Clinical outcomes assessment of three similar hip arthroplasty bearing surfaces

Christopher Parsons,1 Ryan Batson,1 Shane Reighard,1 Stephanie Tanner,2 Becky Snider,2 Thomas B. Pace1,2
1University of South Carolina, School of Medicine, Greenville, SC; 2Department of Orthopedics, Greenville Health System, SC, USA

Abstract

This report examines the clinical performance of three similar total hip arthroplasty bearing designs with distinctly different bearing surfaces used over the course of 10-17 years. Clinical outcomes assessments for each group are compared in the context of varying implant related costs related to the latest technology at the time of surgery. Eighty-one surgeries were studied and differ by bearing surface. In this study, 36 hips are ceramic on polyethylene, 27 are metal on polyethylene and 18 are metal on metal. All polyethylene components are non-highly cross-linked. The ceramic on polyethylene group has younger patients, on average, and higher percentage of patients with significant polyethylene wear. These groups have an average follow-up time of 8.6 years when assessing functional hip scores, thigh pain, groin pain, revision surgeries and radiographic osteolysis. The implant purchasing cost at the time of surgery was assessed to determine if a correlation exists between outcomes and more technologically advanced implants use at the time of surgery. Based on midterm clinical outcome assessment, no correlation between initial hospital cost and clinical outcomes of one bearing surface over another can be found.

Introduction

Total Hip Arthroplasty (THA) is a widely utilized procedure for severe hip arthritis. A recent survey reported over 285,000 procedures annually in the United States alone.1 And, Kurtz et al. predicts the total number of annual primary THAs will reach 572,000 by the year 2030 in the United States.2 Many hip system design changes over the past two decades focused on altering the articulating bearing surface of the implant system with the aim of yielding better results and higher patient satisfaction. More involved manufacturing processes are often associated with an increased cost of processing and subsequently increased hospital related implant purchasing cost.

The three most common hip arthroplasty bearing surfaces used in the United States are metal on polyethylene (MoP), metal on metal (MoM) and ceramic on polyethylene (CoC and CoP) comprising 51%, 35% and 14% respectively.3,4 The expected increased frequency of THA in the US in an environment of more limited health care funds further necessitates the need to evaluate the value of different bearing surfaces relative to associated cost and any improved patient focused clinical outcome.

Reports from arthroplasties performed in the late 1980’s and early 1990’s have suggested that earlier version of ceramic heads (alumina) articulating against conventional polyethylene have improved clinical and radiographic findings when compared with cobalt chrome alloy heads against conventional polyethylene in active patients.5 Subsequent manufacturing techniques produced the current ceramic material (Biolox®, Ceramtec, Plochingen, Germany) which has diminished the fracture problems in the earlier alumina ceramic heads while improving wear rates.5,6 Metal on metal hip bearing surfaces (Metasul®, Zimmer, Warsaw IN, USA) with and without polyethylene backing of the cup metal liner have been used with varying reports of success and failure over the past three decades.7,8

Cobalt chrome heads articulating with conventional polyethylene have been in use since the early 1960’s. To our knowledge, there is no report comparing clinical and radiographic outcomes of the same Natural Hip femoral stem with three different bearing surfaces: metal on conventional polyethylene, Metasul metal on metal and the Biolox ceramic heads on conventional polyethylene in a general population of hip arthroplasty patients. In addition, to our knowledge there is no report assessing the respective hospital related cost for each of these three implant types to assess if any variance in implant cost could be justified based on outcomes assessment.

The present study examines the radiographic and clinical outcomes for each of three different total hip bearing surfaces in cases where the femoral stem, cup, surgical implantation technique, and rehabilitation protocol were non variables. The bearing surfaces were the only variables for each of the three groups. Using patient follow-up data, the clinical and radiographic assessments of each hip articulating bearing surface are compared, and the overall outcome is weighed against the associated cost for each hip implant system. The authors declare that there is no conflict of interests regarding the publication of this article.

Materials and Methods

Institutional Review Board approval was obtained for this retrospective study. Criteria for the study included THA cases with at least two years follow up. All subjects received the Natural Hip press fit femoral stem and acetabular cup (Zimmer) and all received a femoral head size of 28 mm. Eighty-one THAs performed by the same surgeon were performed between February 1996 and January 2003. The cases were separated based on bearing surface of the head and cup into three groups for study: MoM, MoP and CoP. The metal heads (both standard cobalt chrome and MoM Metasul) were manufactured by Sulzer (Winterthur, Switzerland, now Zimmer). The ceramic heads were manufactured by Ceramtec (Plochingen, Germany). The femoral stems in all cases were press fit Natural hip stems, all heads were 28 mm diameter heads and the acetabular cups were modular press fit Intra-Op cups also manufactured by Sulzer (Winterthur, Germany).

Surgical technique

All surgeries were performed by the senior author (TBP) using a postero-lateral approach. The decision to use a metal on metal bearing surface vs. ceramic on polyethylene or cobalt chrome on polyethylene was based both on the availability of the material as an FDA approved arthroplasty option and an assessment of the patient’s age and activity level, anticipating that the younger more active patients might benefit from the latest technology of either

[Orthopedic Reviews 2014; 6:5334]
metal on metal or ceramic heads articulating with polyethylene. Highly cross-linked polyethylene was not used in any of these cases. Other than the choice of bearing surface preference for younger more active patients receiving either MoM or CoP, the patient selection criterion remained the same for each group.

**Post-operative care**

Post-operative rehabilitation included weight bearing as tolerated with walker support for 3–6 weeks. Venous thrombosis event (VTE) prophylaxis was based on individualized patient risk assessment. For the standard at risk patient, this included oral warfarin 5 mg daily beginning the night of surgery until the prothrombin time was 15 seconds or until the INR (international normalized ratio) was 1.2 or greater, at which time the warfarin dosage was reduced to 2 mg daily. This was continued as has previously been reported for 4 weeks as a 2 mg per day mini fixed dose oral warfarin regimen. For the patient without a higher VTE or bleeding risk assessment, once the hospital prothrombin time reached 15 seconds or INR levels reached 1.2-2.0, post discharge monitoring was not done unless signs or symptoms of bleeding occurred. For higher risk VTE patients, higher dose monitored oral warfarin was used (prothrombin time of 18-20 seconds or INR range 2.0-2.5). Early in the series the hospital used a prothrombin time based laboratory system (based on patients bleeding time). The hospital used a prothrombin time based laboratory system (based on patients bleeding time). The hospital used a prothrombin time based laboratory system (based on patients bleeding time). The hospital used a prothrombin time based laboratory system (based on patients bleeding time). The hospital used a prothrombin time based laboratory system (based on patients bleeding time). The hospital used a prothrombin time based laboratory system (based on patients bleeding time).

**Clinical follow up**

Harris Hip Scores (HHS) were collected at follow up office visits by the senior author. The score was recorded before surgery for all patients and at standard follow up intervals after surgery of three, six, and twelve months, and then annually. The presence or absence of thigh or groin pain was determined from the Harris Hip Score functional assessment record.

Anteroposterior radiographs of the hip and pelvis and a Lowenstein lateral radiograph of the hip were obtained at each follow up visit, and compared with the immediate postoperative radiographs. Acetabular components were evaluated for cup inclination (greater or less than 55 degrees), peri-implant radiolucent lines and osteolytic areas in the regions described by DeLee and Charnley. The femoral components were evaluated for radiolucent lines and osteolytic areas described by Gruen et al. and Johnston et al. Calcar round-off (defined as resorption of calcar bone back to the junction of collar and stem in zone VII) was also recorded. Erosion under the collar was defined as radiographic bone loss between 2 to 10 mm extending distally from the junction of the collar and stem. Calcar bone loss greater than 10 mm was considered structurally significant osteolysis. Stem subsidence was determined by measuring the distance from the proximal tip of the greater trochanter to the lateral shoulder of the prosthesis on successive radiographs. This measurement was corrected for magnification using the known diameter of the prosthetic head to its measured diameter on the radiograph. A change of 2 mm or more was considered evidence of subsidence. Heterotopic bone formation was evaluated at the second year following surgery and classified by the method of Brooker et al.

Descriptive analysis consisted of percentages for qualitative data and means and ranges. Statistical analysis was performed using Analysis of Variance for continuous data. Fisher’s exact test was used for qualitative data analysis. Significance was set at P<0.05 for all tests.

**Results**

The MoP, CoP, and MoM groups consist of 27, 36 and 18 subjects respectively. Age, gender, pre-operative HHS and follow up time for each group is shown in Table 1. There is a significant difference between the ages of the three groups (P=0.039) with the youngest group (average age 57.8 years) receiving the CoP implants and the oldest (average age 64.7) receiving the MoM implants. There is also a significant difference between the follow up times (P=0.02) with the longer follow up time (average 9.9 years) in the CoP group and the shortest follow up time (average 7.6 years) in the MoP group. There is no difference between the preoperative Harris Hip Scores of the three groups. There is a trend for more females in the MoP and MoM groups, but these are not statistically significant. The most recent clinical follow up assessment score, incidence of thigh or groin pain and incidence of revision surgery for each group are shown in Table 2. There is no significant difference among the three groups when measuring the most recent Harris Hip Scores. The occurrence of thigh or groin pain was not statistically different between the three groups. Two patients underwent early revision surgery; both are in the MoP group, and the revision surgery was performed for recurrent instability that was unrelated to accelerated polyethylene wear based on radiographic assessment.

Complications are limited to dislocation, hematoma requiring surgical evacuation, heterotopic ossification and surgical site infection requiring surgical irrigation. There are no symptomatic pulmonary emboli or deep venous thrombosis. Complications for all three groups are shown in Table 3. There is no significant difference in the incidence of complications related to articular wear.
between the three groups.

Radiographic assessment performed from standing anteroposterior and lateral radiographs prior to surgery and at each subsequent follow up clinical visit is shown for the three groups in Table 4. There are no cases of femoral stem osteolysis, acetabular cup osteolysis or stem subsidence noted. There are no differences in cup inclination angles between the three groups. Linear wear assessment could not be performed on the MoM group due to the absence of the polyethylene component. For the CoP and MoP groups there were 12 cases and 2 cases respectively of measured linear wear of 2mm or greater for each group respectively, producing a significant difference (P=0.03). While the CoP had a higher incidence of wear, the clinical assessment between the groups did not differ nor did the incidence of osteolysis and wear related revision surgery.

The hospital related implant purchasing cost for each of the three total hip systems is the same for all components aside from the three different bearing surfaces. The implant specific purchasing cost for each is shown in Table 5. These costs reflect the host hospital implant purchasing contracts in effect at the time of the index arthroplasty. These costs reflect the price for the stem, cup and ball components of the implants.

### Discussion

In mid to long term follow up, the primary reason of failure for primary hip arthroplasties is implant related failure as a result of wear debris from the bearing surface. The bearing surfaces of metal on metal, polyethylene on metal heads and ceramic heads on polyethylene have been the main choices for most arthroplasties in the US over the past three decades. There are varied costs associated with each of these designs as well as purported benefits and risks for the use of each.

### Metal on metal

The current generation of MoM THAs, which utilizes better materials and machining techniques, was developed in the late 1980s as an improvement over the older generation of devices and is still in use today. MoM offers increased durability over the long term, compared to MoP THAs. The conventional MoP implants have shown a 50% 25-year failure rate in younger, more active patients. Meanwhile, MoM implants have been shown to have a long durability and low failure rate due to implant complications. MoM implants allow for the use of larger femoral heads, which increases the range of motion for the joint. However, the use of MoM implants has declined significantly in the past several years due to the suspected harmful effects of metal ions generated by implant wear, high revision rates, recalls of two models made by DePuy Orthopedics, and FDA requirements for surveillance of MoM implants.

Additionally, there is conflicting data reported among the literature in regards to the effect of femoral head size on the clinical outcomes. Shimmin et al. discuss the significantly different amongst metal ion concentrations in varying femoral head sizes. Meanwhile, another prospective study shows no significant difference between ion levels of 28 mm and 36 mm heads at two year postoperative levels following a MoM implant. However, we controlled for any effect femoral head size would play in clinical outcome by limiting that variable to 28 mm femoral heads. As a result, we are able to focus on the impact the bearing surface has on the clinical outcome.

### Ceramic on polyethylene

Alumina ceramic (polycrystalline aluminum oxide, Al2O3) femoral heads on polyethylene cups were introduced in the late 1960’s. Contemporary, ceramic femoral heads convey several theoretic advantages over traditional metal articulations. These advantages include: superior lubricating properties, biologically inert behavior, and a smooth surface finish. Additionally, the improvements in manufacturing led to newer ceramic materials that are less likely to fracture, such as BioLOX®. Ceramic femoral heads coupled with conventional polyethylene acetabular cups may also provide advantageous wear characteristics over ceramic on ceramic (CoC) bearings while reducing the risk of fracture reported with CoC. A recent report of ceramic versus cobalt chrome articulating femoral heads found the wear rate to be significantly different between the two groups in young patients (0.09 mm/yr for CoP vs. 0.14 mm/yr for MoP), while the rate of osteolysis was very similar for each group.

This study examines the clinical and radiographic outcomes of patients who underwent THA with three different bearing surfaces. The femoral stem was the same in all groups (natural hip press fit femoral stem) and has a proven track record of excellent ingrowth and function beyond ten years. Some significant differences were found between the groups. There is a significant difference among the average ages of the groups. The CoP group has a younger average age, and the MoP group has the oldest average age. A potential source of the disparity in age stemmed from more years of use expected and higher activity level being selection criteria for the CoP and MoM implants. If patients within that group were expected to retain the implant for a longer period of time, then a bearing surface with a lower wear rate would be favorable with studies suggesting that CoP implants have significantly lower wear rates.

The data presented in the current study demonstrates that the CoP group has a significantly higher percentage of patients with radiographically identifiable wear, with a threshold of ≥2 mm of wear being recognized as clinically significant. This finding differs from reports that ceramic heads with polyethylene cups protect against wear more than metal heads against polyethylene. Some of the discrepancy

**Table 4. Radiographic assessment.**

<table>
<thead>
<tr>
<th>Implant</th>
<th>Stem osteolysis</th>
<th>Cup osteolysis</th>
<th>Poly wear ≥2 mm*</th>
<th>Stem subsidence</th>
<th>Calcar erosion</th>
</tr>
</thead>
<tbody>
<tr>
<td>MoP</td>
<td>0</td>
<td>0</td>
<td>2 (7.4%)</td>
<td>0</td>
<td>1 (3.7%)</td>
</tr>
<tr>
<td>CoP</td>
<td>0</td>
<td>0</td>
<td>12 (33.3%)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>MoM</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

MoP: metal on polyethylene; CoP: ceramic on polyethylene; MoM, metal on metal. *Fisher Exact, P=0.03.

**Table 5. Cost per implant at time of surgery (including stem, head and cup).**

<table>
<thead>
<tr>
<th>Implant</th>
<th>Cost, US Dollars</th>
</tr>
</thead>
<tbody>
<tr>
<td>MoP</td>
<td>4200</td>
</tr>
<tr>
<td>CoP</td>
<td>4600</td>
</tr>
<tr>
<td>MoM</td>
<td>5000</td>
</tr>
</tbody>
</table>

MoP: metal on polyethylene; CoP: ceramic on polyethylene; MoM, metal on metal.
in the current data could be due to the following factors. First, there is a significant difference among the follow up times, with CoP having the largest average follow-up among the groups which would allow more time for wear to occur. There is a large disparity between the gender makeup of the CoP and MoP groups. Furthermore, the aforementioned younger population of the CoP group could contribute to a greater amount of wear since a younger patient population would be expected to be more active. In spite of these findings, the CoP group does not differ from the other groups with respect to clinical outcome.

Given the similar clinical outcomes, the utility of MoM implants must be addressed since they do not have significant radiographically identifiable wear. MoM THA have higher revision rates associated with osteolysis. The reports on longevity of MoM implants are conflicting. Some reports show MoM implants have a higher survival rates than CoP, as well as MoP implants. However, there is still a concern over the formation of pseudo-tumors. Pseudo-tumors have been associated with all THA types, but those associated with MoM THAs are more aggressive.

Given the increased focus on healthcare spending, the cost of these bearing surfaces should also be considered when evaluating their worth alongside clinical outcomes. Kurtz et al. predicts that the total number of primary THAs will increase to 572,000 per year by 2030 in the United States. Furthermore, they go on to predict that THA revisions will increase by 137% between 2005 and 2030. The cost of the implants used in the current study (Table 5) varies by roughly 800 US Dollars from the least expensive to the most expensive, or approximately a 16% difference. This increase in implant cost is of questionable benefit given the similar revision rates of the three groups reported in the study.

The current study has some acknowledged weaknesses. The small size of the groups limits the power of the data analysis, it is a retrospective study, and there is no randomization of the patients to the different implant types. The lack of randomization limits the extent to which one can draw definitive conclusions from the data assessed as surgeon and patient preference may have confounded the data. However, the authors propose that the patient and implant selection criteria of these three groups represent common practice patterns of many orthopaedic surgeons in the United States.

Conclusions

Within any given surgical hospital, newer technologies and their respective manufacturing costs are often justifications for increased purchasing cost in expectations of improved clinical outcomes. The rates of pain, complications and need for revision secondary to articulating wear are similar among the groups compared here. The data presented in this study of one total hip system comparing three differing bearing surfaces with identical femoral stems, suggest that when using patient clinical and radiographic outcomes as endpoints, the increased hospital cost of the more expensive implants may not be justified based on midterm follow up.

References

24. Shimmin A, Beaulé PE, Campbell P. Metal-


