as the criterion to classify the femoral component as mechanically loose or firmly implanted in both directions at both interfaces. The RMD error was ±12 µm and therefore, a value of 45 µm (57 µm minus 12 µm) was determined as the threshold for classifying an implant as mechanically loose or firmly implanted at each interface and in each test direction. The PMMA-implant interface was classified as loose if the bone-implant measurement was more than the bone-PMMA measurement and more than 45 µm.

**Analysis of contact radiographs.** The craniocaudal and ML contact radiographs of the femora were marked with modified Gruen zones as shown in Figure 2. The relevant radiological features measured were the minimum and maximum thickness of the cement mantle in each zone, the presence and size of voids, cracks, and radiolucencies, axial and rotational alignment of the implant within the femur and the presence of contact between the distal tip of the implant and bone. An orthopaedic veterinary surgeon measured and assessed the radiographs.

**Transverse sectioning of femoral components.** The craniocaudal contact radiographs, marked with modiﬁed Gruen zones as shown in Figure 2, were used to plan the sites of transverse slices. Since previous human and canine studies had found a relationship between aseptic loosening and the position of the distal tip of the femoral component, a 2 mm slice taken 3 mm proximal to the distal end of Gruen zone 3/5 was considered to be the most clinically relevant slice for digital image analysis (Fig. 2). Transverse slices were made using a bone saw (Exakt Technologies Inc., Oklahoma City, Oklahoma). The slice of interest and the remainder of the specimen were placed in 70% ethanol. The remainder of the specimen was reserved for future radiological, microscopic, and histological analysis. Contact radiographs were made of the slice of interest (Fig. 3).

**Digital image analysis.** Digitised contact radiographs of the slice of interest for each femur were analysed using Scion Image (Release Beta 3b Scion Corporation, Frederick, Maryland). The measurements made included the areas and centroid locations of the implant, PMMA, and cortical bone, the minimum and maximum thicknesses of the cement mantle and the distances from the implant centroid to the PMMA centroid (distance 1), from the implant centroid to the bone centroid (distance 2), and from the PMMA centroid to the bone centroid (distance 3) (Fig. 3). These distances indicated how well the distal tip of the implant was centralised within the cement mantle. Small distances indicated a well centred implant.

**Statistical analysis.** All statistical analyses were performed using SAS software (SAS Institute, Cary, North Carolina). Non-parametric, Kruskal-Wallis one-way analysis of variance (ANOVA) was used to compare the age and weight among the implant designs. The non-parametric Wilcoxon rank-sum test was used to compare implants with duration of less than three years with those in place longer.

Frequency data were calculated on categorical data (loose vs firmly implanted) based on the mechanical testing results. Categories were compared using the chi-squared test. Measurements of relative movement were compared between manufacturers with all BioMedtrix implants combined into one category using the two-sample t-test for unequal variances. Repeated measures analysis was performed with the subject being each case and the repeated measure being the four measurements of relative movement and the two calculated measurements.

Our study had a relatively low number of samples and a relatively large number of measured variables, which reduced the power of the statistics. The initial statistical analysis showed only a few significant differences between the measured variables.

To reduce dimensionality and enhance statistical power, principal component analysis was used to identify highly correlated variables, some of which were then combined into index variables based on the expert opinion of the authors. Only variables with the same dimensions (e.g. length) were combined. Square-root transformations were
used for those indices which did not have a normal distribution in order to allow the use of parametric statistics. Index variables of note which were defined included the bone-PMMA index (i.e. square-root transformation of the mean of subsidence and rotation measurements at the bone-PMMA interface) and the bone-implant index (i.e. square-root transformation of the mean of subsidence and rotation measurements between bone and implant). The square-root transformations allowed the use of simple linear regression analysis with the bone-PMMA index and the bone-implant index as dependent variables against all independent variables. Other dependent variables used in simple linear regression analysis were the PMMA-metal interface and the bone-PMMA interface (i.e., loose designation in one of four tests or two calculations), or had been classified as grossly loose after manual testing. The remaining 29 were mechanically tested.

Of the eight loose bilateral implants, two were grossly loose and six mechanically loose. Of the 16 loose unilateral implants, seven were grossly loose and nine mechanically loose.

**Statistical analysis.** Using ANOVA, there were no significant differences in age or weight of the dogs for all designs of implants used. When grouped by duration (less than three years vs more than three years) Wilcoxon rank-sum analysis found a significantly higher minimum thickness of PMMA, measured on the transverse slice, in the group with shorter duration (p < 0.05).

The frequency distributions of categorical results of the stability of femoral implants were calculated (Table II). Although not statistically significant by chi-square analysis, there appeared to be a trend of Richards femoral components to be more stable than BioMedtrix components. Using the Wilcoxon rank-sum test, the relative movement measured for unilateral implants was significantly greater than for bilateral implants in subsidence at both interfaces and in rotation at the PMMA-implant interface (p < 0.05; Fig. 4). When relative measurements of movement were compared according to implant design using ANOVA, there were no significant differences, but there was a trend that BioMedtrix 1 implants were less stable than Richards and that BioMedtrix 2 implants were less stable than BioMedtrix 1. When measurements of relative movement were compared between manufacturers with all BioMedtrix implants combined into one category using the two-sample $t$-test for unequal variances, there was significantly more relative movement measured for the BioMedtrix than Richards implants in rotation at both interfaces and in subsidence at the bone-PMMA interface (p < 0.05; Fig. 5). Repeated measures analysis with the subject being each case and the repeated measure being the four relative measurements of movement showed that the subsidence at the PMMA-implant interface was significantly greater (p < 0.05) than the other three measurements, and the rotation at the PMMA-implant interface was significantly greater than rotation at the bone-PMMA interface.

Few significant linear correlations were found, partly because of the low number of specimens and the resulting

<table>
<thead>
<tr>
<th>Table II. Implant stability in subsidence*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>------------------------------------------</td>
</tr>
<tr>
<td>All implants combined</td>
</tr>
<tr>
<td>Implant design</td>
</tr>
<tr>
<td>Richards</td>
</tr>
<tr>
<td>BioMedtrix 1</td>
</tr>
<tr>
<td>BioMedtrix 2</td>
</tr>
<tr>
<td>Unilateral vs bilateral</td>
</tr>
<tr>
<td>Unilateral</td>
</tr>
<tr>
<td>Bilateral</td>
</tr>
</tbody>
</table>

* with the exception of three implants (7.9% of total), when an implant tested firm (or loose) at an interface in subsidence it also tested firm (or loose) in rotation.
low statistical power. There was a strong positive correlation \((p = 0.001, R^2 = 0.348, power = 0.956)\) between the bone-implant index and distance 3, the distance between the PMMA and bone centroids (Fig. 3). There was a weak but significant negative correlation \((p = 0.017, R^2 = 0.192, power = 0.686)\) between the number of implants and the bone-implant index.

**Discussion**

The first hypothesis regarding the event initiating failure in loosening of the femoral implant was supported. Similar to human components, canine femoral components tended to loosen at the PMMA-implant interface before the bone-PMMA interface. In our study, all except one of the implants which had been classified as mechanically loose at the bone-PMMA interface were also mechanically loose at the PMMA-implant interface, but seven of the 14 implants classified as mechanically loose at the PMMA-implant interface were still firmly implanted at the bone-PMMA interface. The single implant which was mechanically loose at the bone-PMMA interface and firmly implanted at the PMMA-implant interface was in contact with the endosteal bone surface at the distal tip of the implant, as seen on the transverse femoral slices, implying that there was no PMMA mantle present at this level.

On analysis, failure in canine femoral components, as in the human specimens, first appeared to be measurable in subsidence at the PMMA-implant interface. The mechanical debonding of the PMMA-implant interface, observed in the dogs with degenerative joint disease in our study, should also be observed in healthy experimental animals, provided that they are allowed normal levels of activity.

The second hypothesis that the incidence of aseptic loosening of the femoral component in dogs would be lower than that in human specimens was not found to be true. Canine femoral components were either grossly or mechanically loose in 63.2% of retrievals. This percentage is much greater than the 6% found in human post-mortem retrieval studies. All femoral components in the human follow-up and post-mortem retrieval studies were made from similar materials, implanted using similar methods, and tested for mechanical stability using methods similar to those used in the present study. No bacterial cultures were made of the retrievals in this study but, none of the contact radiographs showed irregular periosteal proliferation consistent with infection. Based on these findings and the low incidence of septic loosening in previous studies, the loosening found in this study was likely to be aseptic.

THR in human patients of less than 50 years of age has been shown to fail more quickly than that in older, more sedentary patients. The high incidence of loosening of canine femoral components may be due to the expectation of the owner that their dog will enjoy completely unrestricted activity within a few months of surgery. The higher incidence of aseptic loosening observed in the dogs in our study may be apparent in short-term studies using experimental animals since no significant difference was found in the incidence of loosening between groups of short and long duration. The higher incidence of loosening may indicate that the performance of canine THR clinically resembles that of human THR in young patients and that researchers using the canine model may wish to limit the levels of activity of the dogs if interested in simulating THR in more sedentary patients. In addition, age-related changes in human femora, particularly in the female, coupled with the lack of age-related changes in canine femora, reinforce the need to compare the performance of canine THR with that in young human patients.

The third hypothesis that unilateral implants would perform adequately in comparison with bilateral implants was not supported. Bilateral implants were significantly more stable than unilateral implants. Owners and veterinary surgeons have reported anecdotally that, after unilateral THR, dogs with bilateral degenerative joint disease place more weight on the implanted side than on the diseased, intact limb. This may result in overloading and earlier aseptic loosening of the unilateral femoral components. This finding may not directly translate to experimental studies.

<table>
<thead>
<tr>
<th>Relative movement (µm)</th>
<th>Bone/PMMA relative movement</th>
<th>Bone-implant relative movement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subsidence*</td>
<td>Rotation</td>
<td>Subsidence*</td>
</tr>
<tr>
<td>Unilateral</td>
<td>Bilateral</td>
<td>Richards</td>
</tr>
</tbody>
</table>

The mean relative movement of unilateral vs bilateral implants at each interface and in each test direction (unilateral vs bilateral at p < 0.05).

<table>
<thead>
<tr>
<th>Relative movement (µm)</th>
<th>Bone/PMMA relative movement</th>
<th>Bone-implant relative movement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subsidence*</td>
<td>Rotation</td>
<td>Subsidence*</td>
</tr>
<tr>
<td>Richards</td>
<td>BioMedtrix</td>
<td>Richards</td>
</tr>
</tbody>
</table>

The mean relative movement by manufacturer at each interface and in each test direction (Richards vs BioMedtrix at p < 0.05).
because healthy experimental animals would not be as likely to transfer load to an implanted hip if the contralateral hip was not diseased. However, in experimental studies using bilateral implantation of different implants, unequal loading of the hips may occur with differing degrees of pain and function.

The fourth hypothesis suggesting that implants of short duration would perform better than those of longer duration, was not strongly supported and requires further study. When comparing short duration (< three years) vs long, the only significant difference between the groups was a greater minimum thickness of PMMA at the distal tip in the implants of shorter duration, which was probably due to recent improvements in surgical technique including centralisation of the tip.

Improper positioning of the distal tip of the implant, resulting in thin cement mantles or contact of the tip with endosteal bone, has been implicated as a predictor of aseptic loosening of the femoral component in both human and canine THR. Our findings support these previous results. There was a significant positive correlation between the bone-implant index and distance 3, the distance between the PMMA and bone centroids. This indicates a relationship between improper centralisation of the tip and loosening of the implant.

The fifth hypothesis that different implant designs would have differences in the initiating events of failure and/or the incidence of aseptic loosening was supported. While the same initiating event, debonding of the PMMA-implant interface, was found in all three designs, the incidence of aseptic loosening differed among them. The Richards implants were the initial design used and were introduced using earlier surgical techniques resulting in cement mantles of poorer quality, but they were more stable than either of the BioMedtrix designs.

The Richards implant has a more pronounced distal taper in comparison with the BioMedtrix design. Huiskes and Verdonschot stated that in cases in which the femoral stem had a very prominent taper at the tip, distal stresses were reduced, and this may have protected the PMMA-implant interface bond.

However, differences in design may not fully explain this result as there were confounding factors. The uneven distribution of design between unilateral and bilateral cases, with the more stable Richards prosthesis making up 50% of the bilateral cases but only 15% of the unilateral, may explain the higher incidence of loosening in the latter implants. The reverse, on the other hand, could be true. The Richards design may have been more stable because 75% of them were bilateral; whereas, only 35% of the combined BioMedtrix implants were bilateral.

Conclusions
More than half of the canine femoral components (63.2%) were loose, in contrast to the 84% of positive responses perceived by dog owners in a previous survey. The incidence of aseptic loosening observed in dogs was much higher than that seen in humans, suggesting that the clinical performance of canine THRs is more representative of that in young human patients.

The results for implants in place less than three years were similar to those for longer duration. Significant differences in performance were observed for the different designs.

We wish to thank James ZumBrunnen, Donna Wheeler, Lisa Eschbach, Diane Beranek, Daniel O’Connor, and Charles Bradgon for technical assistance, BioMedtrix Corporation for donation of the femoral components and Howmedica Corporation for the donation of PMMA bone cement.

No other 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