EPIDEMIOLOGY, COSTS, AND COST-EFFECTIVENESS OF LOAD-BEARING SURFACE CHOICE IN PRIMARY TOTAL HIP ARTHROPLASTY: CERAMIC-ON-POLYETHYLENE VS. METAL-ON-POLYETHYLENE

by

Keith James Carnes

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Health Services Research

Charlotte

2016

Approved by:

____
Dr. Jennifer Troyer

____
Dr. William Brandon

____
Dr. Christopher Blanchette

____
Dr. Susan Odum

____
Dr. Murray Webster

ABSTRACT

KEITH JAMES CARNES. Epidemiology, costs, and cost-effectiveness of load bearing surface choice in primary total hip arthroplasty: Ceramic-on-polyethylene vs. metal-on-polyethylene (Under the direction of DR. JENNIFER TROYER and DR.WILLIAM BRANDON)

In an age of cost-consciousness in health care, it is essential that we seek to understand the ramifications of the choices where costs and expected clinical performance differ. When the cost differences are evolving and performance differences are not yet clearly defined, it is even more important for physicians and patients to ponder the consequences of these shared decisions. The case of bearing choice in total hip arthroplasty is the quintessential case where the ceramic-on-polyethylene choice generally costs more, but is expected to perform better by having fewer failed hip replacements that require some 10-20 years for meaningful differences to emerge.

This study demonstrated that clinical problems in ceramic-on-ceramic and metal-on-metal surfaces, once thought to be the future of hip arthroplasty, have market consequences. Significantly reduced use and decreased cost premium of these surfaces compared to the cheapest alternative, metal-on-polyethylene, was observed over the study period. Further, in a constant press to increase the bearing life in patients, the use of ceramic-on-polyethylene surfaces grew exponentially across the 2007-2012 study period even before sufficiently rigorous evidence of performance that is just now starting to emerge.

This study has revealed that costs of these bearing choices vary widely across time and contextual variables. Analysis showing that more expensive, presumed longer lasting devices were used more frequently in younger, less sick patients that have more potential life years remaining to overcome the added cost indicated that cost-effectiveness principles are often considered. Cost-effectiveness of choosing the more expensive ceramic-on-polyethylene bearing surface was demonstrated to be highly dependent on patient age, implant cost difference, and level of improved performance of the ceramic-on-polyethylene implant. This study's use of the Premier Research Database's individual-level costs for both implants and the surgical procedure for almost 2 million patients and multiple literature sources for bearing surface performance brings new information and analysis to the question of spending more to get better results in total hip arthroplasty.

ACKNOWLEDGEMENTS

I would like to acknowledge Dr. William Brandon and Dr. Jennifer Troyer for so much of their time and wisdom in completing my dissertation and for their contributions to my education in general. I have enjoyed learning from them through now a second UNC Charlotte degree. I must also thank the UNC Charlotte Graduate Assistant Support Plan for tuition assistance and the Mecklenburg Medical Foundation for providing financial support through a graduate assistantship. I would like to thank Dr. Susan Odum and Dr. Thomas Fehring for the inspiration that led to my dissertation topic selection. Finally, I would like to thank Premier, Inc. and specifically, Scott Robinson for access to the Premier Research Database as the data source for this study.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

Constraining the growth in health care spending is a major objective of the Affordable Care Act. A primary vehicle to help achieve this goal is comparative effectiveness research wherein one assesses clinical effectiveness of alternate treatments for a given ailment.[1, 2] A logical next (or possibly overlapping) step is a cost-effectiveness analysis (CEA) that incorporates present and future costs as well as expectations of the length and quality of the patient's post-intervention life.   A CEA determines if an initially more expensive, but clinically superior alternate treatment is cost-justified by increasing the probability of patient well-being or "utility" in the quasi-technical language of economics.   The study reported here constitutes new cost-effectiveness research on one of the major components of US surgical cost, total hip arthroplasties.

Primary total hip arthroplasty (THA) is a common surgical procedure, comprising 1.3% of all non-diagnostic, non-pregnancy-related hospital procedures in the US in 2010, increasing to 2.6% of procedures in the $\geq$65 year-old population.[a][3] Despite its relatively high cost, the cost-effectiveness of this medical procedure has been demonstrated vs. available alternatives.[4-7] By the time the physician and patient agree on pursuing THA, the principal remaining clinical decision is which implant load-bearing surface materials to use. The decision is important because the upfront cost of the ceramic-on-

---

[a] Author's analysis of 2010 National Hospital Discharge Survey Results.

polyethylene (CoP) implant bearing surface is generally higher than the metal-on-polyethylene (MoP). However, almost all wear studies and most clinical survival studies indicate that the CoP option lasts longer than the MoP option, because it avoids or delays a costly and clinically demanding revision procedure involving the replacement of a failed hip prosthesis. The primary patient impact of the bearing surface choice is the serviceable duration of the implant and the resulting increased quality of life.

The research reported here is the first cost-effectiveness analysis comparing the two dominant bearing surface material combinations in the US today that used a Markov decision model for the CEA, nationally US representative administrative patient cost and clinical data, and revision rates derived from large population joint registries. The study used the Premier Research Database of hospital discharge data from 475 general, acute care, nonfederal hospitals across the US from 2007-2012, which provided cost item charge details and patient level clinical data.

This research began with an epidemiological description of the use of different types of bearing surfaces in the US between 2007 and 2012, attempting to find patterns in patient demographic and provider contextual variables and trends across time. This effort required multivariate logistic regression analysis to determine how demographic, contextual and clinical variables were associated with the choice of bearing surface. Understanding how the different articulations are distributed across geography, hospital descriptors, and patient demographic can help assess the scope of any clinical problems that present later.

The next phase was a detailed cost study of primary THA and the bearing surface choice. In this phase of the research, log-linear regression was conducted on the total hospitalization cost as a function of bearing choice, controlling for demographic, contextual, and clinical variables. Total volume of primary THAs, contextual variation in THA implant use, and trends across time and US geography in the cost of THA surgeries and implants can be used to help estimate national expenditures for these procedures.

National administrative databases are often underutilized for clinical or economic studies, particularly in the area of orthopedic surgery. Conventional wisdom includes a preference for the more precise, less biased results produced by prospective clinical/economic evaluations such as randomized controlled trials. The imperfect precision and insufficient clinical information in retrospective use of administrative data is sometimes thought to be less reliable. Although research on total arthroplasty has not often challenged that conventional wisdom by using administrative data for complex analyses, in 2013 Odum et al.[8] artfully demonstrated that the National Inpatient Sample database[b] (NIS) could be used to compare the cost-effectiveness of simultaneous vs. staged bilateral total knee arthroplasty; a complicated, but successful analysis that accounted for patient clinical and demographic covariates. A conventional study design would be ethically challenging with patients randomized into 2 treatment groups of sufficient size to allow cost and clinical data to lead to valid conclusions. In addition to the normal practical problems connected with this type of prospective study, such a design would also face ethical challenges because patient clinical factors that help

---

[b] The National Inpatient Sample (NIS) is an annually updated all-payer dataset of discharge-level patient demographic, clinical, and total charge records for all hospital discharges in a stratified 20% sample of critical care, non-federal hospitals in each participating state (48 in 2011, for example). Hospital-level weights allow nationally representative incidence estimates to be determined. The Health Care Utilization Project (HCUP) sponsored by the Agency for Healthcare Research and Quality (AHRQ) collects and distributes this and other important hospital discharge information.

determine the choice of staged vs. simultaneous would be difficult to incorporate into the random assignment of patients into treatment groups.

In 2005 and 2006, the Centers for Medicare and Medicaid Services (CMS) introduced ICD-9-CM surgical procedure codes describing the four main bearing surface choices, metal-on-polyethylene , metal-on-metal (MoM), ceramic-on-ceramic (CoC), and ceramic-on-polyethylene.[9] These codes allow hospitals to voluntarily report this information, which then made the surface choice a potential topic for large-scale research using administrative databases.

Specific Aims

1. Quantify the use of the 4 different load-bearing surfaces in primary total hip arthroplasty across appropriate hospital contextual and patient demographic variables between 2007 and 2012 in the US.
2. Describe variation in both the implant bearing surface choice cost and the total hospital stay costs between 2007 and 2012 in the US across hospital descriptors and patient demographic and clinical variables.
3. Conduct a CEA using a Markov decision model that compares the newer ceramic-on-polyethylene (CoP) to the older metal-on-polyethylene (MoP) bearing surface materials.

Chapter 2 provides a review of the literature on total hip arthroplasty as related to bearing surface and cost-effectiveness. Chapter 3 applies relevant theory to the specific aims of the dissertation. Chapters 4 through 6 are essays dedicated to the research questions asked and answered for each of the specific aims just described. Chapter 7 summarizes the research conducted and discusses policy implications.

CHAPTER 2: BACKGROUND

Although total hip arthroplasty is generally successful and very cost-effective, the failure of this surgery, which does involve a low risk of infection and other clinical issues, is a concern. Nonetheless, wear-and-tear, specifically aseptic loosening of the implant over extended use, is the primary source of lost post-surgery functionality. If the replacement hip joint loses enough functionality to cause an unacceptable state of patient well-being, a revision arthroplasty usually replaces the first implant. Revision procedures are more costly and difficult because the patient's remaining natural bone material and soft tissue has now endured a second failure and aging may have caused the patient's general health to deteriorate as well, thereby increasing surgical risk. Accordingly, significant research has focused on ways to increase the expected serviceable life of total hip arthroplasty procedures. Improved bearing surfaces have been the focus of much of this research. I evaluated the interaction of upfront cost and lifetime patient utility and cost of the choice of the load-bearing femoral (head) and acetabular (liner) surface materials.

The review of the literature will focus on the following areas to demonstrate the relevance of the specific aims:

- Trends in surface material choice
- Prior characterizations of bearing surface use
- MoP and CoP bearing performance
- Joint registries
- Hospitalization and hardware costs
- Cost-effectiveness and surface materials

Trends in Surface Material Choice

Four different load-bearing surfaces are currently used for total hip arthroplasty. The oldest is the metal-on-polyethylene (MoP) that has been the mainstay for these procedures since the 1960s due to its generally good short-term performance and acceptable long-term material wear and revision rates.[10] The term “hard-on-soft” is sometimes used to characterize the hard nature of the metal head and the softer nature of the polyethylene liner. The harder metal surface wearing out the softer polyethylene surface is the source of most long-term issues leading to revisions of MoP implants.[11, 12]

By the 1970s concerns about the impact of polyethylene wear on long-term performance prompted investigations into alternate materials, which initially focused on metal-on metal (MoM) and ceramic-on-ceramic (CoC) surfaces. These are often described as “hard-on-hard” surfaces. Replacing the soft polyethylene liner with a harder metal or ceramic material was expected to improve long-term performance through less material wear. Research using both hard-on-hard combinations found much lower material wear rates than for MoP as well as early indications of improved long-term liner life in vivo (in the body).[13, 14]

Unfortunately, significant short-term problems arose with both hard-on-hard surfaces as the use of these implants became more widespread. Premature fracture of the harder but brittle CoC liner component, even though rare, has been a concern.[15] Audible squeaking and clicking of the liner and head during normal use are also quite prevalent (4 to 20%).[13, 16, 17] Metal-on-metal bearing surfaces were plagued by permanently elevated ion levels in the bloodstream caused by corrosion[14] and growing concerns of local tissue reactions to metal wear debris.[18] Ultimately, the US Food and Drug Administration concluded that eroding metal can cause localized soft tissue damage,[19] leading to recalls of MoM components and systems in 2008-2012.[20]

Highly cross-linked polyethylene (XLPE) liners were developed to reduce wear with metal heads. Use of XLPE liners with Cobalt Chrome (Co-Cr) metallic heads had largely replaced conventional polyethylene liners by the mid-2000s.[21] Mid-term results using this articulation have been excellent compared to conventional polyethylene.[22, 23] Trunnionosis, corrosion and fretting at the junction of the Co-Cr head and metal neck of the femoral component, can cause catastrophic failure. This phenomenon was initially thought to be a problem only in MoM articulations but has more recently been observed in MoP bearing surfaces.[24, 25] Belief in better long-term polyethylene wear performance and the failure mechanisms associated with the alternate bearing options discussed here has driven the use of ceramic heads on highly cross-linked PE liners to grow substantially since the mid-2000s.[21, 26]

Prior Characterizations of Bearing Surface Use

The previous description of the evolution of bearing surface is somewhat anecdotal in nature, but efforts to more formally characterize use have been undertaken. Based on the first 15 months of NIS data (Oct 2005-Dec 2006) after voluntary reporting of bearing surfaces in hospital discharge records began, Bozic et al.[27] described use across hospital variables and selected patient demographic descriptors for 112,095 primary THA procedures. They found decreased use of non-MoP bearings in older patients, males, in urban non-teaching and rural hospitals, and in the northeast US region, with 48.8% of reported bearings overall being non-MoP. Unavailability of ICD-9-CM codes for CoP until October 2006 prevented the authors from describing CoP bearing use.

The authors also did not do multivariate analysis to investigate the factors driving the choice of bearing surface as has been conducted here. Further, they did not go into detail to assess potential nonrandom reporting/non-reporting of surface bearing that could potentially bias generalization of results.

Rajaee et al.[28] also used NIS data, focusing on bearing surface use in very young patients ($\leq$ 30 years old) from 2006 to 2009. Although sample size was very small (4,455 THAs; consistent with the age of the population), this rigorous study reported results in a similar manner as Bozic et al.[27] Non-MoP bearing use was much higher in this population at 77.9% than in the general population described earlier, with CoP prevalence growing from 1.6% in 2006 to 25.7% in 2009. Across this timeframe, MoP use changed little in the study sample, indicating that the bulk of CoP growth came at the expense of MoM and CoC surfaces.

Lehil et al.[29] used data from the Orthopedic Research Network (ORN)[c] to report details on bearing surface and femoral head size, along with other THA characteristics, on just over 100,000 THA procedures between 2001and 2012 in 174 hospitals. They reported movement away from CoC and MoM articulations, with CoP picking up the lion's share of the shift. This information is useful but the authors note they provided no details of hospital descriptions and therefore cannot assess the representativeness beyond the study population. They also reported average hospital implant device purchase prices across this time, with a low of $3,800 in 2001 to a maximum of $6,800 in 2007, and then declining prices to $5,842 in 2012. However, the authors provided no cost information on the specific types of bearing surfaces. Moreover, no mention is made as to whether these

---

[c] The Orthopedic Research Network annually reports aggregate data on arthroplasty procedures from hospitals that self-report implant details for the procedures and the prices they pay for components.

values were corrected for inflation during this 12-year time span which, in the general presence of positive inflation would tend to understate earlier reported costs.

MoP and CoP Bearing Performance

The challenge in clinically assessing THA hardware choices rests on the critical test of a particular bearing surface's value: its long-term performance. An implant failure that requires a revision surgery to replace the joint is the ultimate measure of an articulation's value, but it can take 10 years or more to discern performance differences. The need to estimate long-term performance led to the development of sophisticated in vitro (laboratory) wear simulation; automated testing over millions of cycles to provide component wear data more quickly. Sophisticated non-invasive methods to measure in vivo (in the body) liner wear were also developed to estimate long-term performance.

In 2000, Clarke and Gustafson offered an early critical review of the literature on CoP vs. MoP in vitro and in vivo wear rates.[30] Reviewing 6 articles, they found reported values of 1.5 to fourfold reduction in linear polyethylene wear rate for ceramic compared to metal heads. The authors also reported significant sensitivity to test conditions in observed laboratory wear rates: this finding might point to a reason why reliable translations of laboratory derived wear rates to actual revision rates in patients have not yet been developed.

In 2003, Hernigou and Bahrami[31] reported on a non-randomized, prospective polyethylene liner wear-rate study assessing two types of ceramic head (zirconia and alumina) with minimum 10-year patient follow-up. They compared 32mm alumina (56 hips) to 28 mm zirconia (40 hips) comparing to 20 hips each of 32mm and 28mm steel heads as controls. Liner wear in alumina (which would become the dominant ceramic material) was lower than zirconia and outperformed steel in the larger diameter head but was not significantly different from the smaller diameter steel head.

In 1996, Wroblewski et al.[32] reported on a prospective study (19 hips, 6.5 years duration) on the wear rates of cross-linked polyethylene liner with ceramic heads. They reported favorable liner wear rates for the ceramic heads (3-fold reduction vs. steel), but their comparator was a collection of values reported in the literature for metal heads with normal polyethylene liners. The authors also reported on a simultaneous lab wear simulation study that closely aligned with the clinical results. In another prospective study, Urban et al.[33] reported on 64 ceramic-on-poly hip implants followed between 17 and 21 years. They reported survivorship of the hips (95% at 5 years post-THA, declining to 75% at 20 years). They compared wear rates to literature values for steel heads (ceramic had 1.1 to 4-fold reduction in annual liner wear rate). Both of these studies suggest that ceramic implants perform better, but the lack of directly observed comparison treatments makes these two studies less compelling.

Some more recent studies directly compare CoP to MoP. In a 2010 retrospective study using the Norwegian Arthroplasty Registry, Hallan et al.[34] reviewed 9,113 primary THA components implanted with a maximum follow-up of 20 years. They reported several different types of component comparisons, limited to normal polyethylene liners, and observed that alumina ceramic heads resulted in longer life than did steel (1.8-1.9 relative risk of acetabular revision due to liner failure).

Several prospective studies targeted younger patient populations (age 50-55 years at index surgery) to compare CoP and MoP. Ihle et al.[35] reported on a non-randomized study that involved implanting 80 alumina (ceramic) and 13 chromium heads (metal), all with normal polyethylene liners; the 93 surgeries were then followed for an average of 19 years. Six of 13 hips with metal heads required revision vs. 11 of 80 for the ceramic. Liners with ceramic heads also had approximately 2x less measured wear in all patients as measured by x-ray photography. Meftah et al.[36] followed 31 well-matched pairs (similar age, body weight, diagnoses, and activity levels) of hip arthroplasties performed between 1989 and 1992, with alumina ceramic and metal femoral heads being the operative difference; all used normal polyethylene liners. After an average follow-up of 17.7 years, annual liner wear rate for the ceramic heads was 40% lower than for metal. Although not statistically significant, 3 of the metal hips required revision while 1 of the ceramic hips required reoperation.[d] Wang et al.[37] prospectively compared ceramic to metal via performing bilateral primary THA on 22 patients between January and December 2002, with a different bearing surface used in each hip. After 10 years of follow-up, the polyethylene liner wear rate averaged 60% lower against ceramic heads than against steel.

Not all studies reveal improved performance for ceramic. Nakahara et al.[38] conducted a prospective randomized study on 102 hips split equally between alumina ceramic and cobalt-chromium heads and reported no difference in PE liner wear after 5-8 years of observation. Kawate et al.[39] conducted another prospective, randomized study on 32 zirconia heads and 30 cobalt-chromium heads and also found no reduction in PE liner wear after 5 years. Both authors acknowledged that short follow-up time may have contributed to the findings.

---

[d] “Reoperation” is sometimes used interchangeably with “revision” but the former is also often used to describe a surgical procedure that does not involve replacing the implant (personal conversation with Susan Odum of OrthoCarolina Research Institute (OCRI)). The authors’ context implies distinguishing between a traditional wear-driven loosening failure that leads to revision and a non-wear related failure.

Kadar et al.[40] gave a preliminary report on an ambitious randomized controlled trial wherein groups of 30 THA patients, between November 2004 and June 2007, were assigned to one of 5 treatment groups: ceramic and metal heads, each combined with high molecular weight and cross-linked polyethylene liners, respectively, using the original Charnley hip design as a control. Their initial report at 2 years follow-up indicated that the cross-linked liner was the dominant factor in reducing wear rate, of greater importance than the head material. They did acknowledge the very short follow-up time assessed so far and that results may not hold up over time.

Virtually all lab wear evidence and the majority of in vivo wear evidence indicate superior performance of alumina ceramic heads with polyethylene liners. Revision data of sufficient duration comparing metal to ceramic heads is starting to be published from randomized trials and registry data is beginning to be used as well. The growth in use of cross-linked PE liners may prove to mute some of the observed differences between metal and ceramic, but longer observation periods of relevant comparisons will be needed to define the long-term revision rate differences between CoP and MoP.

Joint Registries

Long-term orthopedic clinical trials are costly and difficult to conduct because 10 or more years may be needed for revision rates to show meaningful differences between implant types or bearing surfaces. Generally, an alternative source of meaningful data is a hip arthroplasty registry, a prospective collection of a large sample (ideally a census) of THA procedures that begins when the registry is created. Patients are enrolled continuously and the enterprise continues until the registry sponsor decides to stop enrolling patients and collecting data. The clinical information can be as comprehensive as desired—or at least as extensive as collection costs justify; once entered, hip survival data is collected continuously; and the database is regularly updated at least annually. From these data, Kaplan-Meier (K-M) cumulative probability of first revision graphs can be created. Registries contain many more implants than even the largest prospective clinical trials and, if properly executed, can generate the performance measures that are necessary to evaluate different specific components of the procedure. A joint registry affords collection of effectiveness data in a virtually uncontrolled environment and thus offers advantages over the efficacy data that is collected in tightly controlled randomized trials.

The National Joint Registry (NJR) for England, Wales, and Northern Ireland, which is in its 11th year, receives reports on all hip and other joint implants in those regions of the UK, and stores a broad range of very detailed information.[41] Among its many tables and figures detailing incidence and performance, Table 3.7 (p. 132) provided K-M cumulative first revision probability graphs for the primary total hips by bearing surface. Review of 10 years of available registry data indicated that cumulative probability of first revision is ~25% lower for CoP than MoP bearing surfaces (across 110,000 MoP and 19,000 CoP hips), but varies across different specific articulations.

The Norwegian Arthroplasty Registry (NAR)[42] operates differently from the UK registry just discussed. The NAR entered new joint replacements into their database from 1987-2007 and published reports between 2000 and 2010. Its last report included much incidence information on a variety of arthroplasty-related issues but no bearing surface information and very little performance information. However, the NAR's raw data does contain these types of details and have been made available to researchers (including studies cited in this review).

The final joint registry considered here is the HealthEast Joint Replacement Registry (HEJRR). HEJRR is the oldest operating joint registry in the US; it collects joint replacement information from the greater metropolitan St. Paul-Minneapolis, MN area and has been in operation since 1991. HEJRR continues to collect follow-up data on over 30,000 joints. Figure 11 of the HEJRR report provides THA survival information on the bearing surfaces of interest.[43] The report indicates 87% and 89% cumulative survival probabilities at 15 and 17 years, respectively, for 2 different MoP articulations but data have not been collected long enough on CoP hips for survival rates longer than 4 years to be estimated.

Registries can provide the best information to assess effectiveness in use; however, registries vary widely in what they publish. The K-M figures are helpful to define differences between different hip configurations, but providing access to the raw survivorship data allows researchers to determine variances in the data necessary for probabilistic decision modeling as well as control for potentially confounding variables available in the data.

Hospitalization and Hardware Costs

Two different types of data are generally used to report costs incurred during hospital stays. Articles using both types of data will be discussed here noting positives and negatives of each approach.

Some studies use clinical and patient financial records procured directly from hospitals. One example, which reported on 491 consecutive hip replacements (primary and revision) performed by 2 surgeons at one hospital between 2000 and 2002, described total and department-level hospital resource utilization.[44] Another example is a study that reported on ~5,000 THA procedures conducted in 2008 at hospitals involved in California-based purchasing initiatives.[45] This latter study described variation in hospital costs and reported on the degree to which hardware, patient characteristics, and hospital characteristics contributed to variation. Both studies provided rich detail and useful information on large populations but suffered from the common weakness of limited geographical range and failure to specify bearing surface of the implanted devices. Only with the relatively recent advent of the ICD-9-CM bearing surface procedure modifier codes can large, publicly available national databases such as the HCUP data discussed earlier be utilized to assess questions relating to bearing surface.

Medicare hospital claims databases are considered the gold standard for many types of studies, due to their rich detail and the virtually complete coverage of the Medicare-eligible population. Medicare has and will continue to have great importance as a payer in total hip arthroplasty. A more recreationally active, albeit heavier, baby-boomer population is undergoing THAs at younger ages,[46] i.e., in advance of Medicare coverage; consequently, it is important that databases used for THA research capture cases that have been reimbursed by private insurance as well. An interesting study used the NIS wherein the authors assessed the impact on total costs and device price/Medicare payment ratio of total procedure volume and implant selling price during the 1995-2005 period.[47] The study used published trade information for device pricing and the NIS charge data to develop the analyses of surgical costs. After adjusting for inflation, the authors found that THA device purchase prices increased by 24% across their study period. Further, they found that the implant price increased from 29% to 60% of the total procedure cost. Medicare and NIS data can be useful for national studies, but provide no detail on bearing surfaces prior to 2005 nor hardware-level costs critical to studies as reported here.

Supplementing their epidemiological study described earlier that reported on THA bearing use in very young patients, Raijee et al.[28] also reported average hospitalization charges by bearing surface using NIS data across 2006-2009. They observed a narrow range of average charges across MoP, MoM, and CoC bearings at $51,170-$51,382, while reporting a much higher value of $58,614 for CoP bearings. The authors' use of total charges rather than costs, which could have been determined from NIS data, limits the study's value. The authors also did not control for any patient or hospital variables, which could have produced different comparative charges (and costs) between the different bearing surfaces.

Also following up on an epidemiological study described earlier using the ORN, Lehil et al.[29] reported average hospital implant device purchase prices, with a low of $3,800 in 2001 to a maximum of $6,800 in 2007, and then declining average price to $5,842 in 2012. However, the authors provided no cost information on specific bearing surfaces and no mention is made as to whether the reported prices had been corrected for inflation across this 12-year span.

In light of the historic unavailability of implant level cost data on a wide scale from administrative databases, one can turn to trade literature for this information. The Orthopedic Research Network (ORN) had 165 participating hospitals that accounted for ~22,000 hip surgeries in 2012 and has been used in several studies cited in this review.

The ORN collects and aggregates information on list price (reported publicly by manufacturers), average selling price to participating hospitals (reported confidentially by hospitals), and volume data for arthroplasty components. The 2013 report published by Orthopedic Network news showed that the average CoP implant package cost ~$600 more than the average MoP package in 2012.[21] This source is very informative and rich in THA details, but it is a considerably smaller sample and only began specifically reporting CoP and MoP surface use details in 2012.

Cost-Effectiveness and Surface Materials

The majority of the meager clinical literature available indicates that CoP outperforms MoP, but at a generally accepted higher initial cost. Thus, a decision to utilize a more expensive CoP articulation should be justified by greater expected future utility and/or reduced future revision costs because the CoP choice can be expected to last longer than MoP. Three studies addressed cost-effectiveness of bearing surface options more expensive than the base case metal-on-poly but only one specifically compared CoP to MoP, the critical comparison in current clinical decision-making.

In 2006 Bozic et al. reported on a CEA comparing alternate bearing surfaces to MoP.[48] All model inputs came from the literature and the authors concluded that, alternate bearings may or may not be cost-effective depending on what 20-year survival rate improvement was used in the model for alternate surface, the age of the patient, and the initial cost difference of the implant choice. A significant shortcoming of this study is its use of only % differences in revision rate (baseline 20% revision improvement of alternate surfaces over MoP). Although using a percentage difference is a logical way to compare implant performances, without defined revision rates for either surface compared the audience has difficulty putting the improvement required for cost-effectiveness in context. This study did create a useful framework for reporting CEA results as point estimates of input variables are varied.

Gioe and Tatman used HealthEast Joint Registry data to conduct a retrospective analysis using implant cost and survival data from 2002 through 2008.[49] They compared 1,311 hips with premium CoP, MoM, or CoC surfaces to 868 hips with MoP surfaces. They concluded that more expensive implants (Δ = ~$1,000) were not justified because cumulative revision rates were not significantly lower for the higher priced group (although the revision rates were observed to be lower). While this sample size is quite large compared to prospective trials, the short evaluation timeframe and decision to lump all the non-MoP surfaces together obscured differences between the individual surfaces that might be present if the higher-priced hip articulations were considered separately. For instance, CoC and MoM surfaces are more prone to early, non-wear-related failures than is either a MoP or CoP surface; this difference is especially liable to obscure any differences between CoP and MoP when the time horizon is only 7 years. In this case, aggregating the results lost potentially valuable information specific to each bearing surface.

Pulikottil-Jacob et al. (2015)[50] reported on a cost-effectiveness analysis that used United Kingdom National Joint Registry (NJR) revision results comparing several different articulations, including a direct comparison of cemented MoP to cemented CoP.[e] The authors report that Markov modelling shows neither sufficient utility gain nor enough reduction in lifetime costs to declare CoP more cost-effective than the base case

---

[e] Hip replacement hardware can be placed with our without cement to secure the implanted components. In the case of CoP vs. MoP considered here, both femoral and acetabular components were cemented in place.

(most prevalent) of cemented MoP, although CoP had the highest probability of being cost-effective. The authors acknowledge that their study was limited by short-term revision data (7-8 years). A significant percentage drop in CoP revision rate was seen (20-30% lower than MoP by my visual observation of the authors' K-M graphs). Despite this large observed improvement, the base MoP revision rate was so low that the algebraic difference in actual revisions was too small for the Markov models to demonstrate significant value from a more expensive CoP implant.

Summary of the Literature

The fast-changing world of bio-engineering leaves researchers struggling to find or collect suitable data of adequate size and sufficient detail to examine the latest achievements. This is particularly true for total arthroplasty, as innovations are often put into practice prior to the 10+ year long-term clinical trials required to fully determine the performance differences. The problem is exacerbated in THA by significant problems that arose years after the hopeful introduction of MoM and CoC articulations. Researchers and surgeons were left with only CoP as an alternative to MoP in attempts to improve long-term performance. The growth of joint registries (in both population size and just as importantly, elapsed time since index procedure) is already beginning to influence how researchers conduct investigations; as must be the case for both clinical and cost-effectiveness assessments.

The challenge now is for a researcher to advance the state of knowledge by focusing more narrowly on the epidemiology and costs of bearing surfaces and THA procedures, assessing across time and contextual variables, and conduct a cost-effectiveness analysis of the competing medical devices-CoP and MoP.

CHAPTER 3: THEORETICAL CONSIDERATIONS…AND REALITY

Economic analyses of decisions that require decades for net effects to be realized are a challenge. The simplest analysis requires long-term assumptions that inevitably are subject to second-guessing as well as changing conditions. Health care decisions of this type are also impacted by the multiple entities involved in paying the bills, incurring the benefits, and a clinical decision process that is frequently not straightforward. Chapter 3 discusses in detail how these issues affect the economic analysis of bearing surface choice in THA and how results are operationalized.

Clinical choices made by doctors and patients are often complex, particularly when they include cost considerations in a world where neither reliable cost nor price information is usually at hand. Bearing surface choice in THA is even more complicated because any added quality improvements generated by incurring additional cost will likely take years or even decades to be demonstrated and the implant options available to the surgeon may be limited by the hospital. Earlier conceptual efforts to frame the economic and quality issues have made investigations into these choices easier. However, the questions of the appropriate perspective and the relationship between health and health care remain especially problematic.

Grossman used the theory of human capital to explain how individuals consider health vs. health care[51, 52]; it was extended by Folland et al.[53(p.150)] to explain how the demand for health and health care differ. The insights of Folland et al. can be briefly summarized:

1. Consumers want health and demand health care to get it.
2. Consumers are not merely passive purchasers. They produce health as well, by investing time for personal health improvement and purchasing health services.
3. Health normally has longer duration than just the current period and can therefore be analyzed as a capital good.
4. Health is both a consumption good and an investment good.

Total hip arthroplasty decisions have features of both consumption and investment decisions. The decision to undergo a THA, or forgo the procedure with the consequence of continuing to suffer pain and lost mobility despite significant and costly management efforts, has characteristics of a consumption decision in that there is high probability of virtually immediate gain in functionality. Prior research showing that THA has a positive long-lasting effect on health and superior long-term cost-effectiveness[4] indicates it can be a beneficial investment decision as well. The bearing surface choice is entirely an investment (aka capital) decision because no short-term differences are apparent, but long-term performance differences between ceramic-on-poly vs. metal-on-poly can be expected. The CEA conducted here addressed whether incurring a greater incremental cost for ceramic at the time of surgery will be outweighed by increased patient utility over the remainder of the patient’s life (longer healthy life before revision) and/or reduced cost (delaying revision procedures reduces lifetime costs in present dollars).

The CEA conducted here used several types of cost information and patient utilities to offer a societal perspective. These factors are described below:

- Non-implant hospitalization costs associated with THA procedures,
- cost of implant hardware associated with THA procedures,
- literature values for revision rates and patient utilities,
- the total hospitalization cost associated with revision procedures,
- rehabilitation costs associated with THA and revision procedures,
- surgeon fees associated with THA and revision procedures, and
- anesthesiologist fees associated with THA and revision procedures.

The costs used in this study represent the most significant societal clinical costs. The analysis specifically distinguishes between those costs that vary with the bearing surface choice (index THA hardware cost) and those costs that do not vary at the time of index procedure (all others), but impact the analysis based on revisions occurring at different times for the different surfaces used at index THA. Further, the analysis assumes and incorporates no difference in health utility values between bearing surface immediately following index THA. These assumptions are consistent with previous bearing surface CEAs discussed in Chapter 2 and available literature on health utilities.

The analysis does not account for some frequently considered societal factors, notably lost wages and travel costs. Bearing surface choice does not impact lost wages or travel costs at the index THA procedure, so they do not incrementally impact the upfront cost determination. Lost future wages based on when revisions occur can have an impact but the age and workforce participation of the subject population minimizes the impact. The median age at index THA for the subject population is 66 years old[f] and workforce participation for the 65 and older population was 21.7% in 2014 and will head sharply

---

[f] Author's analysis of the Premier Research Database.

downward[54] across the 10-20 years before significant numbers of implant failures start to occur. Accordingly, the impact of lost wages following a revision procedure will be minimal for the bulk of the study population. Travel costs are excluded, because their impact will be minor compared to the other large factors already included.

How the economic analysis results are eventually operationalized depends upon the nature of the health insurance coverage (for example, patient co-pay vs. insurer-incurred costs), negotiating power of the hospitals that actually incur the incremental cost of ceramic heads, and society's threshold for incremental cost-effectiveness. For example, the choice of bearing surface hardware does not result in different Medicare payments to hospitals,[g] but payments may vary across private insurers and Medicaid; patient co-pays may vary as well in Medicaid programs.

Health Care and Rational Choice

The realities of delivering and financing health care in the US skew how the results of economic analysis of bearing surface choice are put into use. Patients, physicians, hospitals, and insurers are distinct but inter-related entities, each playing a critical role in the decision-making process related to THA in general and to the economics of surface choice specifically. These factors are discussed next.

Compared to many medical situations, the patient's decision to get a hip replaced is straightforward. Stripped of the issues of access, including financial access, the patient and her provider-agent must decide whether the relief granted by THA is worth the required cost-sharing and the risks associated with major surgery. The large and growing prevalence of total joint replacement indicates that this "trigger" is being pulled very

---

[g] During the initial phases of new technology offerings, Medicare sometimes pays a premium to promote adoption of the innovation. The premium to encourage use of newer bearing surfaces over MoP has expired. (Personal conversation with Susan Odum of OCRI.)

often in the US. The next question, which bearing surface to choose, is more subtle and requires more guidance from the surgeon. Although the data are limited, ceramic-on-poly is increasingly regarded as the better clinical choice to delay revisions compared to metal-on-poly, particularly with the high failure rates with metal-on-metal and ceramic-on-ceramic bearings described earlier. Physicians receive the same surgical fee regardless of the bearing surface although the physician's role as "agent" to the patient's role as "principal" can possibly conflict when a physician has designed an implant and receives royalties. Even if the surgeon is a diligent and selfless agent, recent research suggests that few orthopedic surgeons know the upfront hardware costs of the devices they implant.[55]

Moreover, the incremental costs incurred by patients based on the choice can vary as discussed earlier. In most cases, the patients will not know the difference in co-pay for different surfaces or whether there is a co-pay difference at all. Even if they do, the actual difference felt from the typical 20%-30% co-pay may not be great enough to constrain additional spending in cases where it is not justified. In light of the confusing, conflicting and unclear cost-sharing information, the societal perspective that distinguishes CEA from self-interested profit maximization and/or utility provides the ethical grounding necessary to propose solutions to problems normally managed within the doctor-patient shared decision-making process.

The situation becomes more complicated when one considers that the hospital purchases and incurs the cost of arthroplasty hardware inventory until it is needed for use. Most hospitals will consult their physicians about what items to stock, but administrative pressures are likely to lead hospitals to reduce the number of different items stocked.

Eventually, cost of the hardware is likely to influence what implants hospitals choose to purchase altogether.

Hospital responsibility for inventory costs, varying physician knowledge of costs, and varying/poorly understood patient cost-sharing can easily get lost in a cost-effectiveness analysis; yet, they are key factors in the real-world bearing surface decision process. Moreover, a hospital will receive the same reimbursement for all surface choices in most cases and cannot expect to benefit from a patient who experiences longer lasting implants and delayed revisions. At the margin, the hospital, which stocks more expensive, longer lasting devices, can suffer financially a second time by having helped the patient enjoy good hip health longer if a patient delays or avoids the revision that he might have undergone at the same hospital. Even the most aggressive accountable care "value purchasing" incentive program will find that the lengthy timeframe to observe ceramic- vs. metal- differences makes it difficult to incorporate payments for better outcomes of THA.

Insurers also play an essential role in almost all health care decisions. As stated earlier, Medicare does not cover the incremental cost of more expensive bearing surfaces and private insurance often follows Medicare's lead. Sometimes however, device reimbursement and coverage differ between Medicare and private insurers. A recent study[56] found that across 47 devices (spanning a wide range of clinical areas), the 16 largest US private insurers that publicly reveal their coverage policies agreed with Medicare half the time, were more restrictive one quarter of the time, and were less restrictive one-quarter of the time. Hypothetically, for an insurer considering alternate bearing surface costs (whether for a specific patient or as policy), the decision to spend more is difficult, because as a practical matter a private insurer is unlikely to reap the benefit of the incremental expenditure. Any benefit of higher quality is likely to accrue to another insurer which will have to pay for a revision 10 or more years later.

The reimbursement system obscures and does not efficiently reward decision-makers; in the contorted financing system short-run profit maximization is the only rational strategy. The objective analysis of a CEA asks whether incurring extra expense upfront for a potentially better choice results in less total expenditure and/or increased utility for the patient, independent of which entity reaps the financial benefit. The societal perspective is the most neutral among the entities incurring costs and/or financial benefits, thereby overcoming many of the reimbursement issues previously raised. The societal perspective is the appropriate stance for a CEA, because it alone internalizes all costs and benefits (or effectiveness measures) by incorporating all negative and positive externalities along with individual actor's utility maximization.

In 1955 Nobel Lauriat Herbert Simon[57] described rational choice as the process by which "rational" individuals weigh options and make choices according to their expectation of their greatest present and future utility. The choice of bearing surface in THA fits this theoretical model. Unfortunately, the various entities discussed earlier are not vested in making a decision that benefits society, nor can the patient be expected to align his or her "rational" interests with those of society. Determining how society's interests are protected is the purview of health economists. In a world of scarce resources, society's best choice is the one with the lowest incremental cost-effectiveness (ICER) ratio. The ICER is a measure of how many dollars each additional year of life at full health costs society, when comparing two alternative treatments. Cost-effectiveness analysis is a formal analytical tool that can help the decision-making. Its value is in providing discipline to counteract the fragmented and impressionistic thinking characteristic of humans--whether they are patients, profit-driven corporate managers, or physicians.

In the broad world of health economics, the decision about which bearing surface to use is not a particularly complicated one: does the delay or avoidance of a well-defined failure point (need for revision) offset increased upfront cost for a better bearing surface? The varied ways in which the US chooses to finance health care imposes burdens on this decision process that make it complicated. This section has demonstrated how and why the cost-effectiveness analysis conducted here can help overcome these burdens to facilitate the best decision process, thereby helping to protect the interests of both society and patients.

Other Hypotheses to Consider

In addition to providing information for the CEA and for detailed descriptive studies of bearing surface use and cost, the Premier data set can be analyzed in ways that shed light on issues of technology dissemination and the nature of current medical device markets. A brief review of these research issues is offered here, but detailed specification of the analysis, results, and implications will be found in Chapters 4 and 5.

Conventional wisdom has long assumed that academic medical institutions are early adopters of new technologies. A logical extension of the epidemiological description of bearing use in the first specific aim is to test this anecdotal claim by measuring the association between use of the different bearing surfaces and teaching status and other characteristics of the hospitals where the THA procedures were done, while controlling for patient and payer factors. If conventional wisdom holds true, the academic institutions should adopt non-MoP surfaces faster and with greater prevalence.

Just over half of all THAs in the study population have no indicator of bearing surface noted (reporting is voluntary); therefore, it is possible that the device types that are reported are not representative of those that are not reported.[h] The data are examined to determine if reporting is associated with insurance payer, patient descriptors, and hospital descriptors. Absence of observed bias will help bolster the argument that the epidemiological results can be generalized to produce national estimates.

The second specific aim explores how the costs of THA surgery and bearing surface hardware vary across time, hospital and patient variables. If one uses hardware costs in the data set as an indication of prices paid by hospitals, the market relationship between purchasers and sellers of these devices can be explored. Specifically, economies of scale would generally result in larger purchasers of a given type of item being able to negotiate lower prices, given that there are multiple suppliers in the arthroplasty device market. The Premier national research data base allows several indicators of hospital size and surgical volume to be determined: procedure volumes give a direct indicator of purchase volume of bearing surface hardware. Thus, we can test whether hospital size and arthroplasty volumes are associated with reported hardware cost (as a proxy for price paid).

Group purchasing cooperatives that assist hospitals to attain competitive purchase prices are also prevalent, but available data does not allow testing of how group purchasing practices might affect hospitals' bearing surface purchase prices. It is also not

---

[h] Author's analysis of Premier Research database indicates approximately 46% of surgeries included the voluntary ICD-9-CM procedure code that indicates which bearing surface was used.

possible to test impacts of hospital mergers, particularly the recent evolution of large integrated health care systems purchasing hospitals.

This chapter has provided the reader a discussion of ways in which classical economic and market theory provides important background for this research. Specifically, these theories suggest worthwhile avenues of research into the complex issues of the choice and purchase of the different bearing surfaces used in total hip arthroplasty. The peculiarities of health care financing and delivery makes economic considerations challenging, but not impossible.

CHAPTER 4: FACTORS ASSOCIATED WITH BEARING SURFACE CHOICE IN TOTAL HIP ARTHROPLASTY

Primary total hip arthroplasty (THA) is a common surgical procedure that most often is employed to overcome steadily worsening pain, functionality and quality of life due to osteoarthritis of the hip joint. In 2010 THA comprised 1.3% of all non-diagnostic, non-pregnancy-related hospital procedures in the US, increasing to 2.6% of procedures in the $\geq$65 year-old population.[i][3] Despite its relatively high cost, the cost-effectiveness of this medical procedure vs. available alternatives has been demonstrated.[4-7] The original Charnley hip replacement design introduced in the 1960s, using a metal-on-polyethylene (MoP) bearing surface articulation, has demonstrated upwards of 80% survival 20 years after original implantation.[58] Nonetheless, efforts to improve on the original design have continued in a number of areas. Improvements to the bearing surface materials have been at the forefront of these efforts.

This paper will detail the recent evolution of use of the four dominant bearing surface materials used in THA: MoP, metal-on-metal (MoM), ceramic-on-ceramic (CoC), and ceramic-on-polyethylene (CoP), the most recent surface material innovation. Figure 4.1 shows the components involved (metal-on-polyethylene example), demonstrating where and how the materials just noted are used. Previous efforts to characterize bearing surface use[27,29,28] (detailed in Chapter 2) were informative, but

---

[i] Author’s analysis of 2010 National Hospital Discharge Survey Results, conducted by the Centers for Disease Control and Prevention. http://www.cdc.gov/nchs/fastats/inpatient-surgery.htm accessed 9 January 2016.

suffered from one or more of the following limitations: narrow geographic coverage, small number of hospitals contributing cases, narrow populations studied, lack of descriptive information, only bivariate analysis reported, or short length of time studied. The study reported here overcomes these limitations in addressing the following research question:

> What were the drivers of use of the 4 main bearing surface combinations in primary total hip arthroplasty in the US across time, hospital, and patient variables from 2007-2012?

Consistent with the general belief that teaching hospitals[j] are early adopters, it is hypothesized that greater prevalence of newer, non-MoP surfaces will be seen in these hospitals. Further, it is hypothesized that whether or not a given procedure has the bearing surface reported will be randomly distributed across the available and tested patient, hospital, and physician variables. Finally, it is hypothesized that, across the study period, CoP use will grow substantially and that CoC and MoM use will decrease.

Methods

Data Source

In 2005 the Centers for Medicare and Medicaid Services (CMS) introduced International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) surgical procedure codes making it possible to denote each of the bearing surface choices investigated here.[9] These codes allow voluntary reporting of the bearing surface for each procedure, which then made the surface choice a potential topic for nationally representative research using large-scale administrative databases. This study used the

---

[j] Per the Healthcare Cost and Utilization Project, teaching status is conferred on a hospital if it has “an AMA-approved residency program, is a member of the Council of Teaching Hospitals (COTH), or has a ratio of full-time equivalent interns and residents to beds of 0.25 or higher.” http://www.hcup-us.ahrq.gov/db/vars/hosp_bedsize/nisnote.jsp accessed 9 January 2016.

Premier Research Database (PRD; Premier, Inc., Charlotte, NC) of hospital discharge data for 2007-2012 from general, short-term, nonfederal, acute care US hospitals that participate by choosing to provide patient level demographic, cost and clinical data, as well as hospital demographic information.

Premier is a large group purchasing organization (GPO), with approximately 3,400 participating US hospitals.[k] The PRD includes a group of hospitals, some of which are GPO members and some not, that participate in a range of quality improvement activities with Premier. Participating PRD hospitals provide clinical, demographic, and cost data for a 100% census of all their discharges. Across the study period of 2007-2012, 523 hospitals spanning 44 states reported hip arthroplasty procedures. These procedures were the data source for this study. Premier calculated and provided weighting values for all discharge records based on comparing its mix of hospitals to those reported in the American Hospital Association Survey,[l] incorporating hospital staffed bed count, location, rurality, and teaching status; thus nationally representative prevalence estimates can be made. This THA study period began in 2007 the first full year that ICD-9-CM procedure codes were available for reporting of all four bearing surfaces of interest.

Sample Characteristics

Primary THA case records for patients ≥ 18 years of age were selected from the PRD based on one incidence of the 81.51 ICD-9-CM procedure code and no other hip arthroplasty codes present except for ones denoting bearing surface. Specifically,

---

[k] As reported by the Health Group Purchasing Industry Initiative, a trade association of eleven of the leading GPOs, http://www.healthcaregpoii.com/signatorycompanies/premier.html accessed 7 January 2016.

[l] The American Hospital Association Survey is conducted annually and collects over 1000 data fields describing hospitals, including the hospital-level descriptor variables used in this study. http://www.ahadataviewer.com/book-cd-products/aha-survey/ accessed 8 January 2016.

bilateral THAs, revisions and resurfacing procedures were excluded from this analysis. The bearing surface used in a THA was determined by the additional presence of one of the following ICD-9-CM procedure codes: 00.74 (MoP), 00.75 (MoM), 00.76 (CoC), or 00.77 (CoP). In order to estimate the total number of THAs conducted nationally, discharge records missing the voluntarily reported bearing surface were included as well to create a sample data set of 278,179 primary THAs described in the PRD from 2007 through 2012. Approximately 46% of all discharges in this study included a bearing surface procedure code. This is slightly better than the 41% bearing surface reporting rate seen in the National Inpatient Sample (NIS).[m]

Analytic Methods

The study conducted here began with descriptive characterizations of bearing surface use in primary THAs across PRD-provided variables that provide the year of the surgery, patient and hospital demographic indicators, and patient clinical indicators. The Agency for Healthcare Research and Quality (AHRQ) Healthcare Cost and Utilization Project (HCUP) has suggested logical groupings of hospitals by hospital bed count, rurality of location, teaching status, and US census region (all provided in the PRD) to facilitate comparisons by uniform hospital characteristics. Rurality and teaching status were combined into rural, urban teaching, and urban nonteaching. Hospital size strata were created as small, medium and large based on HCUP definitions that incorporate the number of staffed beds, hospital location rurality, teaching status, and US Census region.

Bivariate tables were created showing choice of bearing surface (MoP, MoM, CoC, or MoP) by the year of procedure, patient descriptors, and hospital descriptors

---

[m] The NIS is a 100% census of all discharges in each hospital of a sample of approximately 1,100 general, short-term, nonfederal, acute care US hospitals (~20% of total), stratified to be nationally representative. The author's analysis indicated 41% reporting of bearing surface in the 2011 NIS dataset.

where the weight values were used to create national prevalence estimates. Chi-square statistics were employed to determine if deviations in proportion from the expected value in the bivariate tables were significant and thus denoted a relationship between the independent variables (year, patient and hospital indicators, tested individually) and dependent variable proportions (bearing surface). The expected value is the population proportion of each bearing surface in the sample studied with a null hypothesis that there is no difference between observed and expected proportions. The Chi-square statistical tests were judged to be statistically significant at $p<0.001$.

The PRD provides many patient- and hospital-level indicators. Patient demographic variables considered include year of procedure, age, sex, and primary payer. Hospital type, hospital size, and US census region were used to describe conditions of hospitalization. Primary diagnosis, Medicare Severity Diagnosis Related Group[n] (MS-DRG), AHRQ comorbidity count,[o] and patient discharge status were used to characterize patient clinical condition as a proxy for health status.

Finally, multivariate regression was conducted to achieve two purposes. First, the reported bearing surface was regressed against hospital and patient variables to determine which associations seen in bivariate analyses remained significant when all available information was considered in one model. Next, the ability to produce meaningful national estimates was assessed by creating an indicator noting if the bearing surface was

---

[n] Developed originally for prospective Medicare reimbursement, the MS-DRG is a value assigned to each inpatient stay that incorporates the patient's primary diagnosis, secondary diagnoses, the principal procedure and any additional procedures, sex, and discharge status. MS-DRG values related to THA are 469 and 470, which stand for major joint replacement associated with lower extremity with and without, respectively, major complications and comorbidities.

[o] AHRQ offers software that can be used to review the hospital discharge record; considering MS-DRG, the principal ICD-9-CM diagnosis code, and secondary diagnoses to create a binary indicator of the presence of each of 29 different comorbid conditions. The conditions are generally unrelated to the patient's orthopedic condition. https://www.hcup-us.ahrq.gov/toolssoftware/comorbidity/comorbidity.jsp accessed 14 January 2016.

reported which was then regressed on the previously listed hospital and patient variables. Clustering of patients within hospitals was accounted for in each regression analysis to avoid underestimating standard errors. Absence of significant association between reporting status and hospital, physician, and patient variables would suggest that non-reporting of bearing surface is random.

Results

Descriptive statistics were tabulated to describe bearing use across the study time period and across patient and hospital descriptors. Tabulations across contextual descriptors also show the sample prevalence for each descriptor variable value in the leftmost column. The next 4 columns show the relative prevalence of each bearing surface for each descriptor variable strata. In all cases, PRD-provided weighting values were used to create estimates of US prevalence.

Bearing Surfaces across Time

Table 4.1 indicates an estimated 1,851,893 primary total hip arthroplasty procedures were conducted in the US during 2007-2012, as determined from the PRD. This total includes both those procedures with the bearing surface noted (852,610; 46.0%) and those with no bearing surface reported (999,824; 54.0%). Total procedure volume grew 19.3% across the study period. MoP bearing growth outpaced total growth with an increase of 28.3%, however an almost 4-fold increase in CoP bearing surface use was observed from a small base at the beginning to approaching parity with MoP. From a position approaching parity with MoP at the outset, MoM use fell dramatically (greater than 4-fold decrease in volume) across the study period. From a small base at the beginning of the study period, CoC's use fell to well under half of its starting value.

Henceforth, descriptive results will be reported based only on the THA procedures with reported bearing surface values, weighted to provide national US estimates of THA procedures with reported bearing surface. Table 4.2 shows the trends in prevalence of the 4 different bearing surfaces annually through the study period. Table 4.2 reveals that, although MoP gained some position lost by MoM and CoC, CoP picked up the lion's share of general procedure growth and the share lost by the bearing surfaces that declined in use. The total line of Table 4.2 also indicates the overall proportion values of each bearing surface across the entire study period; thus, creating an "average" study period value against which subsequent stratified segments can be compared (MoP: 47.7%, MoM: 25.7%, CoC: 4.3%, CoP: 22.3%). Figure 4.2 graphically shows the large swings in bearing surface prevalence across the study period.

Patient Demographics

Table 4.3 examines bearing surface and patient demographic characteristics. A steady trend upward in MoP prevalence (from 28.8% to 64.9%) and downward in CoP prevalence (from 28.9% to 10.2%), respectively, was seen for older age groups relative to younger. CoC's use plummeted going from the youngest to the oldest patients (almost 7-fold drop), while MoM prevalence decreased only slighter in older strata.

Table 4.3 also indicates that 55.7% of study THA patients were female. MoP bearings were significantly more prevalent in females than males while MoM was significantly less prevalent in females. No significant differences were seen between the sexes in the use of CoC and CoP bearings. Medicare was the dominant primary payer, accounting for 50.1% of all THAs during the study with MoP being disproportionately high in prevalence (57.7%) for these patients. MoP, CoC, and CoP bearings were used in greater prevalence when private insurance was the payer as compared to Medicare. Medicaid was the payer for a small number of primary THAs and its bearing surface prevalence profile was virtually identical to that of private insurance.

Hospital and Physician Descriptors

Table 4.4 describes the relationship between bearing surface choice and descriptors of the hospitals in which procedures were done. A preponderance of THAs were conducted at urban nonteaching hospitals (608,027; 71.3%) with little difference across the hospital types in MoP prevalence, hovering around the “average” MoP prevalence at 46.9%-49.6% use. Rural hospitals used the highest prevalence of MoM bearings (35.0%) and the lowest of CoC (2.6%) and CoP (13.5%). Urban teaching hospitals used the greatest prevalence of CoC bearings (5.0%), with prevalence almost as high in urban nonteaching hospitals (4.1%).

The vast majority of THA procedures were performed at large hospitals (698,008; 80.8%). Large hospitals used the lowest prevalence of MoP (39.8%) and CoC bearings (3.9%) and the greatest prevalence of MoM (28.0%) and CoP bearings (28.3%). Medium size hospitals used the greatest prevalence of MoP bearings (49.1%). Small hospitals used the lowest prevalence of MoM (23.9%), while using the highest prevalence of CoC (10.5%), more than twice as often as medium and large hospitals.

Although total procedure volumes varied relatively little across the US census regions, surface prevalence did vary across regions with the Midwest highest (56.9%) and West lowest (43.1%) in MoP prevalence. The South region (31.8%) was highest in MoM prevalence while the Northeast region (18.3%) was lowest. Cross-region prevalence variance was not substantial for either the CoC or CoP surfaces.

Patient Clinical Indicators

Bearing surface and patient clinical characteristics are shown in Table 4.5. Osteoarthritis of the pelvic region[p] (739,533; 86.7%) dominated the ICD-9-CM primary diagnosis code for the population with avascular necrosis of the femoral head (AVN)[q] (46,432; 5.4%), and fracture of the pelvis/upper femur[r] (26,491; 3.1%) completing the THA-related diagnosis codes. The bearing surface prevalence profile of the dominant osteoarthritis primary diagnosis generally mirrored the overall study population prevalence profile (see Table 4.2). MoP is of considerably lower prevalence when AFN has been diagnosed (39.3%) and considerably greater prevalence where fracture has been diagnosed (55.6%).

DRG 470 (absence of major comorbidities or complications) dominated (820,630; 96.2%) and, by virtue of such dominance, had a bearing prevalence profile virtually identical to the total population profile (see Table 4.2). MoP was used in greater prevalence (52.7%) in those patients for whom DRG 469 was in place with CoP used in lower prevalence (17.0%).

The count of AHRQ comorbidity presence indicators is an indicator of the general health of the patient. The groupings were selected based on low count being worthy of distinction and the desire to create similar sized strata. Patients with 1 or 2 diagnoses straddled the population average for MoP prevalence at 46.2% and 49.6, respectively, while patients with 0 and 3 or more comorbidities were significantly lower (39.2%) and

---

[p] Includes one of the following ICD-9-CM diagnosis codes as primary: 715.25, 715.25, 715.35, or 715.95 from CMS listing at https://www.cms.gov/medicare/coding/ICD9providerdiagnosticcodes/codes.html, accessed 14 January 2016.

[q] Denoted by ICD-CM diagnosis code 733.42 as primary.

[r] Includes one of the following ICD-9-CM diagnosis codes as primary: 73.21, 82.08, 82.09, 715.96, 733.14, 733.82, 82.00, 820.01, 820.02, 820.03, 820.09, 820.19, 820.20, 820.21, 820.22, 820.31, 820.32, 821.00, 821.01, 821.20.

higher (53.1%), respectively, than the population average. Very little difference in MoM prevalence was noted across comorbidity counts. Both CoC and CoP prevalence trended downward as comorbidity counts increased.

The discharge status describes the patient's destination post-surgery. The noted discharge status values generally indicate poorer health and/or less patient independence going down the table. MoP prevalence increased sharply from discharge to home (41.9%) down to discharge to skilled nursing/critical care (SNF/CCF; 59.3%). CoC and CoP trend in the opposite direction, where patients going home have the highest prevalence of these surfaces at 5.5% and 27.9%, respectively. MoM prevalence does not show a pattern with discharge status.

Regression Analysis of Bearing Choice

Table 4.6 presents the results of multinomial regression using SAS Proc Surveylogistic (SAS Institute; Cary, NC). This analysis allowed one to assess impact of each of the predictor variables on the choice of MoM, CoC, and CoP compared to the referent MoP while accounting for clustering of surgeries within hospitals. The burden for statistical significance was established as $p<0.01$.

CoC surface preference over MoP steadily moved downward (odds ratio (OR) = 0.553 to 0.327) across the study period while MoM's steep downward trend became statistically significant in 2010, dropping all the way to OR = 0.184 in 2012. COP's increasing preference over MoP was steady and became statistically significant with big moves upward in 2011 (OR=1.990) and 2012 (OR=2.365).

Steep, significant trends downward in use were seen with increasing age for CoC (OR=0.544 to 0.149) and CoP (OR=0.776 to 0.195) with a more modest trend in MoM

(OR=0.743 to 0.441). Females were significantly less likely to have used MoM over MoP, but registered no significant difference from males in regard to the other two surfaces vs. MoP. The only significant result observed in the payer descriptor was greater use of CoC over MoP in those patients with private insurance.

No significant surface preferences were seen in hospital type (rurality-teaching status). Small hospitals showed strong and significant preferences for MoM (OR=2.097) and CoC (4.275) over MoP. This could be related to orthopedic surgeons have disproportionate input into the procurement decisions at small hospitals. Midwest (OR=0.427) and northeast (OR=0.274) census regions showed strongly reduced MoM preference compared to MoP. The Midwest region (0.451) also showed significantly reduced preference for CoP over MoP.

A primary diagnosis of hip/femur fracture was associated with reduced use of both CoC (OR=0.625) and CoP (OR=0.610) compared to MoP. No significant associations were seen between MS-DRG and any of the bearing surfaces. Increased AHRQ comorbidity count was associated with decreasing prevalence of CoC bearings (OR=0.786 down to 0.579) compared to MoP. One comorbidity was associated with a reduced preference for CoP (OR=0.902) vs. MoP. In discharge status, general trends down were seen in CoC and CoP use as discharge status indicated sicker patients. The observations were only significant in CoC for other (OR=0.387) and in CoP for SNF/CCF (OR=0.663) and other (OR=0.565).

Assessing Randomness of Surface Reporting

A surface reporting variable was created that captured whether the bearing surface was reported in each record. This was then used as the dependent variable in a regression using all hospital and patient descriptors as independent variables, accounting for clustering of observations within hospitals. Table 4.7 reveals that 3 of 11 independent variables, age, payer, and US census region, showed significant associations with whether or not bearings were reported. Patients 75-84 and 85 and older had 13.1% and 16.7% lower odds, respectively, of having bearing surface reported. Bearings were more likely to be reported for privately insured patients (16.7% higher odds) and less likely for those with Medicaid (22.3% lower odds). Although all 3 census regions were much more likely to have bearings reported than the referent South region, only the Northeast was statistically significant (242% higher odds).

Discussion

The major contribution of this research is to report on trends in bearing surface use across 2007-2012 applying rigorous quantitative methods to the large, nationally representative PRD database. Both bivariate tabulation and multivariate regression revealed pronounced trends in the different bearing surfaces across time, punctuated by the fall in use of MoM and CoC accompanied by large growth in use of CoP. The large growth in annual volume of primary THA procedures seen across 2007-2012 (19.3% growth) was predominantly taken by CoP surfaces (290% growth), with MoP growing at a much slower pace (28.5% growth).

The large decreases in prevalence of MoM (37.4% in 2007 to 8.1% in 2012) and CoC (7.0% in 2007 to 2.8% in 2012) are consistent with significant recent observations of clinical issues with these surfaces. Excellent in vivo wear rates led physicians and scientists to believe these hard-on-hard surfaces were a path to longer lasting implants.[13, 14] Clinical issues arose in both surfaces and ultimately led to the results observed in this study. Infrequent but catastrophic CoC bearing failures related to ceramic liner fracture[15] combined with frequent audible squeaking and clicking (4-20%)[16, 17] have raised significant concerns. Metal-on-metal bearing surfaces were plagued by permanently elevated ion levels in the bloodstream caused by corrosion[14] and concerns about local tissue reactions.[18] Ultimately, the US Food and Drug Administration concluded that eroding metal can cause localized soft tissue damage,[19] leading to recalls of MoM components and systems in 2008-2012.[20] The timing of these MoM recalls and equally importantly, the inevitable anecdotal knowledge that would have spread in the clinical world, almost certainly drove the sharp fall in MoM bearing use. With MoM and CoC out of favor, CoP, whose rapid growth this research has confirmed, was left as the leading option to extend implant life.[26]

Sharp trends downward in use of CoC (OR=0.544 to 0.149) and CoP (OR=0.776 to 0.195) as age increased are consistent with more expensive, presumed longer lasting, implants being a better fit for patients with more years of life remaining and confirm results seen in bivariate analysis. The use of MoM also trended downward with increasing patient age, but not nearly as sharply as did CoC and CoP. Of the many significant relationships between payer class and surface seen in bivariate analysis, only CoC's association with private insurance (46.2% greater odds than Medicare) remained significant in the multivariate regression (possibly related to privately insured population being younger).

The absence of significant multivariate association between hospital teaching status and bearing choice refutes the hypothesis that teaching hospitals would adopt more advanced surfaces at greater rate than non-teaching. Small hospitals adopted newer bearings at much greater rates than MoP (OR=2.097 for MoM, OR=4.275 for CoC) and thus, represent the hospitals that statistically led use of new technology in hip arthroplasty. Significantly reduced odds ratios seen for MoM and CoC bearings across two US census regions (Midwest and Northeast) confirm bivariate relationships.

Clinical variables can also give an indirect indication of potential years of life remaining. Bearing surface use varied with AHRQ comorbidity count very similarly to the observed relationship between surface and age. In bivariate analysis, use of MoP increased significantly (39.2% to 53.1%) as comorbidity count increased from 0 to 3 whereas CoC and CoP use dropped significantly. MoM use dropped very little. Multivariate analysis showed similar trends although the results were not as consistently significant. The discharge status and MS-DRG show some of the same tendencies but the results were of mixed significance. Since the MS-DRG is assigned after the hospitalization is complete, it is difficult to know if a less "healthy" discharge arose as a result of the surgery or from a pre-existing condition.

Both bivariate and multivariate analyses show trends of greater MoP bearing use and lesser use of CoC and CoP bearings on patients in older and sicker patients; thus, focusing higher priced implants on patients with greater potential life years remaining. MoM bearings (also generally more expensive than MoP) had much flatter profiles across the same indicators. There is no obvious explanation for the MoM phenomenon.

Another significant contribution is the multivariate regression on surface reporting conducted to test whether bearing surface reporting was random; in light of the fact that only 46% of surgeries in the sample have a bearing surface reported. Absence of widespread variation in reporting status is essential to generalizing beyond the sample data available. The initial look into the epidemiology of bearing surface use by Bozic et al[27] indicated that they saw no bias but did not detail how they came to this observation. In the study presented here reporting was higher for private insurance (OR=1.126) and lower for Medicaid (OR=0.787) than the referent Medicare payer, although Medicaid accounted for only 2.5% of THAs. Mathematically large impacts of US census region on surface reporting were also seen, but only the Northeast region's odds ratio (OR=3.424) was significantly different from the referent South region. Age also showed statistically significant associations with reporting. The 75 and older population segments were less likely to have surface reported (13.1% and 15.7% reduced odds for 75-84 and 85 and up, respectively) but accounted for only 22.9% of all surgeries and the observed impact was not very large; thus the practical impact on the results is limited. On balance, the relationships observed here between the indicators and surface reporting should not unduly limit generalization of the results.

Limitations

Administrative records like those used in this study are generated primarily for remuneration purposes; thus, using them for other purposes always offers the potential that data not critical for reimbursement is not reported accurately or at all. The voluntary reporting of bearing surface would be of particular concern, given its importance to this study and particularly with surface reported in only 46% of cases. Fortunately, relatively little nonrandom reporting behavior was seen.

Conclusions

The trends in reported bearing surface use across time are consistent with clinical information and market recall activities related to MoM and CoC. The trend to use fewer of the more expensive devices as patient age increased indicates that the patient-physician decision process could be informally incorporating cost-effectiveness considerations of more expensive devices. The general trend of poorer clinical status (AHRQ comorbidity count, and discharge status) being associated with use of the cheaper MoP bearing surface is another potential indication that cost-effectiveness is being considered. Finally, the multivariate regression attenuated the relationship between bearing surface choice and the descriptor variables. The bivariate analysis produced many significant associations with bearing surface choice, the most important of which (bearing surface with year of procedure and age) were strongly confirmed in the regression while others were not (sex, hospital type, and hospital size only partially confirmed). Producing only bivariate results, as was the case in 3 previous studies cited,[27-29] would have overstated the association between descriptors and bearing surface.

MoM bearings consistently showed use patterns different from those of the other two newer surfaces, CoC and CoP, particularly as related to age and patient clinical indicators. Specifically, use of CoC and CoP tended to decrease in older and sicker patients, consistent with aligning more expensive devices with patients more likely to benefit. MoM's profile across these 2 variables was much flatter. Increase in MoP's use was consistently associated with sicker and older patients. Were MoM bearings to regain a significant share of the total hip arthroplasty market, its unique associations related to age and clinical indicators could lead to interesting future research.

The absence of systematic nonrandom behavior in bearing surface reporting is an important observation as it indicates that, even though only 46% of the surgeries included a bearing surface, results can be generalized from this data set.

Table 4.1: All THA procedures

Year	Surface not reported	MoP	MoM	CoC	CoP	Total
2007	146,203	57,413	48,423	9,097	14,461	275,597
2008	156,691	60,014	51,649	5,382	17,509	291,246
2009	170,773	65,954	49,712	6,158	24,582	317,178
2010	172,563	75,796	36,625	6,777	32,524	324,284
2011	170,242	73,432	21,001	5,320	44,731	314,726
2012	182,812	73,678	11,830	4,032	56,511	328,863
Total	999,284	406,287	219,240	36,765	190,318	1,851,893
	54.0%					

Table 4.2: THAs with reported bearing surfaces - annual procedures
(% prevalence within the year)

Year	MoP	MoM	CoC	CoP	Total
2007	57,413 (44.4%)*	48,423 (37.4%)*	9,096 (7.0%)*	14,461 (11.2%)*	129,393
2008	60,014 (44.6%)*	51,649 (38.4%)*	5,382 (4.0%)*	17,509 (13.0%)*	134,555
2009	65,954 (45.0%)*	49,712 (34.0%)*	6,158 (4.2%)	24,582 (16.8%)*	146,406
2010	75,796 (50.0%)	36,625 (24.1%)	6,776 (4.5%)	32,524 (21.4%)*	151,722
2011	73,432 (50.8%)*	21,001 (14.5%)*	5,320 (3.7%)	44,731 (31.0%)*	144,484
2012	73,678 (50.4%)*	11,830 (8.1%)*	4,031 (2.8%)*	56,511 (38.7%)*	146,051
Total	406,287 (47.7%)	219,239 (25.7%)	36,765 (4.3%)	190,318 (22.3%)	852,610

Table $p<0.001$; Cell values with * have $p<0.001$.

Table 4.3: Patient demographics and bearing surfaces

Age group		MoP	MoM	CoC	CoP
18-45	(43,522; 5.1%)	28.8%*	31.3%*	11.0%*	28.9%*
45-54	(133,295; 15.6%)	33.0%*	30.5%*	8.1%*	28.5%*
55-64	(241,891; 28.4%)	39.4%*	26.4%	5.5%*	28.7%*
65-74	(238,381; 28.0%)	54.1%*	23.8%*	2.2%*	19.9%*
75-84	(159,745; 18.7%)	64.1%*	22.5%*	1.3%*	12.2%*
≥ 85	(35,776; 4.2%)	64.9%*	23.4%	1.6%*	10.2%*
Total	(852,610)				

Table $p<0.001$; Cell values with * have $p<0.001$.

Sex		MoP	MoM	CoC	CoP
Male	(377,928; 44.3%)	44.4%*	28.6%*	4.7%*	22.3%
Female	(474,591; 55.7%)	50.2%*	23.4%*	4.0%	22.4%
Total	(852,519)				

Table $p<0.001$; Cell values with * have $p<0.001$.

Primary payer		MoP	MoM	CoC	CoP
Medicare	(427,358; 50.1%)	57.7%*	23.4%*	1.9%*	16.9%*
Private	(368,711; 43.2%)	38.0%*	27.5%*	6.8%*	27.7%*
Medicaid	(20,990; 2.5%)	37.5%*	27.7%	6.3%*	28.5%*
Other	(35,549; 4.2%)	33.3%*	32.9%*	6.3%*	27.6%*
Total	(852,610)				

Table $p<0.001$; Cell values with * have $p<0.001$

Table 4.4: Hospital and physician descriptors and bearing surface

Hospital type		MoP	MoM	CoC	CoP
Urban nonteaching	(608,027; 71.3%)	46.9%*	26.5%*	4.1%*	22.4%
Urban teaching	(220,526; 25.9%)	49.6%*	22.4%*	5.0%*	23.0%*
Rural	(24,057; 2.8%)	48.9%	35.0%*	2.6%*	13.5%*
Total	(852,610)				

Table p<0.001; Cell values with * have p<0.001

Hospital size		MoP	MoM	CoC	CoP
Large	(689,008; 80.8%)	39.8%*	28.0%*	3.9%*	28.3%*
Medium	(116,431; 13.7%)	49.1%*	25.4%	4.0%	21.5%*
Small	(47,171; 5.5%)	46.3%*	23.9%*	10.5%*	19.3%*
Total	(852,610)				

Table p<0.001; Cell values with * have p<0.001

Hospital US census region		MoP	MoM	CoC	CoP
South	(230,529; 27.0%)	40.3%*	31.8%*	4.0%*	24.0%*
Midwest	(175,511; 20.6%)	56.9%*	21.8%*	3.9%*	17.5%*
Northeast	(206,720; 24.2%)	53.3%*	18.3%*	5.5%*	22.8%*
West	(239,840; 28.1%)	43.1%*	29.2%*	3.9%*	23.8%*
Total	(852,610)				

Table p<0.001; Cell values with * have p<0.001

Table 4.5: Patient clinical variables and bearing surfaces

Primary diagnosis		MoP	MoM	CoC	CoP
Osteoarthritis-pelvis	(739,533; 86.7%)	47.9%	25.3%*	4.2%	22.6%
AVN	(46,432; 5.4%)	39.3%*	30.7%*	6.1%*	23.9%
Fracture-pelvis/upper femur	(26,491; 3.1%)	55.6%*	26.9%*	2.9%*	14.6%*
Other	(39,684; 4.7%)	48.0%	27.2%	4.7%	20.2%
Total	(852,610)				

Table p<0.001; Cell values with * have p<0.001

MS-DRG code		MoP	MoM	CoC	CoP
470	(820,630; 96.2%)	47.5%	25.7%	4.3%	22.5%
469	(29,011; 3.4%)	52.7%*	27.0%	3.3%*	17.0%*
Other	(2,969; 0.3%)	49.6%	33.6%*	4.6%	12.2%*
Total	(852,610)				

Table p<0.001; Cell values with * have p<0.001

AHRQ comorbidity count		MoP	MoM	CoC	CoP
0	(156,132; 18.3%)	39.2%*	27.9%*	6.4%*	26.4%*
1	(249,033; 29.2%)	46.2%*	26.0%	4.8%*	23.1%
2	(211,666; 24.8%)	49.6%*	25.7%	3.7%*	21.1%*
3 or more	(235,799; 27.7%)	53.1%*	24.0%*	3.0%*	19.9%*
Total	(852,610)				

Table p<0.001; Cell values with * have p<0.001.

Discharge status		MoP	MoM	CoC	CoP
Home	(193,195; 22.7%)	41.9%*	24.8%*	5.5%*	27.9%*
Home w/formal care	(369,562; 43.3%)	44.4%*	27.1%*	4.9%*	23.7%*
Rehab facility	(86,107; 10.1%)	47.0%*	29.9%*	3.5%*	19.7%*
SNF/CCF	(197,108; 23.1%)	59.3%*	22.2%*	2.6%*	15.9%*
Other	(6,638; 0.8%)	59.7%*	27.0%	1.3%*	12.0%*
Total	(852,610)				

Table p<0.001; Cell values with * have p<0.001.

Table 4.6: Multinomial regression of bearing surface

Effect	Surface-Odds Ratio-vs. MoP		
Procedure year-ref: 2007	MoM	CoC	CoP
2008	1.019	0.553**	1.133
2009	0.864	0.532*	1.306
2010	0.580**	0.480*	1.453
2011	0.335**	0.409**	1.990*
2012	0.184**	0.327**	2.365**
Age cohort – ref: 18-54			
55-64	0.743**	0.544**	0.776**
65-74	0.542**	0.242**	0.434**
75-84	0.431**	0.124**	0.235**
85 and older	0.441**	0.149**	0.195**
Sex-ref: male			
Female	0.778**	1.083	1.064
Payer-ref: Medicare			
Private Insurance	1.179	1.462**	1.154
Medicaid	1.176	1.391	1.086
Other	1.368*	1.437	1.194
Hospital type-ref: urban nonteaching			
Rural	1.451	0.679	0.644
Urban teaching	1.213	1.020	1.127
Hospital size-ref: large			
Medium	1.573	1.738	1.456
Small	2.097*	4.275**	0.902
US census region-ref: south			
Midwest	0.427*	0.509	0.451*
Northeast	0.274**	0.720	0.631
West	1.040	1.122	1.012
Primary diagnosis-ref: osteoarthritis			
AVN	0.952	1.011	0.892
Fracture pelvis/upper femur	0.760	0.625*	0.610**
Other	0.850	0.878	0.814
MS-DRG-ref: 470			
469	0.912	1.030	1.019
Other	1.297	1.070	0.638
AHRQ comorbidity count-ref: 0			
1	0.941	0.786*	0.902*
2	0.959	0.678**	0.860
3 or more	0.900	0.579**	0.819
Discharge status-ref: home			
Home w/formal care	1.033	0.928	0.913
Rehab facility	1.401	0.855	1.024
SNF/CCF	0.954	0.741	0.663*
Other	1.080	0.387*	0.565*

*p<0.01; **p<0.0001

Table 4.7: Bivariate logistic regression on surface report status

Effect	Surface report v. not report	
Procedure year-ref: 2007	Odds ratio	Table p-value
2008	0.965	0.7770
2009	0.943	
2010	0.945	
2011	0.928	
2012	0.880	
Age cohort – ref:18-54		
55-64	0.953	0.0071
65-74	0.927	
75-84	0.869*	
85 and older	0.843*	
Sex-ref: male		
Female	1.004	0.8000
Payer-ref: Medicare		
Private insurance	1.167*	<0.0001
Medicaid	0.787*	
Other	0.926	
Hospital type-ref: urban nonteaching		
Rural	0.666	0.3551
Urban teaching	0.969	
Hospital size-ref: large		
medium	1.177	0.5918
Small	0.891	
US census region-ref: south		
Midwest	1.392	<0.0001
Northeast	3.424**	
West	2.166	
Primary diagnosis-ref: osteoarthritis		
AVN	0.948	0.2894
Fracture pelvis/upper femur	0.897	
Other	0.966	
MS-DRG-ref: 470		
469	0.933	0.3124
Other	0.972	
AHRQ comorbidity count-ref: 0		
1	0.970	0.4614
2	0.966	
3 or more	0.975	
Discharge status-ref: home		
Home w/formal care	1.065	0.0909
Rehab facility	0.797	
SNF/CCF	0.996	
Other	0.876	

**p<0.01, ***p<0.0001

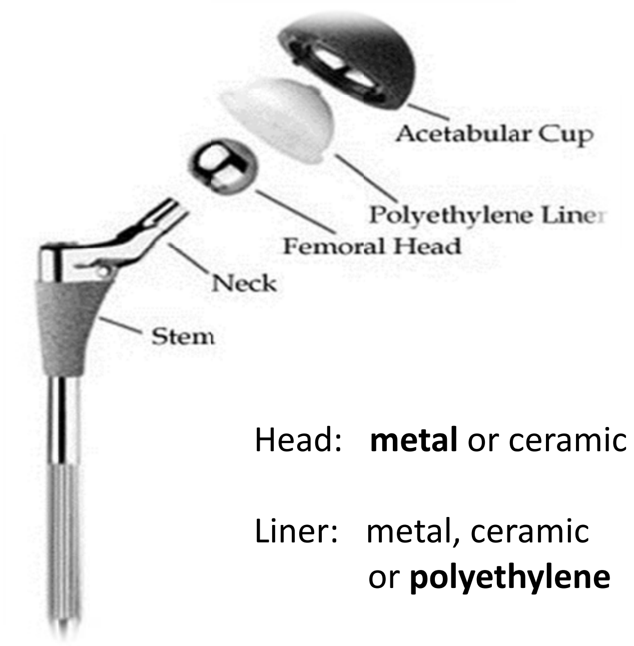

Figure 4.1: Hip arthroplasty hardware and bearing surface possibilities

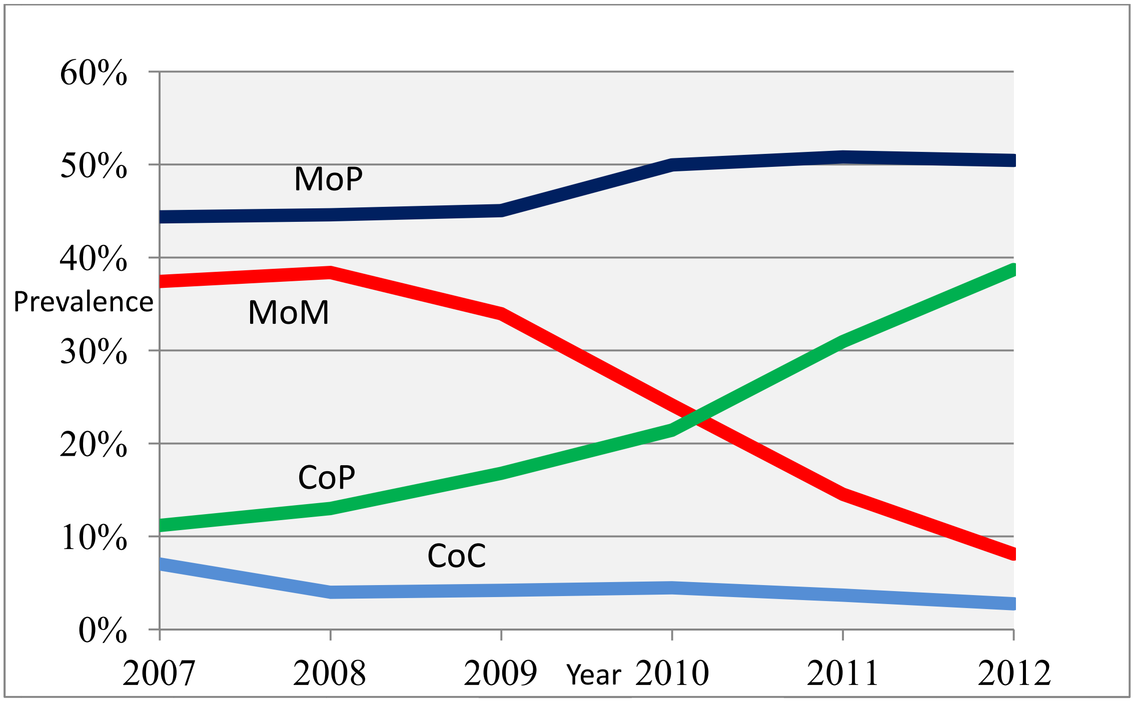

Figure 4.2: THA bearing surface prevalence across time

CHAPTER 5: COSTS OF BEARING SURFACE AND SURGERY IN TOTAL HIP ARTHROPLASTY

Primary total hip arthroplasty (THA) is a common surgical procedure that most often is employed to overcome steadily worsening functionality and quality of life due to osteoarthritis of the hip joint. In 2010 THA comprised 1.3% of all non-diagnostic, non-pregnancy-related hospital procedures in the US, increasing to 2.6% of procedures in the ≥65 year-old population.[s][3] Despite its relatively high cost, the cost-effectiveness of this medical procedure has been demonstrated vs. available alternatives.[4-7] Nonetheless, efforts to improve on the original design continue in a number of areas with bearing surface materials at the forefront of these efforts.

Improvements to the bearing surface have historically come at higher prices[21] which are almost always incurred by hospitals. Explorations into the long-term cost-effectiveness of materials with different cost/performance expectations are beginning to be conducted but have not yet provided clear direction.[48-50] Until such time as clear guidance is given on which materials are most cost-effective for which types of patient, primary insurance and public policy focus are likely to remain on upfront costs, in large part because 10 or more years can be required before appreciable outcome differences are observed between surfaces. This paper addresses issues that are of significant concern at the time of the procedure: the cost of implant materials and the procedure's total hospitalization costs. Previous efforts to describe implant costs[28, 29, 44, 45, 47] (detailed in

---

[s] Author's analysis of 2010 National Hospital Discharge Survey Results, conducted by the Centers for Disease Control. http://www.cdc.gov/nchs/fastats/inpatient-surgery.htm accessed 9 January 2016.

Chapter 2) were informative, but suffered from one or more of the following limitations: narrow geographic coverage, small number of hospitals contributing cases, aggregate implant costs were reported, reporting of charge data (rather than cost), or short length of time studied. The study reported here describes trends across time and demographic descriptors of costs of the four dominant bearing surfaces in use today: metal-on-polyethylene (MoP), metal-on-metal (MoM), ceramic-on-ceramic (CoC), and ceramic-on-polyethylene (CoP) as well as for total hospitalization costs. The following research question will be addressed:

> What were the drivers of cost of the 4 main bearing surface combinations and surgical costs in primary total hip arthroplasty in the US across time, hospital, and patient variables from 2007-2012?

It is hypothesized that the cost incurred for bearing surface implants (as a proxy for price paid) will vary inversely with hospital size and/or purchased volume; i.e., larger volume purchasers will pay lower prices.

Methods

This study used the Premier Research Database (PRD; Premier, Inc., Charlotte, NC) of hospital discharge records for 2007-2012 from general, short-term, nonfederal, acute care US hospitals that chose to provide patient level demographic, cost and clinical data, as well as hospital demographic information. In 2005 the Centers for Medicare and Medicaid Services (CMS) introduced International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) surgical procedure codes making it possible to denote each of the bearing surface choices investigated here.[9] Availability of these voluntarily reported codes made the surface choice a potential topic for nationally representative research using large-scale administrative databases.

Premier is a large group purchasing organization (GPO), with approximately 3,400 participating US hospitals.[t] The PRD includes a smaller subset of hospitals, some of which are GPO members and others not that participate in a range of quality improvement activities with Premier. PRD hospitals provide clinical, demographic, and item-level cost data for a 100% census of all discharges. Across the study period of 2007-2012, 523 hospitals spanning 44 states reported hip arthroplasty procedures. These procedures were the data source for this study. This THA study period began in 2007 because this was the first full year that ICD-9-CM procedure codes were available for reporting of all four bearing surfaces of interest.

The PRD provides many patient- and hospital-level indicators. Patient demographic variables used were year of the procedure, age, sex, and primary payer. Staffed bed count, teaching status, rural/urban status, and US census region were used to describe conditions of hospitalization. Medicare Severity Diagnosis Related Group[u] (MS-DRG), discharge status, length of hospital stay, and ICD-9-CM primary diagnosis codes (osteoarthritis,[v] avascular necrosis of the femoral head (AVN),[w] pelvis/upper femur fracture[x]) were used as proxies for the patient's health status.

---

[t] As reported by the Health Group Purchasing Industry Initiative, a trade association of eleven of the leading GPOs, http://www.healthcaregpoii.com/signatorycompanies/premier.html accessed 7 January 2016.

[u] The MS-DRG is a value assigned to each inpatient stay that incorporates the patient's primary diagnosis, secondary diagnoses, the principal procedure and any additional procedures, sex, and discharge status. MS-DRG values related to THA are 469 and 470, which stand for major joint replacement associated with lower extremity with and without, respectively, major complications and comorbidities.

[v] Includes one of the following ICD-9-CM diagnosis codes as primary: 715.25, 715.25, 715.35, or 715.95 from CMS listing at https://www.cms.gov/medicare/coding/ICD9providerdiagnosticcodes/codes.html, accessed 14 January 2016.

[w] Denoted by ICD-CM diagnosis code 733.42 as primary.

[x] Includes one of the following ICD-9-CM diagnosis codes as primary: 73.21, 82.08, 82.09, 715.96, 733.14, 733.82, 82.00, 820.01, 820.02, 820.03, 820.09, 820.19, 820.20, 820.21, 820.22, 820.31, 820.32, 821.00, 821.01, 821.20.

Cost Determination

The PRD contains descriptive detail for each item and service incurred during each hospitalization. Appendix A details the procedure by which implant and total hospitalization costs were determined for each discharge record included in this study. All costs were indexed to 2015 dollars using the Bureau of Labor Statistics consumer price index inflation calculator.[y] Charge and cost information are included for each item although this study used only cost data. The cost for each item is described as the cost to deliver the item to the point of use and thus can be expected to include the purchase price of the item and overhead or other costs a given hospital chooses to apply to that item.

Study Sample Characteristics

Primary THA case records for patients $\geq$18 years of age were selected from the PRD based on one incidence of the 81.51 ICD-9-CM procedure code and no other hip arthroplasty codes present except for ones denoting bearing surface. Specifically, revisions, bilateral THAs, and resurfacing procedures were excluded from this analysis. The PRD included 278,179 THA procedures across the study period, of which 128,526 (46.2%) included an indicator of bearing surface. The bearing surface used in a THA was determined by presence of one of the following ICD-9-CM procedure codes: 00.74 (MoP), 00.75 (MoM), 00.76 (CoC), or 00.77 (CoP). To minimize the impact of data entry errors, $0 cost values for implants and total hospitalization costs were excluded as well as those in the top and bottom 1% of the total cost distribution, leaving 121,128 cases in the study from 485 hospitals.

---

[y] The US Department of Labor, Bureau of Labor Statistics charts movements in consumer prices and provides a CPI Inflation Calculator that can be used to equate buying power across different time periods. Accessed on 1 February 2016 at http://www.bls.gov/data/inflation_calculator.htm.

Analytic Methods

The study began by describing the use of different bearing surfaces across time in tabulated and graphical form. Next was graphical characterization of total hospitalization cost and implant cost across time and bearing surface. This was followed by graphical presentation of implant costs by bearing surface and hospital descriptors. To support these efforts, Agency for Healthcare Research and Quality (AHRQ)-suggested logical groupings of hospitals by bed count, rurality, teaching status, and US census region were used to create variables for hospital type (urban nonteaching, urban teaching, and rural) and size (large, medium, and small).

Regression Analysis

Regression of implant cost was conducted on bearing surface, procedure year, hospital descriptors (type, size, and US census region, and staffed bed count), surgical volume variables (annual volume and annual volume/staffed bed), primary payer, primary diagnosis, MS-DRG, discharge status, patient age, AHRQ comorbidity count,[z] length of hospital stay, and sex. This analysis was to determine if significant relationships might indicate some logical advantages in price negotiation, particularly if volume purchased or hospital size might play a role. To supplement the HCUP size variable and staffed beds as volume indicators, the annual volume of THA procedures for each hospital was calculated and used in the model. Further, the annual volume was divided by the number of staffed beds to obtain a possible indicator of the level of "focus" a given facility might have on THAs.

---

[z] AHRQ software was used to review the hospital discharge record, considering MS-DRG, the principal ICD-9-CM diagnosis code, and secondary diagnoses to create a binary indicator of the presence of each of 29 different comorbid conditions. https://www.hcup-us.ahrq.gov/toolssoftware/comorbidity/comorbidity.jsp accessed 14 January 2016.

In order to characterize the drivers of total hospitalization cost, regression was conducted of total hospitalization cost on bearing surface, procedure year, hospital descriptors (type, size, US census region, and staffed bed count), surgical volume variables (cumulative volume and cumulative volume/staffed bed), primary payer, primary diagnosis, MS-DRG, discharge status, patient age, AHRQ comorbidity count, length of hospital stay, and sex.

General estimating equation (GEE) methods were used for both regression analyses, accounting for clustering of observations within hospitals to generate robust standard errors. Consistent with many sources of hospital discharge cost data, the visual appearance of both implant and total costs were non-normal, with many low values and a smaller but numerically influential concentration of high values as well, even after the data cleaning described in “Study Sample Characteristics.” Per the method suggested by Manning and Mullahy,[59] log-scale residuals of both regression models were evaluated for kurtosis. Given a statistically significant observation of kurtosis >3, both implant and total hospitalization costs were log-transformed and generalized linear models were constructed.

Results

Bivariate tabular and graphical results describing bearing use across time are reported followed by implant cost and total surgery cost across time and bearing surface. Next, implant costs across bearing surface and hospital descriptors will be presented using graphical depictions. Finally, multivariate regression analyses on implant and total costs will be reported. All costs are reported in 2015$.

Bearing Use Across Time

Table 5.1 shows the evolution in use of bearing surfaces in THA in the study population across the study period. It reveals a greater than 3-fold decrease in MoM use and a 28.4% decrease in CoC use across the study period while use of MoP and CoP grew significantly. MoP use grew 61.2% while CoP use grew almost 4-fold, compared to a 38.3% increase in number of THA procedures across the study period. Figure 5.1 shows relative prevalence of each bearing surface in each year and even more clearly shows that CoP is rapidly becoming the choice of doctors and patients, whereas MoP showed moderate growth, and MoM has plummeted in use.

Hospitalization Costs Across Time

Figure 5.2 shows the trends in average hospitalization cost across bearing surface and time. Appendix B tabulates details of these costs, reporting mean, standard deviation, and median values of total hospitalization cost in each stratum of bearing surface and year. In each year, statistically significant ($p<0.01$) differences in hospitalization costs between procedures using each non-MoP surface and those using MoP are noted with an asterisk. THA hospitalizations were significantly more costly in each year for those using MoM and CoP bearings than for those using MoP. Procedures utilizing CoC bearings were significantly more expensive than those using MoP in each year except for 2008 and 2009.

Figure 5.2 shows that hospitalization costs across all bearings decreased in 2008, increased in 2009, but then the cost trends by the surfaces began to differ. Total hospitalization costs for procedures using CoC bearings had the steepest drop in 2008, increased for 2 years, then had a steady decline for the last 2 years, showing the largest overall study period change going from \$19,376 to \$17,425 (decrease of \$1,951). Hospitalization costs for procedures using MoM bearings had a much flatter profile dropping by \$583 across the study period, significantly growing the gap in total costs for MoM procedures vs. those using other surfaces. Hospitalizations that used MoP bearings showed a steady decline in cost after the brief rise in 2009, with an overall cost decrease of \$898 across the study period. Hospitalizations utilizing MoP bearings had fairly flat costs across 2008-2011 before a sharp decline in 2012; the cost of surgeries using MoP implants decreased \$1,068 across the study period.

Implant Costs Across Time

Figure 5.3 shows the implant costs for each bearing surface across the study period. Appendix C tabulates the mean, standard deviation, and median values for implant cost in each bearing-year stratum, with mean implant costs significantly different ($p<0.01$) from those of MoP in a given year noted by an asterisk. MoM implants had significantly higher cost than MoP across all study years. CoC implants had significantly higher costs than MoP in each year but 2009 and 2012 while CoP implants had higher cost than MoP in 2009-2012. In 2007, MoM and CoC implants had almost identical costs, whereas MoP and CoP's costs were close to each other. By 2012, CoP and MoM bearing costs were virtually identical as were those of CoC and MoP.

Figure 5.3 showed a steady decline in MoP implant cost with time, dropping \$1,201 across the study period. Similarly, MoM implant costs steadily decreased between 2007 and 2012, for a total decrease of \$1,531. CoC implant costs dropped sharply in 2008 and 2009, increased sharply in 2010 and then sharply dropped again 2011 and 2012, for a total downward movement of \$2,073. CoP implant cost dropped sharply in 2008, followed by a steady increase through 2011 with a sharp drop in 2012 for a total study period decrease of $758.

Two studies allow comparisons of overall average implant costs to those presented here. Robinson et al.[45] reported $6,072 median 2008 implant costs (updated to $6,679 in 2015$ here). This compares to 2008 overall median implant costs of $7,404 (2015$) found in this study. Lehil and Bozic[29] reported average implant selling prices of $6,800 ($7,773 in 2015$) and $5,842 ($6,031 in 2015$) in 2007 and 2012, respectively. This study found average implant costs of $8,492 and $7,131 in 2007 and 2012, respectively, in 2015$.

Implant Costs Across Hospital Descriptors

Figure 5.4 displays implant costs by surface and hospital type, collectively across the entire study period. The asterisks above the columns denote significant differences in average cost from urban nonteaching hospitals ($p<0.01$) within each bearing surface. Appendix D tabulates the same information, reporting mean, standard deviation and median cost values for each stratum. MoP, MoM, and CoC surfaces costs were highest in urban nonteaching hospitals (MoM highest at $9,238), with much smaller cost differences between urban teaching and rural hospitals. The situation changed in CoP where rural hospitals reported the highest implant costs with urban nonteaching being 2nd highest. Urban teaching hospitals consistently reported the lowest costs although not all differences were statistically significant, with the lowest reported value of $6,679 for MoP.

Figure 5.5 reports implant costs by surface and hospital size across the entire study period. Asterisks denote significant differences in cost from large hospitals at

$p<0.01$ within each surface type. Appendix E tabulates the results, reporting mean, standard deviation and median for each hospital size-surface stratum. Across all surfaces, large hospitals reported the highest implant costs. Small hospitals reported the lowest implant costs within each surface with the exception of MoP, and the cost difference between small and medium was not significant in CoC. The full range of reported implant cost was narrower across hospital types than hospital size. MoP implants in medium size hospitals were the lowest cost at $5,914 and MoM surfaces in large hospitals were the highest cost at $8,883.

Figure 5.6 describes implant cost vs. bearing surface and US census region. Asterisks denote significant differences ($p<0.01$) in reported implant costs from hospitals in the South region within each surface type. Appendix F tabulates these results in detail. The most noticeable observation is that for each of the surfaces, the Northeast region reported the lowest implant costs. The Northeast costs differed from the other 3 regions by large margins and also was the only region significantly different in costs from the referent South across all surfaces. The Northeast implant cost values ranged from $6,556 for MoP surfaces to $7,459 for MoM surfaces. Further, the highest value for the Northeast was still significantly lower than the lowest remaining bearing surface-region combination ($7,708 for MoP in the South).

General Population Descriptive Statistics

General population descriptions were created for each of the variables used in the regression analyses to be reported next. Table 5.2 describes the implant costs and total hospitalization costs per procedure. Important observations here are the mean and median values for implant costs of $7,858 and $7,084 respectively. The mean and median costs for the total hospitalization costs were $18,273 and $17,283, respectively. The substantial difference between mean and median values in both cost measures is consistent with the positively skewed distribution often seen in hospitalization cost data. Tables 5.3 and 5.4 provide the population descriptions for the continuous and discrete independent variables, respectively, used in the regression analyses.

Regression Analysis

Table 5.5 details the regression analyses conducted. Estimated dollar impacts for each model parameter were created by exponentiation of the original GEE estimates, converting to a percentage change, and then applying this change to the mean implant cost and total hospitalization cost, respectively, across all procedures described in Table 5.2. Estimates with one asterisk are significant at the $p<0.05$ level and those with two asterisks are significant at the $p<0.01$ level. Note the annual and cumulative volume variables report the impact of 100 surgery increments, to give the estimates more practical meaning. Likewise, the staffed bed variable reports the impact in 100 bed increments for the same reason. Finally, age reports the impact of 10 year increments.

Use of MoM or CoC bearing surfaces was significantly associated with increased implant costs of $717 and $685, respectively, compared to MoP use. Surgery in 2008 or 2012 was significantly associated with decreased implant cost of $371 and $907, respectively, compared to surgery in 2007. None of the hospital descriptors, size, or surgery volume indicators showed any significant association with implant cost. Clinical variables did produce significant results as a primary diagnosis of upper femur/pelvis fracture was significantly associated with a $660 greater implant cost, compared to an osteoarthritis diagnosis. Further, a discharge status of “other” was associated with $1,041 increased implant cost,[aa] compared to being discharged to home without formal assistance. Patient demographics also played a role with 10 years in increased age associated with $95 decrease in implant cost whereas being female associated with an $87 decrease in implant costs.

In total hospitalization cost regression, bearing surface was associated with increased total cost of $1,024, $954, and $717, respectively, for MoM, CoC, and CoP compared to MoP. Year of procedure had no significant association with total costs, nor did the hospital type or size, with bearing surface dominating the regression results. The West census region had a large impact on total cost at $2,981 compared to the referent South region. The only volume variable that had a significant association with total cost was the cumulative volume/bed at $75.

Patient clinical and demographic variables showed substantial impact on total costs. A primary diagnosis of AVN was associated with decreased total cost of $276 and the "other" category (a collection of a large number of low prevalence diagnoses) was associated with increased total cost of $790.[bb] The observation of major complications and/or comorbidities was associated with increased total cost of $937. Discharge to a rehabilitation facility was associated with $770 increased total cost. Each increase in AHRQ comorbidity count by 1 was associated with $185 increased total hospitalization cost and each additional day of hospital stay was associated with $1,514 increased total cost. Each 10-year increase in patient age was associated with $194 decrease in total cost while being female was associated with a decrease in cost of $151.

[aa] This large estimate appears related to 1/3 of the discharges collectively labelled "other" being either to a psych facility or to simply "other institution" and having the highest average cost of all categories.

[bb] The "other" category was populated with some very high cost diagnoses in significant number (traumatic arthropothy, 808; congenital deformity, 414; and rheumatoid arthritis, 475) that skewed this category's cost value high.

Discussion

Figure 5.1's depiction of evolving use of the 4 primary bearing surfaces across the study period is consistent with clinical observations in the literature. Concerns about corrosion and local tissue reactions in metal-on-metal surfaces, which ultimately led to recalls in 2008, 2010, and 2012,[20] have been well documented by researchers and the FDA.[14, 18, 19] These problems led directly to a more than 4-fold drop in MoM prevalence (36.4% to 8.6%) across the study period. Infrequent but catastrophic issues with liner fracture[15] and frequent squeaking concerns[16, 17] contributed to a drop by half in the prevalence (6.7% to 3.4%) of CoC liners across the study period. The market place was left with only CoP bearings for patients and surgeons who desired greater probability of longer implant life through bearing selection than MoP surfaces provided.

Costs Across Time

Figures 5.2 and 5.3 are best viewed in tandem to describe the movements in average (2015$) total hospitalization cost and implant cost across time and bearing surface. Generally speaking, total costs and implant costs moved down over the study period but the relative cost movement differed significantly by bearing surface in both measures of cost.

MoP implant costs decreased by $1,201 whereas the total hospitalization costs dropped by only $898 in constant 2015 dollars (Figs. 5.2 and 5.3). MoM's moves were in similar direction but had a greater drop in implant costs ($1,531) and a smaller decrease in hospitalization cost ($583). The visual impact of MoM's performance vs. the other 3 makes it seem as if the total costs did not move at all while the implant costs dropped precipitously, the latter of which is consistent with the dramatic shift away from

MoM in the market across the study period. CoC's move downward in hospital and implant costs were virtually identical (at around $2,000) and the largest across all surfaces. CoP's relative moves were different in that hospital costs dropped by $1,068 while the implant cost using these surfaces dropped by only $758, the smallest drop of the group by over $400. It is not possible to know from these results what is driving these relative changes, but case-mix shifts that are associated with bearing surface choice seem a likely source. CoP's perceived better clinical situation could also be creating demand that helps device manufacturers maintain CoP pricing relative to MoP against externally payer-driven price pressure.

Regression results confirmed the bivariate results on bearing surface as all 3 were more costly than MoP, both in implant costs and total hospitalization costs. When controlling for covariates, implant costs still trended down across time although the results were only statistically significant in 2008 ($-371) and in 2012 ($-907), compared to referent 2007. The steady downward trend in total costs in bivariate analysis was not confirmed with significant results in regression analysis. The time trends observed in this study may not be universal but study period dependent. The overlap with the significant market moves in the different bearing surfaces likely play a major role in the time observations. For instance, the trends in bearing surface cost may not extrapolate backward in time depending on exactly when the market began responding to the clinical problems.

Costs Across Hospital Variables

Regression results clearly display how little association there was between hospital variables and either implant cost or total hospitalization cost. Strong bivariate relationships indicating urban nonteaching hospitals and large hospitals incurred higher implant costs were not borne out in regression analysis. This is most likely because the regression accounts for clustering of observations within hospitals which bivariate analysis cannot do; the net effect is to increase standard errors on hospital-level variables. The same general trends were observed although none of the results were statistically significant and $p<0.05$ is not a large barrier to overcome in an analysis with 121,128 cases. The only hospital descriptor that showed statistical significance was the West region at $2,981 higher in total costs compared to the referent South. Like the bivariate analysis, the Northeast did show lower implant costs in the regression but the result was not statistically significant. The only volume variable with significant results was the THA concentration variable, cumulative volume/bed, where each unit move upward in this measure was associated with a $75 increase in total costs.

Clinical and Patient Demographic Covariates

Significant associations between implant cost and primary diagnosis (-$660 for fracture) seem odd. The same can be said of the associations observed between implant cost and age ($-95/each 10 years of increased age) and sex ($-87 for female). Review of Chapter 4 indicates that younger patients and those with a primary diagnosis of fracture were more likely to use the cheaper MoP surfaces. Female patients were also more likely to use the cheaper MoP surfaces. Even controlling for these effects in a multivariate cost regression may not be enough to account for all cost behavior in the complex environment found in this study.

Not surprisingly, total cost showed large associations with patient-level covariates, greater than with bearing surface in some cases. For example, an MS-DRG notation of major complications and/or comorbidities was associated with an increased total hospitalization cost of $937, while discharge to a rehab facility was associated with $770 increased total cost. In this context, it seems that discharge to a skilled nursing or critical care facility (SNF/CCF) also would have been associated with greater total cost but such was not the case. Not surprisingly, each day of additional hospital stay was associated with a large increase in total cost ($1,514).

Robinson et al.[45]also observed that hospital variables explained little implant cost variation in their work attempting to describe variation in implant and surgical costs in total hip and knee replacements. They suggested that within-hospital variation might play an unaccounted for role as well. They did not incorporate different bearing surfaces in their analysis, which have just been demonstrated to cause wide variation in hardware cost between surfaces as well as over the 6-year time span of this study. Further since implant costs accounted for 41.0% of total costs in this study (calculated from Table 5.2), unexplained variation in implant costs will lead to unexplained variation in total costs. Robinson et al.[45] found that implant costs accounted for 33.5% of total costs, but they were much more restrictive on what implant related hardware items were included in their assessment.

Conclusions

It is surprising that neither hospital descriptors nor surgical volume were often significantly associated with either implant cost; this finding directly refutes the hypothesis that larger and/or higher surgical volume hospitals would report lower implant costs (as proxy for price paid). This lack of association may be an indictment of reported cost being a suitable proxy for price paid: at the very least, hospital accounting practices can make reported implant costs a variable of doubtful precision. Further, group purchasing organizations may impact pricing outside traditional drivers of purchaser pricing power but this is not testable from the data used.

Given the considerable although seemingly illogical association between some of the clinical and patient demographic indicators and implant costs, it also seems possible that endogeneity is playing a role as well. Chapter 4 of this dissertation, in finding strong associations between patient age, clinical variables, and implant choice seems a likely source of unexplained variation between costs and implant choice. Lower cost implants were generally utilized more often in older and sicker patients, i.e., those with potentially fewer years of life remaining.

Limitations/Future Research

As in any secondary data analysis, data generated primarily for remuneration purposes can be inaccurate or incomplete in reporting information that is not essential for reimbursement. Further, the method detailed in Appendix A to describe implant and total cost is imperfect. The key study inclusion criterion was the presence of an ICD-9-CM procedure indicating a total hip arthroplasty; thus, all included patients had undergone a hip implant procedure. Accordingly, I adopted an aggressive “opt in” policy to categorize cost items as hip implant related whose description included the word “implant” but nothing of more helpful detail. This choice could have overstated hip implant costs, but if so, systematic bias between surface choice and other variables seems unlikely. Finally, because only 46.2% of eligible cases included an identified bearing surface, the potential for surface reporting bias exists although Chapter 4’s analysis indicates such bias does not appear to be present.

Future work in this area might be better served by focusing on even more detailed use and cost descriptives to quantify case-mix association with bearing choice and assess how that influences cost. Also, analyzing implant costs and non-implant costs (which together add up to total hospitalization costs) should be more sensitive to the multivariate regression analysis variables than considering implant cost and total cost. This should better manage the issue with physicians tending to use lower cost devices in sicker and older patients, i.e., those with fewer years of life remaining (noted in Chapter 4).

Despite substantial remaining opportunities, this chapter has nonetheless established that implant hardware costs vary widely across implant choice and across 2007-2012 as well. This work also established that CoP bearings have maintained their premium over the historical baseline MoP bearings, an important consideration given that clinical issues have taken MoM and CoC bearings out of favor.

Table 5.1: THA procedures in study by surface and year

Year	MoP	MoM	CoC	CoP	Total
2007	7,299	6,018	1,101	2,101	16,519
2008	7,614	6,460	662	2,466	17,202
2009	8,905	6,713	863	3,402	19,883
2010	10,578	5,512	967	4,658	21,715
2011	11,587	3,552	947	6,872	22,958
2012	11,766	1,974	788	8,323	22,851
Total	57,749	30,229	5,328	27,822	121,128

*Procedures with $0 cost and outside +/- 99% range have been excluded.

Table 5.2: 2007-2012 THA cost descriptions - 2015$

	Implant cost	Total hospitalization cost
Mean	7858	18273
Standard Deviation	3596	5520
Minimum	631	8293
Lower Quartile	5349	14532
Median	7084	17283
Upper Quartile	9788	20858
Maximum	24161	50931

Table 5.3: Descriptive statistics of continuous independent regression variables

Variable	Mean	Std Dev	Minimum	Lower Quartile	Median	Upper Quartile	Maximum
Staffed beds	398	204	23	256	382	532	1171
Annual THA volume	478	499	1	171	353	533	2504
Annual volume/bed	1.94	4.06	0.0088	0.52	0.85	1.32	21.22
Cumulative THA volume	1770	1880	2	525	1130	2341	9687
Cumulative volume/bed	6.45	12.38	0.0132	1.59	3.06	5.37	82.09
Patient age, years	64.3	12.19	18	56	65	73	89
Comorbidity Count	1.8	1.45	0	1	2	3	12
Length of stay, days	3.17	1.5	0	2	3	3	30

Table 5.4: Descriptive statistics of discrete independent regression variables

Variable	Value	Sample prevalence
Bearing surface	MoP	57749 (47.7%)
	MoM	30229 (25.0%)
	CoC	5328 (4.40%)
	CoP	27822 (23.0%)
Year	2007	16519 (13.6%)
	2008	17202 (14.2%)
	2009	19883 (16.4%)
	2010	21715 (17.9%)
	2011	22958 (19.0%)
	2012	22851 (18.9%)
Hospital type	Urban nonteaching	57014 (47.1%)
	Urban teaching	54875 (45.3%)
	Rural	9239 (7.63%)
Hospital size	Large	92942 (76.7%)
	Medium	18164 (15.0%)
	Small	10022 (8.27%)
Hospital US census region	South	35226 (29.1%)
	Midwest	19360 (16.0%)
	Northeast	35515 (29.3%)
	West	31027 (25.6%)
Primary Payer	Medicare	60019 (49.6%)
	Private insurance	52907 (43.7%)
	Medicaid	3155 (2.60%)
	Other	5047 (4.17%)
Primary diagnosis	Osteoarthritis	105399 (87.0%)
	AVN	6631 (4.5%)
	Fracture	3652 (3.0%)
	Other	5446 (4.5%)
MS-DRG	470	117000 (96.6%)
	469	3747 (3.09%)
	Other	381 (0.31%)
Discharge status	Home	26751 (22.1%)
	Home health care (HHC)	54111 (44.7%)
	Rehab facility	11972 (9.88%)
	SNF/CCF	27201 (22.5%)
	Other	1093 (0.9%)
Sex	Male	53750 (44.4%)
	Female	67378 (55.6%)

Table 5.5: Multivariate regression on costs (2015$)

Parameter	Value or unit	Implant costs	Total hospitalization costs
Bearing surface-ref: MoP	MoM	717**	1024**
	CoC	685*	954*
	CoP	427	717*
Year-ref: 2007	2008	-371*	-231
	2009	-256	386
	2010	-245	483
	2011	-418	474
	2012	-907**	123
Hospital type-ref: Urban nonteaching	Urban teaching	-682	-568
	Rural	-912	-209
Hospital size-ref: Large	Medium	-1051	-658
	Small	-866	-373
US census region- ref: South	Midwest	605	291
	Northeast	-704	232
	West	457	2981**
Staffed beds	100 bed increments	209	293
Annual THA volume	100 surgery increments	-155	n/a
Annual volume/bed		245	n/a
Cumulative THA volume	100 surgery increments	n/a	-38
Cumulative volume/bed		n/a	75**
Primary payer-ref: Medicare	Private insurance	174	162
	Medicaid	-210	110
	Other	-213	-67
Primary diagnosis- ref: Osteoarthritis	AVN	-237	-276*
	Fracture	-660**	-231
	Other	132	790**
MS-DRG- ref: 470	469	-114	937**
	other	-732	-598
Discharge status- ref: Home	Home health care	-29	-18
	Rehab	150	770*
	SNF/CCF	-119	27
	Other	-1041*	555
AHRQ comorbidity count		11	185**
Length of hospital stay	Days	22	1514**
Age	10 year increments	-95**	-194**
Sex- ref: Male	Female	-87**	-151**

*Indicates statistically significant estimate at $p<0.05$.

**Indicates statistically significant estimate at $p<0.01$.

Implant cost basis is 7858.

Total cost basis is 18273.

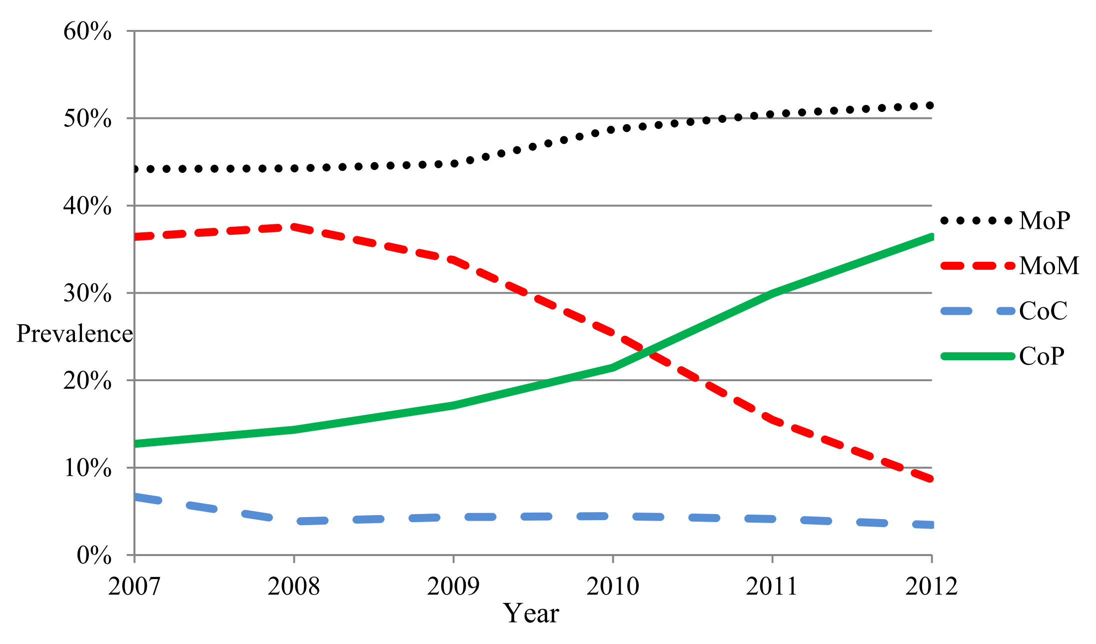

Figure 5.1: THA bearing surface prevalence across time in PRD, 2007-2012
Source: Author's analysis of PRD

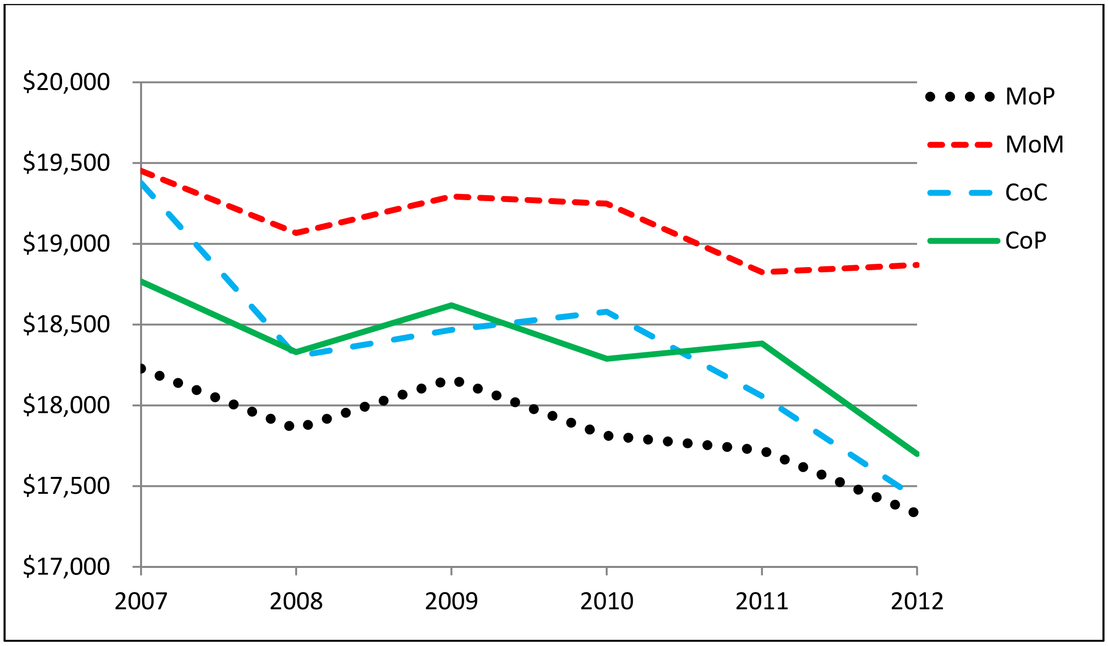

Figure 5.2: Total hospitalization cost by surface and year (average 2015$)
Source: Author's analysis of PRD

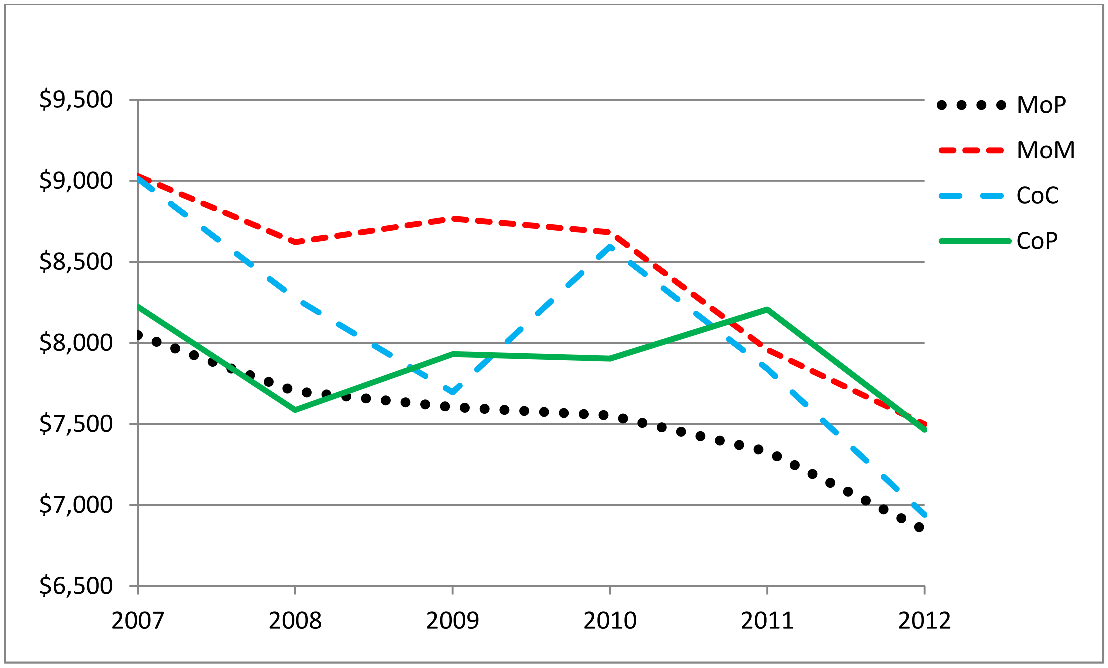

Figure 5.3: Implant cost by surface and year (average 2015$)
Source: Author's analysis of PRD

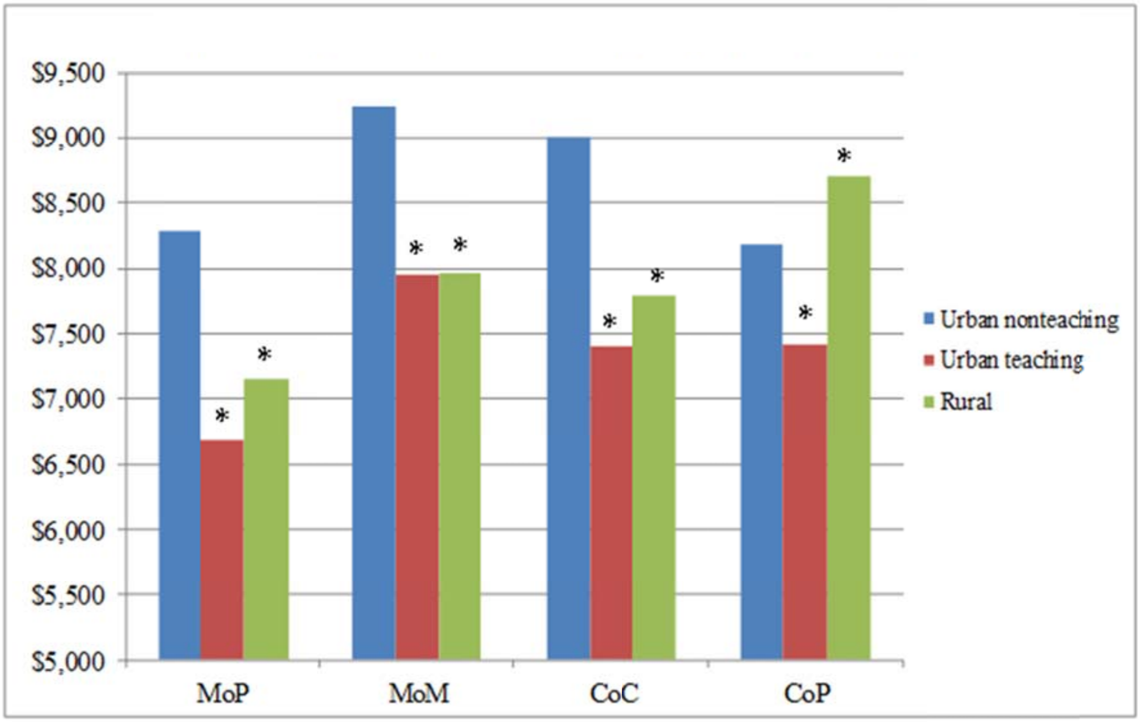

Figure 5.4: Implant cost by surface and hospital size (Average 2015$)
Source: Author's analysis of PRD

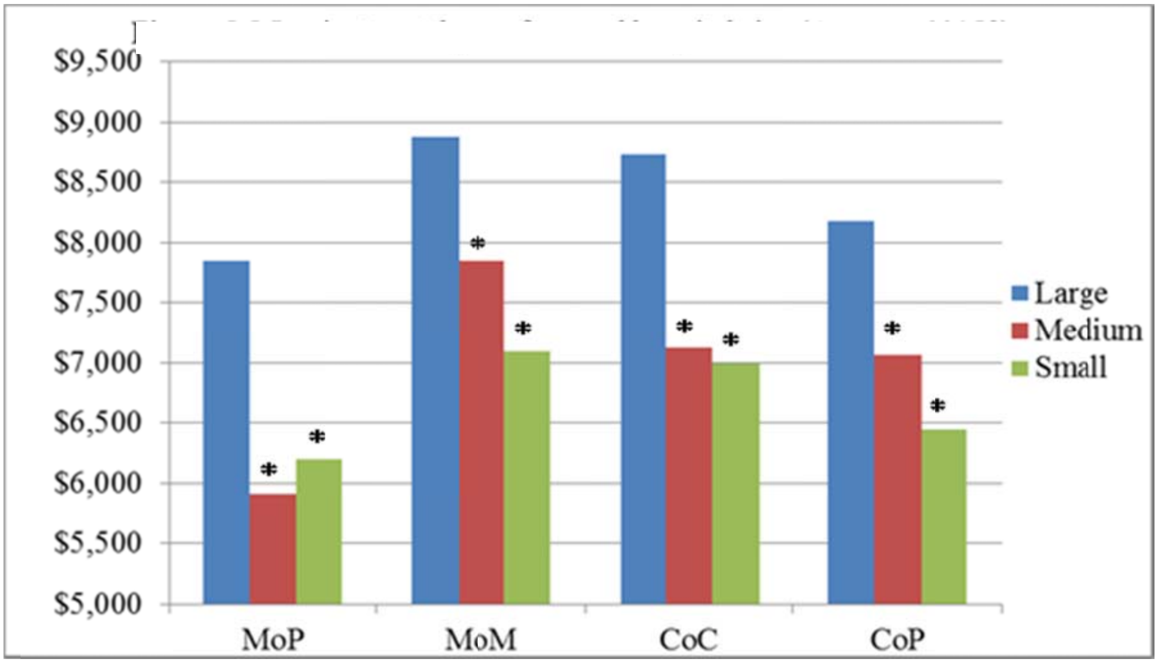

Figure 5.5: Implant cost by surface and hospital size (Average 2015$)
Source: Author's analysis of PRD

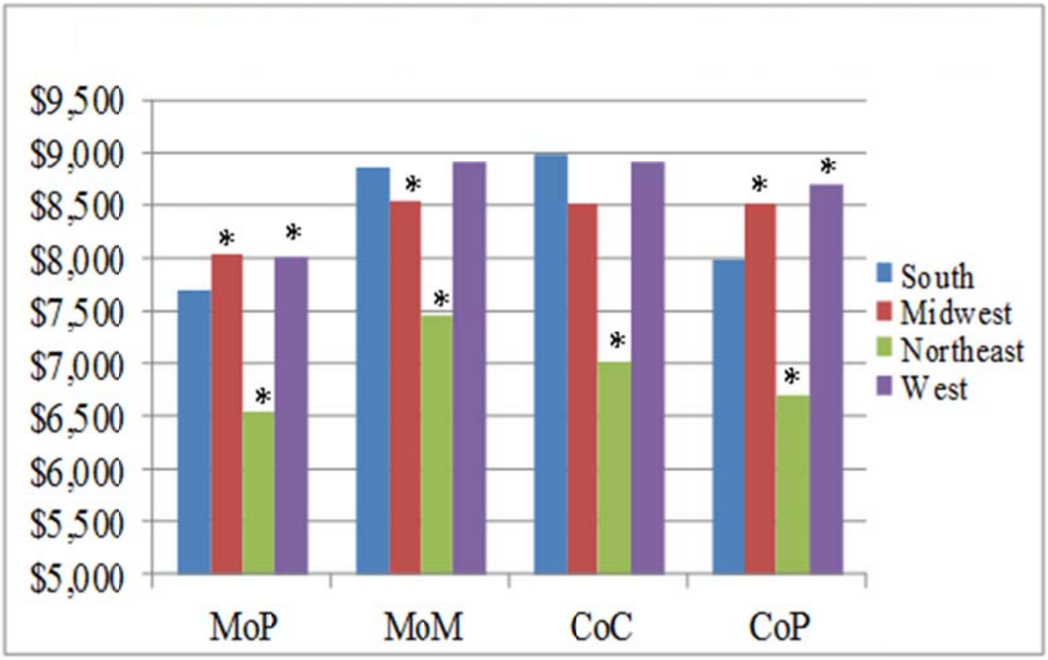

Figure 5.6: Implant cost by surface and US census region (Average 2015$)
Source: Author’s analysis of PRD

CHAPTER 6: COST-EFFECTIVENESS DRIVERS IN CHOOSING CERAMIC VS. METAL HEADS IN TOTAL HIP ARTHROPLASTY

Primary total hip arthroplasty (THA) is a common surgical procedure that most often is employed to overcome steadily worsening functionality and quality of life due to osteoarthritis of the hip joint. In 2010 THA comprised 1.3% of all non-diagnostic, non-pregnancy-related hospital procedures in the US, increasing to 2.6% of procedures in the ≥65 year-old population.[3][cc] Despite its relatively high cost, the cost-effectiveness of this medical procedure has been demonstrated vs. available alternatives.[4, 6, 7]

Despite long success of the Charnley metal-on-polyethylene (MoP) implant introduced in the 1960s, scientists have worked to reduce the need for revisions, often focusing on the load bearing surfaces. Unfortunately, results have been mixed. Metal-on-metal bearings showed early promise to reduce wear; however, concerns about corrosion and local tissue reaction, which ultimately led to recalls in 2008, 2010, and 2012,[20] have been well documented by researchers and the FDA.[14, 18, 19] Ceramic-on-ceramic (CoC) bearings also showed excellent lab wear results but squeaking[16, 17] and infrequent but catastrophic bearing fracture[15] have tempered interest in this surface as well.

Highly cross-linked (XL) polyethylene liners had virtually replaced high molecular weight polyethylene liners in use with metal heads by 2007.[21] Coincident with

---

[cc] Author's analysis of 2010 National Hospital Discharge Survey Results, conducted by the Centers for Disease Control and Prevention. http://www.cdc.gov/nchs/fastats/inpatient-surgery.htm accessed 9 January 2016.

reduced wear of XL polyethylene liners by metal heads, a new mode of metal failure has been discovered—fretting and corrosion at the junction of a cobalt chromium femoral head and the trunnion in a MoP hip.[24, 25] Multiple failure modes on the bearing surface types that have been in use the longest have prompted many surgeons to move toward ceramic-on-polyethylene (CoP) to increase implant life before revision is required.[26] Much lower wear rates for CoP have been reported than for MoP both in the laboratory and when conducting in vivo (in the body) evaluations.[30, 31, 35] In 2001, twenty-year cumulative CoP revision rates were reported to be 25 revisions/100 THAs (prospective evaluation, no direct comparison to MoP).[33] Pulikottil-Jacob et. al. (2015),[50] modelled 8-year survival data to report a 20-year cumulative CoP average revision rate of 6.8/100 THAs (compared to 9/100 THAs for MoP) based on retrospective review of the National Joint Registry for England and Wales (NJR). The reduced revision rates that CoP bearings are expected to bring can take quite long to be realized. With CoC and MoM no longer regarded as viable alternatives, and despite the lack of well-defined improvements from CoP implants, aggressive movement toward CoP has already begun as reported in chapter 4.

Although costs for all implants have been changing recently, CoP bearings were more expensive than MoP bearings as of 2012 (reported in chapter 5). In light of ongoing movements toward CoP, cost-effectiveness of the higher priced bearings must be assessed even in the absence of complete clinical information. Bozic et. al.[48] framed the question well in a theoretical exploration of the impact of age and implant cost on cost-effectiveness of a general “alternate” bearing surface vs. MoP. Unfortunately, as very little comparative data on either cost or revision rates were available at that time, that work did not provide much guidance of a practical nature.

Pulikottil-Jacob et. al. (2015)[50] used revision rates from the NJR and patient utilities from a patient-reported outcome database in a cost effectiveness analysis stratifying patient populations by age (60, 70, 80 years) and sex, looking at several formulations including CoP vs. MoP. They reported that cost savings and utility gains were too small to make a clear case for the more expensive devices, although they acknowledged that the registry data and their source of patient utilities may not have been robust enough for this type of analysis.

This paper explores the total lifetime cost and utility performance of CoP bearings compared to MoP, stratifying on the most important considerations of age, CoP revision rate improvement vs. MoP, and CoP vs. MoP implant cost difference. The purpose of this analysis is to define conditions under which CoP offers sufficient improvement in incremental cost-effectiveness ratio (ICER) expressed as dollars per quality adjusted life year ($/QALY) to be considered preferable to MoP. The incremental cost-effectiveness ratio is the difference in lifetime costs divided by the difference in lifetime patient utility values, which allows one to incorporate both costs and quality of life values into one measure comparing the two medical devices for each condition tested.

Methods

This study involved the following steps to collect inputs for the cost-effectiveness analysis: identifying relevant cases, collecting revision rates and health utility values from the literature, and determining THA implant costs, THA non-implant hospitalization costs, revision hospitalization costs, and costs for the surgeon, anesthesiologist, and rehabilitation. Finally, Markov models were created to determine the ICER for relevant age sub-populations at observed population-based implant and hospitalization costs and literature-sourced revision rates.

Sample Source

Patient-level discharge data for THA procedures in 2012 were obtained from the Premier Research Database (PRD). The PRD includes clinical, demographic, and cost data from a large nationwide sample of general, non-federal acute care US hospitals that chose to participate. CoP and MoP implants were identified using International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) procedure codes. Patients with a 2012 primary THA with either a CoP or MoP implant that were 45 years or older at the time of the primary surgery were subject to inclusion in the study.

Cost Item Determination

The PRD contains descriptive detail for each item and service incurred during each hospitalization. Appendix A details how implant costs and non-implant hospitalization costs were determined for each discharge record in this study. All costs were indexed to 2015 dollars using the Bureau of Labor Statistics consumer price index inflation calculator.[dd] The cost for each item is described as the cost to deliver the item to the point of use and thus can be expected to include the purchase price of the item and overhead or other costs a given hospital chooses to apply to that item.

Sample Characteristics

A total of 52,181 primary THAs were identified in the 2012 PRD across 471 US hospitals and 44 states where the patient was at least 45 years old. Primary THA case

---

[dd] The US Department of Labor, Bureau of Labor Statistics charts movements in consumer prices and provides a CPI Inflation Calculator that can be used to equate buying power across different time periods. Accessed on 1 February 2016 at http://www.bls.gov/data/inflation_calculator.htm.

records were selected based on one incidence of the 81.51 ICD-9-CM procedure code and no other arthroplasty-related codes except for ones denoting bearing surface. Of the 52,181 primary THAs, 28,828 (55.2%) were excluded due to missing ICD-9-CM bearing surface codes. An additional 2,944 THA records were excluded for the following reasons: MoM surface (n=2,112; ICD-9-CM code 00.75), CoC surface (n=726; ICD-9-CM code 00.76), or no cost data (n=106). To minimize the impact of data entry errors, cases were excluded if the total hospitalization costs were in the top or bottom 1% of the nonzero total cost distribution. The final analytic file included 19,165 patients (aged 45+) that underwent a primary THA at one of 377 PRD hospitals with a CoP (n=7,734; ICD-9-CM code 00.77) or MoP (n=11,431; ICD-9-CM code 00.74) bearing surface in 2012.

Revision procedures were selected from the PRD based on the presence of ICD-9-CM procedure codes 81.51, 81.53, 00.70, 00.71, 00.72, or 00.73. Applying the same age and cost value inclusion criteria as in primary THA cases, the PRD contained 7,955 revision cases in 2012.

Cost Data

The PRD was the source for CoP and MoP implant costs and non-implant hospitalization costs for the entire study population of 19,165 THA patients, and hospitalization costs for 7,955 revision cases. Across the entire study population average implant cost was $6,846 (95% CI: $6,788-$6,903) for 11,431 MoP surgeries and $7,466 (95% CI: $7,384-$7,548) for 7,734 CoP surgeries; thus, CoP implants were more expensive than MoP by $620 (95% CI; $523-$717, $p<0.0001$). A more conservative case where the CoP implant costs $1,000 more than the MoP implant was included, consistent with the implant cost difference used in a cost-effectiveness analysis reported by Gioe et. al.[49] Hospitalization costs not associated with the implant hardware averaged $10,393 (95% CI; $10,337-$10,449) for the entire study population of THA cases. Total hospitalization costs for each revision case were calculated to be $23,149 (95% CI: $22,885-$23,413).

Surgeon fees were extracted from the Centers for Medicare and Medicaid Services 2015 Online Physician Fee Schedule.[ee] The surgeon's fee was $1,739 and $1,990, respectively, for a primary THA and a revision surgery. Anesthesiologist fees for both procedures were estimated to be $857 (personal conversation with Susan Odum, (OrthoCarolina Research Institute: Charlotte, NC). Post-surgery rehabilitation costs out to 120 days post-surgery were estimated based on the Rand Institute's work on Medicare beneficiary costs in 2005,[60] updated to $19,314 (2015$), and applied to both THA and revision procedures. Table 6.1 summarizes all cost information used in this study.

Health States and Utility Values

The model used here incorporates the following exhaustive and mutually exclusive unique end-states in each period, with all patients starting from a successful primary THA procedure in the first period:

- Successful THA
- Revision required
- Subsequent revision required
- Death as a result of revision surgery
- All cause death

[ee] The THA surgeon's fee was taken from Current Procedural Terminology (CPT) code 27132 and revision surgeon's fee was taken from CPT code 27134-2015 Medicare payment rates. https://www.cms.gov/apps /physician-fee-schedule/ accessed on 24 February 2016.

Figure 6.1 shows the Markov model paths that patients can take in the model used in this study, indicating that patients can cycle indefinitely in successful primary THA, successful revision, or successful re-revision until death. The model described in Figure 6.1 is consistent with neither surgical complication during THA procedures nor THA-related death rates differing based on the choice of MoP or CoP bearings; no published research has been found to bring this into question.

Health utility values were used for 3 non-expiring states in this model based on a review of previous cost-effectiveness work and literature on assessment of health utilities.[61-64] These health states, and their associated utility values, are osteoarthritic hip before THA (0.50), following successful THA surgery (0.92), and following successful revision surgery (0.80).

Health State Probability Data

Two data sources were used to determine baseline revision rates for use in the study. The HealthEast Joint Replacement Registry (HEJRR)[43] was the data source for the baseline MoP revision rates defined in the Markov model. The 17 years of observed MoP data in the HEJRR is the longest span of survival results found thus far. The 20-year cumulative revision rate for MoP was estimated to be 14.5 revisions/100 THAs, equivalent to an incremental annual failure probability of 0.00724. The HEJRR contained only 5 years of results for THAs using CoP surfaces and was thus not used as a source for CoP revision rates.

The CoP revision rate base case was constructed by reviewing Kaplan-Meyer figures contained in the NJR.[65] Survival rates of CoP and MoP are compared directly out to 10 years. Considering all four reported methods of fixation (cemented, uncemented, hybrid, and reverse hybrid), a procedure volume-weighted revision rate improvement of 25% was observed for CoP over MoP at 10 years. For this analysis, the 25% difference was applied to the baseline MoP rate found in the HEJRR as follows to estimate a 20-year CoP cumulative revision rate: 14.5 x 75% = 10.9 revisions/100 THAs, equal to an incremental annual failure probability of 0.00544. For comparative purposes, cases using a 10% improved CoP revision rate were also modelled (14.5 x 90% = 13.0 revisions/100 THAs at 20 years post-THA equating to an annual failure probability of 0.00650).

Gender-adjusted probabilities of all-cause death were calculated using the Social Security Administration Period Life Tables[66] and incorporated into the Markov model for each patient age (50, 60, 70, 80, and 87 years old). The probability of death associated with revision hospitalization was calculated from the PRD to be 0.0069 using the discharge status variable provided for each record. Table 6.2 summarizes the health state utility values and the health state probabilities used in this study.

Analytical Model

A Markov decision model (TreeAge Software, Williamsburg, MA) was designed to show how age, CoP-MoP implant cost difference, and CoP-MoP revision rate difference drive lifetime cost, quality-adjusted life years, and the resulting incremental cost-effectiveness ratio (ICER, $/QALY). See Figure 6.2 for the decision tree. This study is designed to frame the question in terms that can be operationalized for consideration by both practitioners and researchers. The cost of each THA case included the non-implant hospitalization costs, the surgeon’s fee, the anesthesiologist’s fee, rehabilitation costs, and the implant cost for the appropriate bearing choice. The cost for each revision case included the cost for the entire hospitalization, the surgeon's fee, the anesthesiologist's fee, and the rehabilitation costs.

In a Markov model, individual theoretical patients are transitioned through Markov cycles using mutually exclusive health states,[67] in this case: primary total hip arthroplasty, a potential revision hip procedure, a potential re-revision procedure, and finally, death. A 1-year cycle time was used with 10,000 simulated patients each cycling through the model each year after the initial MoP or CoP THA until death. The outcomes are revisions, and the QALY gains from CoP use are the result of the probabilistic avoidance of revisions. Both future costs and health outcomes were discounted to the present as recommended by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR)[68] at a value of 3.5%/year as recommended by The Panel on Cost-Effectiveness in Health and Medicine[69] convened by the US Public Health Service.

The patient ages considered in this study were 50, 60, 70, 80, and 87 years at the time of the index THA procedure. The implant cost premiums of CoP vs. MoP surfaces were $620 and $1,000. CoP revision rate reduction of 25% and 10% below MoP's revision rate were tested, with annual MoP revision probability fixed at 0.00724 in each case. All other drivers of cost and outcome (both initial and future) were kept constant so that only the above variables drove the results. Frontier curves of ICER vs. age and CoP implant premium cost were created for each CoP revision rate improvement. A balanced orthogonal design with additional age cases was used to evaluate all possible combinations of the 5 chosen patient ages, 2 annual CoP revision probabilities, and 2 implant cost differences. Table 6.3 shows each of the 20 cases on which a Markov model was run.

Results

Table 6.3 shows the results from each case modelled following the test conditions for each case. For each of the 20 modelled cases, these results include the lifetime costs (2015$) of MoP and CoP patients and the difference between each cost value, the lifetime utility values (2015 QALYs) of MoP and CoP patients and the difference between each utility value, and the ICER (Δ$Cost/ΔQALY)[ff] for each case. An initial assessment of Table 6.3 finds the CoP-MoP cost difference to be negative in 6 cases. In these most favorable cost cases, the lifetime discounted cost savings of avoiding revisions in the CoP population vs. revisions incurred in the MoP population more than offset the higher initial cost incurred for the CoP device. In 1 of these 6 cases, the more expensive CoP device was used in the youngest age (50) while the other 5 were spread over 50, 60, and 70 yo patients using the less expensive CoP device. All but one of these negative lifetime cost difference cases were observed at the more aggressive 25% reduced CoP revision rate.

Figures 6.3 and 6.4 best describe the key output measure of the study, the ICER at each of the cases, for the 25% improved CoP revision and the 10% improved revision rate, respectively. These frontier diagrams describe the sensitivity of the ICER to different CoP implant costs and to the relative reduction in the revision rate from the CoP implant at different patient ages. One can work back and forth between these figures and Table 6.3 to get exact Markov model values for specific cases of interest. Note in all cases the MoP baseline 20-year cumulative revision rate was held constant at 14.5/100 THAs (0.00724 annual revision probability).

[ff] Δ stands for delta and represents difference, with MoP's cost and utility being subtracted from CoP's cost and utility, respectively, in the numerator and denominator.

Figure 6.3 shows results for the more aggressive 25% reduced CoP revision rate case (0.00544 annual revision probability; 20-year cumulative rate of 10.9 revisions/100 THAs). At this revision rate, the highest ICER observed was $100,584/QALY for an 87 year-old that incurred a $1,000 incremental cost for a CoP implant. The second highest ICER observed was $43,055/QALY for an 87-yo that incurred $620 incremental cost for a CoP implant. Seven of the total 10 cases displayed in Figure 6.3 had ICERs that were either below or effectively $0. At 25% reduction in CoP revision rate, the ICER was quite insensitive to age until 70 years at either CoP implant cost premium. Only at patient age of 80 years did the CoP implant cost premium and age begin to sharply increase the ICER.

Figure 6.4 shows results for the more conservative 10% reduced CoP revision rate case (0.00650 annual revision probability; 20-year cumulative rate of 13.0 revisions/100 THAs). Although the frontier curves are of similar shape to those in Figure 6.3, note the very different ICER scale. At the $1,000 incremental COP implant cost, the highest 2 ICER values are seen at $375,905 and $238,132, respectively, for an 87 and 80 year-old patient. At ICERs of $153,470 and $101,316 came the 87 and 80 yo patient, respectively, both at the $620 CoP incremental implant cost. The same flatter profile with patient age up through age 70 was seen at this more conservative CoP revision improvement as in the more aggressive case described in Figure 6.3, followed by steep increases in the cost of each additional QALY past 70.

Discussion

The cost society should be willing to incur for an increased quality-adjusted life year will likely never be a fully settled issue, nor should it be. Braithwaite et. al.[70]

investigated whether the traditional value of $50,000 is consistent with modern societal preferences in the US as to how much cost society should be willing to incur for an additional QALY. They concluded that more appropriate ICER bounds to consider were $109,000/QALY-$297,000/QALY. At a threshold of $150,000/QALY, Figure 6.3 reveals that using CoP implants would be justified under all cases at the more aggressive 25% reduced CoP revision rate assumption. If one chooses $250,000 as the ICER threshold, then all age and CoP implant costs are justified at 10% reduced CoP revision rate except for the $1,000 CoP cost delta in an 87 yo patient, which had a cost of $375,905 per QALY gained.

Several general observations stand out and should be noted. Across the 20 modelled conditions, the lifetime CoP-MoP cost difference ranged from -$851.37 to $824.36 while the observed lifetime health utility difference ranged from 0.0022 QALYs to 0.0516 QALYs. To large degree, these small differences are an artifact of a low probability event (a future revision) and a small incremental cost decision at the initial surgery compared to the overall costs both at the initial THA and in the event of a revision.

Pulikottil-Jacob et al[50] (2015) reported that lifetime cost differences (<£3000 = $4,344[gg]) and utility differences (<0.0039 QALYs) prevented them from making a recommendation of CoP over MoP. Their observed cost differences were much higher than observed here but their utility differences were not dissimilar. Their CoP vs. MoP revision rate difference of 25% (my calculation from their Kaplan-Mayer figures) came from the NJR as did the more aggressive case considered in this study, but their baseline

---

[gg] Using current conversion rate of $1.44788/£ http://www.xe.com/currencyconverter/convert/?From=GBP&To=USD accessed 17 March 2017. Should be used for broad illustrative purposes only.

revision MoP rate was approximately 40% lower than the revision rate used in this study. They concluded that with cost and QALY differences as small as they observed, the results could be too sensitive to relatively small changes in non-implant costs and quality of life estimates. Their very low ICER threshold of £20,000 ($28,958) also was a likely stumbling block to their being able to make a recommendation but low observed cost and utility differences between the implants here could prompt the same concerns.

Conclusions

This study has concluded that, under all cases considered here except the more expensive CoP implant, the lower CoP revision improvement, and the oldest patient age, the choice of CoP will be justified in cost/QALY gained at the higher threshold of $250,000 for each QALY. Consistent with more life years remaining, the ICER results up to age 70 are much less sensitive to age at either implant cost difference although the revision rate assumption does have a significant impact on the actual ICER values observed.

The study presented here builds from the framework put forth by Bozic in 2006[48] that allows for the reality that key economic decision criteria of revision rates are yet to be clearly defined and that implant costs vary. This study has applied greater context by using population-based cost information for both the surgical costs and the implant hardware (separately vs. being combined as in most administrative databases). It has also brought in revision rates from the literature, and thus produces information that is more actionable by researchers. Unlike previous studies, this study also incorporated rehabilitation costs for both THAs and revisions. Including these into the total financial cost of a future revision event increased the cost of this event by almost 75%.

The reality of the base case “lesser performing” MoP bearing surface failing only 9-15 times in 20 years for every 100 surgeries (depending on literature source) renders the future cost and utility loss rare. This rare occurrence of a very disruptive need for another hip surgery and significant cost both makes it difficult and important to continue to quantify the costs and utility impact of the available choices.

The most important area of future work in this area is to continue to improve and quantify the useful life of hip implants. This should then lead to more finely detailed analyses than the one conducted here. The small differences in upfront bearing costs and very small differences seen in QALYs across the lifespan of the theoretical patient populations assessed in this Markov model (or other similar efforts) likely do make the results sensitive to considerations outside those being tested. A good next step would be to take a very limited number of the cases assessed here and methodically determine the impact of variables held fixed in this (and other) studies, like quality of life estimates for example.

Limitations

A key variable to determine participation in this study was whether a voluntarily reported ICD-9-CM code indicated use of a MoP or CoP bearing surface. At just under 50% reporting of this variable, a large segment of the study population is unaccounted for and results may not be representative, although Chapter 4 reported that little evidence of biased reporting was found. This study also did not incorporate cost of lost work or travel costs related to a revision or rehabilitation. At a median patient age of 66 years, the practical impact of this shortcoming should be quite small.

Table 6.1: Markov model cost inputs

Cost Item	MoP	CoP Cost 1	CoP Cost 2
Non Implant Cost-THA[1]	$10,393	$10,393	$10,393
Implant Cost-THA	$6,846[2]	$7,466[3]	$7,846[4]
Surgeon Fee-THA[5]	$1,739	$1,739	$1,739
Anesthesiologist Fee-THA[7]	$857	$857	$857
Rehabilitation Costs-THA[8]	$19,314	$19,314	$19,314
Total THA Cost	$39,149	$39,769	$40,149
Total Hospitalization-Revision	$23,149	$23,149	$23,149
Surgeon Fee-Revision[6]	$1,990	$1,990	$1,990
Anesthesiologist Fee-THA[7]	$857	$857	$857
Rehabilitation Costs-THA[8]	$19,314	$19,314	$19,314
Total Revision Cost	$45,310	$45,310	$45,310

[1] Average value of all non-implant costs for each THA patient, 2012 PRD.
[2] Average value of all implant costs for THA patients with MoP bearing, 2012 PRD.
[3] Average value of all implant costs for THA patient with CoP bearing, 2012 PRD.
[4] Higher CoP cost value used for comparison purposes (Gioe et. al).
[5] 2015 Medicare fee schedule for CPT code 27132.
[6] 2015 Medicare fee schedule for CPT code 27134.
[7] Personal conversation with S. Odum (OrthoCarolina Research Institute)
[8] Bunton et. al.

All costs updated to 2015$ by BLS CPI Calculator

Table 6.2: Markov model health state probabilities and utility values

Description	Value	Range	Comment/Reference
Baseline MoP annual revision probability	0.00724	n/a	Derived from HEJRR[43]
Baseline CoP annual revision probability	0.00544	n/a	25% improvement compared to MoP-derived from NJR[41]
Lower CoP annual revision probability	0.00650	n/a	10% improvement compared to MoP-included for comparison
All cause death probability	Varies by patient age	n/a	Social Security Period Life Tables[66]
Revision death probability	0.0069	n/a	2012 PRD revision population
Osteoarthritis-pre THA QOL	0.50	0.32-0.85	[61-64]
Post successful THA QOL	0.92	0.66-0.98	[61-64]
Post successful Revision QOL	0.80	0.60-0.95	[61, 63]

Table 6.3: Markov model cases and results

Model case inputs			Model case results						
Patient age, yrs	THA CoP Implant cost difference	Annual CoP revision probability	CoP lifetime cost ($)	MoP lifetime cost ($)	ΔCost, CoP-MoP ($)	CoP utility (QALYs)	MoP utility (QALYs)	ΔUtility, CoP-MoP (QALYs)	ICER, ΔCost/ ΔUtility, $/QALY
60	$620	0.00544	$43,420	$44,075	-$655.22	13.6944	13.6589	0.0355	-$18,473
50	$620	0.00544	$44,523	$45,374	-$851.37	16.5825	16.536	0.0466	-$18,282
70	$620	0.00544	$42,715	$42,975	-$259.87	10.3466	10.3265	0.0201	-$12,948
50	$1,000	0.00544	$44,592	$45,141	-$549.66	16.5581	16.5065	0.0516	-$10,646
60	$1,000	0.00544	$43,853	$44,159	-$305.44	13.7321	13.6948	0.0374	-$8,178
50	$620	0.0065	$45,176	$45,177	-$1.11	16.503	16.4834	0.0196	-$57
80	$620	0.00544	$41,686	$41,671	$14.51	6.7811	6.7705	0.0106	$1,370
70	$1,000	0.00544	$42,863	$42,819	$44.04	10.3627	10.3424	0.0203	$2,171
60	$620	0.0065	$44,170	$44,008	$161.94	13.7471	13.7352	0.0119	$13,654
50	$1,000	0.0065	$45,464	$45,027	$437.50	16.4441	16.4246	0.0195	$22,424
80	$1,000	0.00544	$42,009	$41,662	$346.78	6.8787	6.8678	0.0109	$31,897
60	$1,000	0.0065	$44,524	$44,009	$514.57	13.7594	13.7469	0.0125	$41,034
87	$620	0.00544	$41,111	$40,902	$208.73	4.6103	4.6055	0.0048	$43,055
70	$620	0.0065	$43,275	$42,992	$282.30	10.3205	10.314	0.0065	$43,767
70	$1,000	0.0065	$43,354	$42,731	$622.74	10.3582	10.35	0.0082	$75,851
87	$1,000	0.00544	$41,573	$41,036	$537.52	4.6157	4.6104	0.0053	$100,584
80	$620	0.0065	$42,067	$41,689	$378.01	6.8592	6.8554	0.0037	$101,316
87	$620	0.0065	$41,424	$41,034	$390.58	4.6181	4.6156	0.0025	$153,470
80	$1,000	0.0065	$42,291	$41,524	$767.26	6.8841	6.8809	0.0032	$238,132
87	$1,000	0.0065	$41,624	$40,799	$824.36	4.5167	4.5145	0.0022	$375,905

Annual MoP revision probability = 0.00724 in all cases.

$620 = average CoP vs. MoP implant cost difference in THA procedures in 2012 PRD.
$1000 = higher CoP vs. MoP implant cost difference for comparison purposes.
All cost values in 2015$.

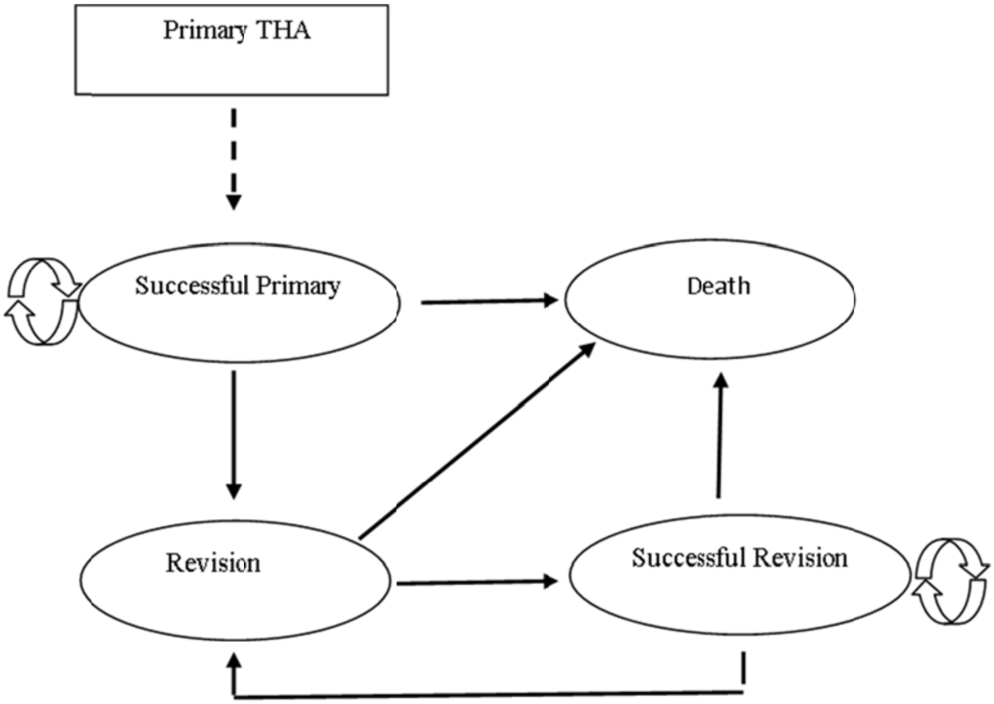

Figure 6.1: Markov model structure

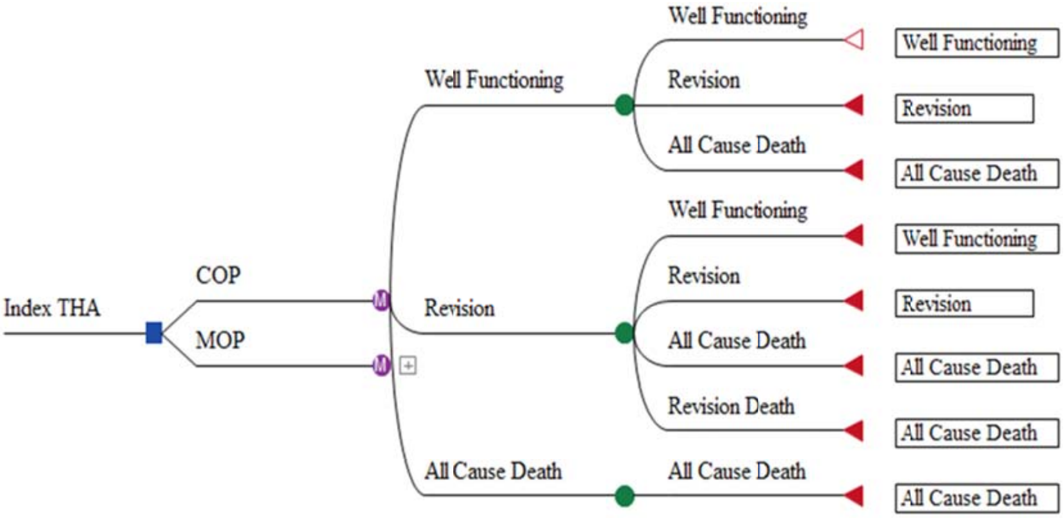

Figure 6.2: Markov model decision tree

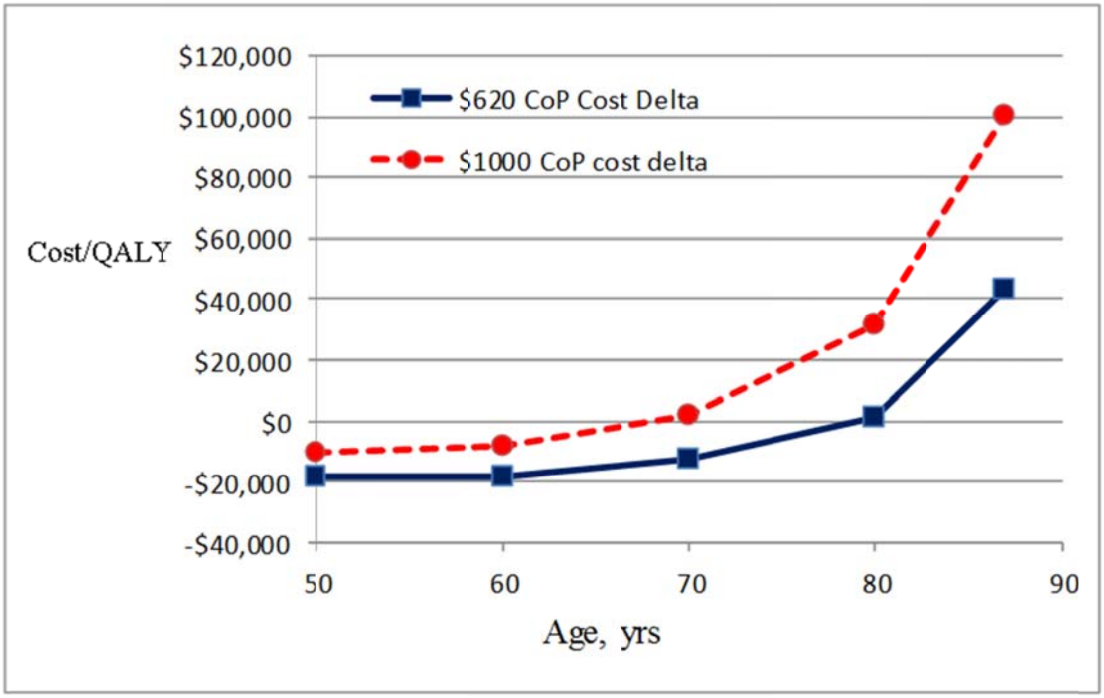

Figure 6.3: Markov results at 25% improved CoP rate

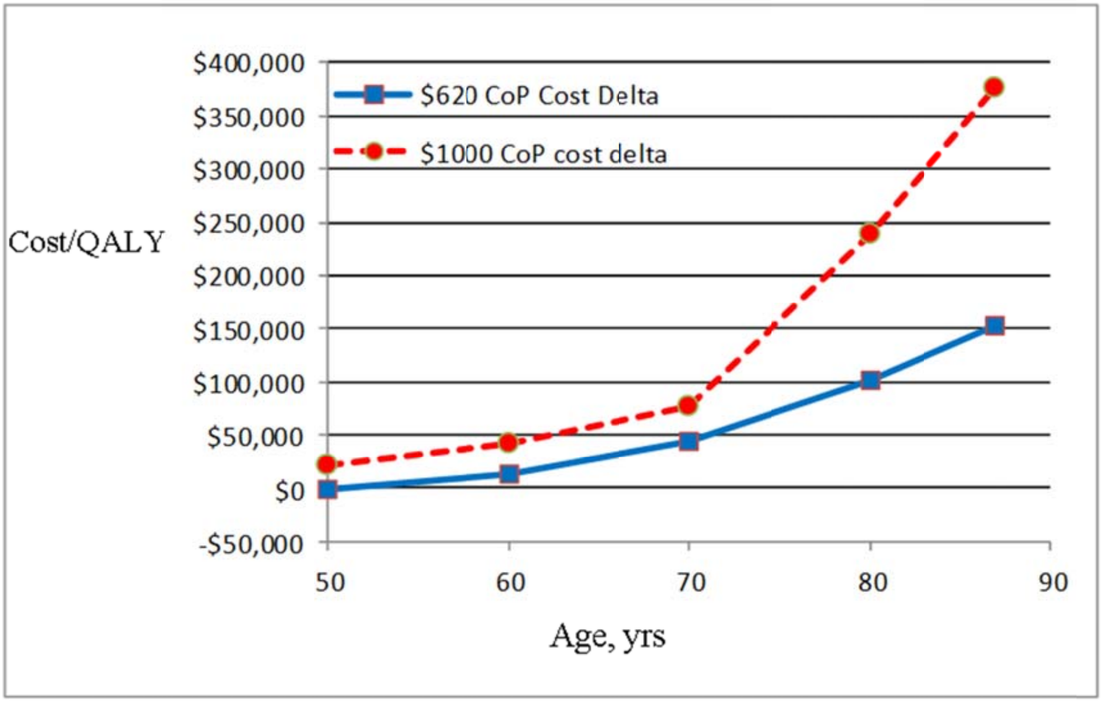

Figure 6.4: Markov results at 10% improved CoP rate

CHAPTER 7: CONCLUSIONS AND POLICY RAMIFICATIONS

Once a decision is made between doctor and patient to undergo a total hip arthroplasty (THA), bearing surface is the most important choice remaining. Hip arthroplasty improvement efforts have focused extensively on these surfaces, with the presumably longer-lasting options generally costing more. This study has quantified use of the 4 major devices and cost of the devices and surgery over the study period, 2007-2012, based on the Premier Research Database of hospitalizations in US general acute care, non-federal hospitals. Next, the costs of the two devices of most importance, ceramic-on-polyethylene (CoP) and metal-on-polyethylene (MoP) were used as part of a cost-effectiveness analysis (CEA). The CEA compared lifetime cost and patient utility of the more expensive, but better performing CoP device to the historical standard MoP.

This study has brought two new findings to light regarding use of bearing surfaces. This study has quantified the degree to which clinical problems in ceramic-on-ceramic (CoC) and metal-on-metal (MoM) bearings that began to surface early in the study period resulted in a sharp decline in the use of these devices. CoP took up the lion’s share of lost CoC and MoM market position. Second, use of newer, more expensive devices is strongly associated with younger patients and those with fewer or less severe clinical issues. These trends are consistent with the general cost-effectiveness principle of placing improved, more expensive devices in patients more likely to be able to take advantage of the greater longevity.

The study revealed new information related to implant cost trends across time. Implant costs tended downward across the study period with the largest drops observed in the MoM and CoC surfaces, followed by MoP. CoP implants, whose use grew exponentially as use of MoM and CoC bearings fell, had the lowest drop in costs. These observations demonstrate that the market responded to negative information about devices with problems and to belief that the newest device, CoP, will offer longer serviceable life. Bivariate relationships revealed substantial variation across hospital variables although few of these relationships held through multivariate regression. The research hypothesis that larger, higher surgical volume hospitals would incur lower implant costs did not prove true.

Total surgical costs also trended down across the study period except for surgeries utilizing MoM surfaces, which had a much smaller drop. This may be explained by MoM surfaces having demonstrated a less sharp prevalence drop as patient age increased.

The CEA, which used more complete surgical procedure and implant cost data, improves upon previous studies by introducing post-surgery rehabilitation costs and taking revision rate differences from joint registries; thus, it brings new context to the bearing surface cost-effectiveness discussion. With only a few exceptions, the use of CoP implants was found to be cost-effective as a replacement for MoP. Only the more expensive CoP implant, placed in the oldest patient, at the lower CoP revision improvement failed to meet a $250,000/QALY incremental cost-effectiveness ratio (ICER) threshold. Only 2 additional cases were found to not be cost-effective at a more demanding $150,000/QALY ICER threshold: both at the lower CoP revision rate improvement in an 80 year old patient ($1000 CoP cost delta) and in an 87 year old patient ($620 delta). Information to help physicians and patients frame bearing surface decisions has been provided.

Medicare is the primary payer for one-half of the THA surgeries and has already decided not to pay a premium for more expensive bearing surfaces, despite the fact that the broad age range of cost-effectiveness means that a CoP implant can be justified in many Medicare-served patients. Hospitals bear the cost of purchasing the devices. The Medicare prohibition on balance billing means almost half the subject population is not subject to bearing the incremental cost of a better implant. Thus, it is not clear how the incremental cost for a better device would be borne by the appropriate party.

Although hospitals actively engage in price negotiation with private insurers, one can think insurers are in a strong position to push back from paying more for surgeries using better devices where 10-20 years can be needed for substantial return to be seen on higher cost implants. Given even modest rates of insurance plan turnover, such a delayed pay-back is well beyond reasonable time frames for the avoided or delayed revision cost resulting from using better devices to likely be recovered by the insurance firm that paid the higher price at the time of the index arthroplasty procedure. So, as a practical matter, a pseudo-societal perspective is used in this study to analyze a decision where the entity incurring the greater cost is unlikely to benefit in any way from making the decision to offer more expensive, better devices.

The patient benefits from any increased utility resulting from choosing a hip implant that lasts longer. Current health financing practice in the US means the patient probably will feel little or no increased cost for this greater benefit while the hospital that is likely to shoulder the cost will get no benefit. Reference pricing might be a possible solution. Reference pricing is being tested for overall pricing of services, like hip and knee replacements, to help private payers financially incentivize patients to go to preferred facilities that have a history of low cost and have agreed to a baseline, reference price.[71] If a patient chose to go to a more expensive facility or selected a more expensive medical device among acceptable alternatives, he or she would incur a much larger portion of the total greater cost (usually 100% of the difference), rather than just being responsible for the 20% co-pay portion (for example) of the above reference cost. Generally, this increased patient cost-sharing would not be subject to out-of-pocket maximums that limit cost exposure. This same logic could be applied to allow the patient to more actively participate in the process by agreeing that, for instance, a $620 increased CoP cost share is justified for a chance at a longer lasting new hip joint. The $620 may seem trivial compared to the ~$40,000 total cost (including rehabilitation) of a THA discussed in chapter 6. On the other hand, given out of pocket limits, the patient may not even face the full 20% co-pay of $8,000 and thus, the $620 extra for a chance at improved performance becomes a more relevant consideration.

Continued efforts to control rising health care costs have led to experiments with “bundled payments” in general and in total joint arthroplasty specifically. Bundled pricing in arthroplasty describes a fixed total reimbursement for the “episode-of-care” that includes the fees for all providers: the surgeon, anesthesiologist, the hospital, physical therapists, and other facilities.[72] The belief is that shared responsibility for the total cost of all aspects of an arthroplasty procedure will force providers to become more efficient and improve quality of care as a method to save cost. Lengths of both 30 and 90 days are being considered for the length of episode for which the bundled pricing applies and for which providers will receive no additional payment; minimizing rehabilitation costs and avoiding the need for hospital re-admission during this time period are obvious areas of focus. The considerations of 30- vs. 90-day episode of care lengths have no impact on the decade it takes for performance differences to begin to emerge based on bearing choice and will thus make it even harder to justify incurring extra cost for improved bearings surfaces.

The most important area for future research in total hip arthroplasty, particularly as related to bearing surface choice, is neither finely tuned cost-effective analyses nor resolving payment issues. The most important area is to better quantify the actual revision performance of the MoP and CoP hip implants. This research has used information from two reputable joint registries that place the 20-year cumulative failure rate of the baseline MoP hip implant at 9 and 14.5 revisions/100THAs, respectively, and long term rates for CoP performance are just beginning to be reported. The vast difference in baseline MoP revision rates reported has a profound impact on assessing economic performance of alternate, better surfaces. Ensuring that the accurate clinical performance is considered in answering these questions is most important.

REFERENCES

1. Dentzer S. Comparative Effectiveness: Coherent Health Care At Last? *Health Affairs.* 2010;29(10):1756.

2. Donnelly J, Garber AM, Wilensky GR. Health Policy Brief: Comparative Effectiveness Research. *Health Affairs.* October 8-web exclusive 2010:5.

3. Inpatient Surgery. *Centers for Disease Control and Prevention.* Available at: http://www.cdc.gov/nchs/fastats/inpatient-surgery.htm. Accessed 23 September 2015.

4. Chang RW, Pellissier JM, Hazen GB. A cost-effectiveness analysis of total hip arthroplasty for osteoarthritis of the hip. *JAMA.* 1996;275(11):858-865.

5. Lavernia CJ, Alcerro JC. Quality of Life and Cost-Effectiveness 1 Year After Total Hip Arthroplasty. *The Journal of Arthroplasty.* 2011;26(5):705-709.

6. Daigle ME, Weinstein AM, Katz JN, Losina E. The Cost-effectiveness of Total Joint Arthroplasty: A Systematic Review of Published Literature. *Best Practice & Research Clinical Rheumatology.* 2012;26(5):649-658.

7. Jenkins PJ, Clement ND, Hamilton DF, Gaston P, Patton JT, Howie CR. Predicting the Cost-effectiveness of Total Hip and Knee Replacement: A Health Economic Analysis. *Bone & Joint Journal.* January 1, 2013 2013;95-B(1):115-121.

8. Odum SM, Troyer JL, Kelly MP, Dedini RD, Bozic KJ. A Cost-Utility Analysis Comparing the Cost-Effectiveness of Simultaneous and Staged Bilateral Total Knee Arthroplasty. *The Journal of Bone & Joint Surgery.* 2013;95(16):1441-1449.

9. Updates and Revisions to ICD-9-CM Procedure Codes (Addendum). *Centers for Medicare and Medicaid Service*. Available at: http://www.cms.gov/Medicare/Coding/ICD9ProviderDiagnosticCodes/Downloads/icd9addendafy06.pdf Accessed 5 June 2014, Version 3 (effective October 5, 2005).

10. Caton J, Prudhon J. Over 25 years Survival After Charnley’s Total Hip Arthroplasty. *International Orthopaedics.* 2011;35(2):185-188.

11. Raphael I, Parvizi J, Rothman R. *The Influence of Head Material on Polyethylene Wear*. Heidelberg: Springer; 2013.

12. Fisher J. Tribology of Hip Prostheses. In: Knahr K, ed. *Tribology in Total Hip Arthroplasty*. Heidelberg: EFORT; 2011:3-9.

13. Hannouche D, Zaoui A, Zadegan F, Sedel L, Nizard R. Thirty years of experience with alumina-on-alumina bearings in total hip arthroplasty. *International Orthopaedics.* 2011;35(2):207-213.

14. Triclot P. Metal-on-metal: History, State of the Art (2010). *International Orthopaedics.* 2011 2011;35(2):201-206.

15. Abdel MP, Heyse TJ, Elpers ME, et al. Ceramic Liner Fractures Presenting as Squeaking After Primary Total Hip Arthroplasty. *Journal of Bone & Joint Surgery Am.* 2014;96(1):27-31.

16. Brockett CL, Williams S, Jin Z, Isaac GH, Fisher J. Squeaking Hip Arthroplasties: A Tribological Phenomenon. *The Journal of Arthroplasty.* Jan 2013;28(1):90-97.

17. Kiyama T, Kinsey TL, Mahoney OM. Can Squeaking With Ceramic-On-Ceramic Hip Articulations In Total Hip Arthroplasty Be Avoided? *The Journal of Arthroplasty.* June 2013;28(6):1015-1020.

18. Fehring T, Odum S, Sproul R, Weathersbee J. High Frequency of Adverse Local Tissue Reactions in Asymptomatic Patients With Metal-on-Metal THA. *Clinical Orthopaedics and Related Research.* Feb 2014;472(2):517-522.

19. FDA Safety Communication: Metal-on-Metal Hip Implants. Silver Spring, MD: US Department of Health and Human Services, Food and Drug Administration; 2013.

20. Trucillo M. Hip Replacement Recall. *American Recall Center.* Available at: http://www.recallcenter.com/hip-replacement/recall/. Accessed 8 April 2015.

21. Mendenhall S. 2013 Hip and Knee Implant Review. *Orthopedic Network News.* July 2013;24(3):7-9.

22. Fukui K, Kaneuji A, Sugimori T, Ichiseki T, Matsumoto T. Wear Comparison Between Conventional and Highly Cross-Linked Polyethylene Against a Zirconia Head: A Concise Follow-Up, at an Average 10 Years, of a Previous Report. *The Journal of Arthroplasty.* Oct 2013;28(9):1654-1658.

23. Babovic N, Trousdale RT. Total Hip Arthroplasty Using Highly Cross-Linked Polyethylene in Patients Younger Than 50 Years With Minimum 10-Year Follow-Up. *The Journal of Arthroplasty.* May 2013;28(5):815-817.

24. Cooper J, Della Valle C, Berger R, et al. Corrosion at the Head-Neck Taper as a Cause for Adverse Local Tissue Reactions After Total Hip Arthroplasty. *Journal of Bone & Joint Surgery* Sep 2012;94(18):1655-1661.

25. Scully W, Teeny S. Pseudotumor Associated With Metal-on-Polyethylene Total Hip Arthroplasty. *Orthopedics.* May 2012;36(5):e666-e670.

26. Cash DJW, Khanduja V. The Case for Ceramic-on-polyethylene as the Preferred Bearing for a Young Adult Hip Replacement. *Hip International.* Sep 2014;24(5):421-427.

27. Bozic KJ, Kurtz S, Lau E, et al. The Epidemiology of Bearing Surface Usage in Total Hip Arthroplasty in the United States. *The Journal of Bone & Joint Surgery.* 2009;91(7):1614-1620.

28. Rajaee SS, Trofa D, Matzkin E, Smith E. National Trends in Primary Total Hip Arthroplasty in Extremely Young Patients: A Focus on Bearing Surface Usage. *The Journal of Arthroplasty.* 2012;27(10):1870-1878.

29. Lehil MS, Bozic KJ. Trends in Total Hip Arthroplasty Implant Utilization in the United States. *The Journal of Arthroplasty.* 2014;29(10):1915-1918.

30. Clarke IC, Gustafson A. Clinical and Hip Simulator Comparisons of Ceramic-on-Polyethylene and Metal-on-Polyethylene Wear. *Clinical Orthopaedics and Related Research.* 2000;379:34-40.

31. Hernigou P, Bahrami T. Zirconia and Alumina Ceramics in Comparison With Stainless-steel Heads: Polyethylene Wear After a Minimum Ten-year Follow-up. *Journal of Bone & Joint Surgery, British Volume.* 2003;85-B(4):504-509.

32. Wroblewski BM, Siney PD, Dowson D, Collins SN. Prospective Clinical and Joint Simulator Studies of a New Total Hip Arthroplasty Using Alumina Ceramic Heads and Cross-linked Polyethylene Cups. *Journal of Bone & Joint Surgery, British Volume.* 1996;78-B(2):280-285.

33. Urban JA, Garvin KL, Boese CK, et al. Ceramic-on-Polyethylene Bearing Surfaces in Total Hip Arthroplasty Seventeen to Twenty-one-Year Results. *The Journal of Bone & Joint Surgery.* 2001;83(11):1688-1694.

34. Hallan G, Dybvik E, Furnes O, Havelin LI. Metal-backed Acetabular Components With Conventional Polyethylene: A Review of 9113 Primary Components With a Follow-up of 20 Years. *Journal of Bone & Joint Surgery, British Volume.* 2010;92-B(2):196-201.

35. Ihle M, Mai S, Siebert W. Ceramic Versus Metal Femoral Heads in Combination With Polyethylene Cups: Long-Term Wear Analysis at 20 Years. *Seminars in Arthroplasty.* Dec 2011;22(4):218-224.

36. Meftah M, Klingenstein GG, Yun RJ, Ranawat AS, Ranawat CS. Long-Term Performance of Ceramic and Metal Femoral Heads on Conventional Polyethylene in Young and Active Patients: A Matched-Pair Analysis. *The Journal of Bone & Joint Surgery.* Jul 2013;95(13):1193-1197.

37. Wang S, Zhang S, Zhao Y. A Comparison ofPpolyethylene Wear Between Cobalt-chrome Ball Heads and Alumina Ball Heads After Total Hip Arthroplasty: A 10-year Follow-up. *Journal of Orthopaedic Surgery and Research.* Jul 2013;8(20):1-4.

38. Nakahara I, Nakamura N, Nishii T, Miki H, Sakai T, Sugano N. Minimum Five-Year Follow-Up Wear Measurement of Longevity: Highly Cross-Linked Polyethylene Cup Against Cobalt-Chromium or Zirconia Heads. *The Journal of Arthroplasty.* Dec 2010;25(8):1182-1187.

39. Kawate K, Ohmura T, Kawahara I, Tamai K, Ueha T, Takemura K. Differences in Highly Cross-Linked Polyethylene Wear Between Zirconia and Cobalt-Chromium Femoral Heads in Japanese Patients: A prospective, randomized study. *The Journal of Arthroplasty.* 2009;24(8):1221-1224.

40. Kadar T, Hallan G, Aamodt A, et al. Wear and Migration of highly Cross-linked and conventional Cemented polyethylene Cups with Cobalt chrome or Oxinium Femoral heads: A Randomized radiostereometric Study of 150 Patients. *Journal of Orthopaedic Research.* 2011;29(8):1222-1229.

41. National Joint Registry for England, Wales, and Northern Ireland. Available at: http://www.njrcentre.org.uk/njrcentre/Portals/0/Documents/England/Reports/10th_annual_report/NJR%2010th%20Annual%20Report%202013%20B.pdf. Accessed 15 May 2014.

42. Norwegian Arthroplasty Registry. *Haukeland University Hospital*. Available at: http://nrlweb.ihelse.net/eng/Report_2010.pdf. Accessed 15 May 2014.

43. HealthEast Joint Replacement Registry. Available at: http://www.healtheast.org/orthopaedics/registry.html Accessed 15 May 2014.

44. Bozic KJ, Katz P, Cisternas M, Ono L, Ries MD, Showstack J. Hospital Resource Utilization for Primary and Revision Total Hip Arthroplasty. *The Journal of Bone & Joint Surgery.* 2005;87(3):570-576.

45. Robinson JC, Pozen A, Tseng S, Bozic KJ. Variability in Costs Associated with Total Hip and Knee Replacement Implants. *The Journal of Bone & Joint Surgery.* 2012;94(18):1693-1698.

46. Ravi B, Croxford R, Reichmann WM, Losina E, Katz JN, Hawker GA. The Changing Demographics of Total Joint Arthroplasty Recipients in the United States and Ontario From 2001 to 2007. *Best Practice & Research Clinical Rheumatology.* 2012;26(5):637-647.

47. Kelly MP, Bozic KJ. Cost drivers in Total Hip Arthroplasty: Effects of Procedure Volume and Implant Selling Price. *American Journal of Orthopedics* 2009;38(1):E1-4.

48. Bozic KJ, Morshed S, Silverstein MD, Rubash HE, Kahn JG. Use of cost-effectiveness analysis to evaluate new technologies in orthopaedics: The case of alternative bearing surfaces in total hip arthroplasty. *The Journal of Bone & Joint Surgery.* 2006;88(4):706-714.

49. Gioe T, Sharma A, Tatman P, Mehle S. Do “Premium” Joint Implants Add Value?: Analysis of High Cost Joint Implants in a Community Registry. *Clinical Orthopaedics and Related Research.* 2011;469(1):48-54.

50. Pulikottil-Jacob R, Connock M, Kandala N-B, et al. Cost Effectiveness of Total Hip Arthroplasty in Oosteoarthritis: Comparison of Devices With Differing Bearing Surfaces and Modes of Fixation. *Bone & Joint Journal.* 2015;97-B(4):449-457.

51. Grossman M. On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy.* 1972;80(2):223-255.

52. Grossman M. *The Demand for Health: A Theoretical and Empirical Investigation.* New York: Columbia (for the National Bureau of Economic Research); 1972.

53. Folland S, Goodman AC, Stano M. *The Economics of Health and Health Care.* 6th ed. Saddle River, NJ: Prentice Hall (for Pearson Education, Inc.); 2010.

54. Toossi M. Labor Force Projections to 2024: The Labor Force is Growing, but Slowly. December 2015 ed. Washington, DC: United States Department of Labor, Bureau of Labor Statistics; 2015.

55. Okike K, O’Toole RV, Pollak AN, et al. Survey Finds Few Orthopedic Surgeons Know The Costs Of The Devices They Implant. *Health Affairs.* 2014;33(1):103-109.

56. Chambers JD, Chenoweth M, Thorat T, Neumann PJ. Private Payers Disagree With Medicare Over Medical Device Coverage About Half The Time. *Health Affairs.* August 1 2015;34(8):1376-1382.

57. Simon HA. A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics.* 1955;69(1):99-118.

58. Berry DJ, Harmsen WS, Cabanela ME, Morrey BF. Twenty-five-Year Survivorship of Two Thousand Consecutive Primary Charnley Total Hip Replacements. *The Journal of Bone & Joint Surgery.* 2002;84(2):171-177.

59. Manning WG, Mullahy J. Estimating Log Models: To Transform or Not to Transform? *Journal of Health Economics.* 2001;20(4):461-494.

60. Buntin M, Deb P, Escarce J, Hoverman C, Paddock S, Sood N. *Comparison of Medicare Spending and Outcomes for Beneficiaries with Lower Extremity Joint Replacements.* Santa Monica, CA: Rand Corporation; 2005.

61. Bozic KJ, Chiu VW, Slover JD, Immerman I, Kahn JG. Health State Utility in Patients with Osteoarthritis of the Hip and Total Hip Arthroplasty. *The Journal of Arthroplasty.* 2011;26(6, Supplement):129-132.e122.

62. Katz JN, Phillips CB, Fossel AH, Liang MH. Stability and Responsiveness of Utility Measures. *Medical Care.* 1994;32(2):183-188.

63. Laupacis A, Bourne R, Rorabeck C, et al. The Effect of Elective Total Hip Replacement on Health-related Quality of Life. *The Journal of Bone & Joint Surgery.* 1993;75(11):1619-1626.

64. Rissanen P, Aro S, Sintonen H, Asikainen K, Slätis P, Paavolainen P. Costs and Cost-Effectiveness in Hip and Knee Replacements: A Prospective Study. *International Journal of Technology Assessment in Health Care.* 1997;13(4):574-588.

65. National Joint Registry for England, Wales, Northern Irelan, and the Isle of Man. 1February 2016. Available at: http://www.njrreports.org.uk/Portals/0/PDF downloads/NJR%2012th%20Annual%20Report%202015.pdf.

66. Bell FC, Miller ML. Life Tables for the United States Social Security Area 1900-2100. Vol Actuarial Study No. 120: Office of the Chief Actuary; 2005.

67. Gray AM, Clarke PM, Wolstenholme JL, Wordsworth S. *Applied Methods of Cost-Effective Analysis in Health Care*. Oxford: Oxford University Press; 2011.

68. Ramsey SD, Willke RJ, Glick H, et al. Cost-Effectiveness Analysis Alongside Clinical Trials II—An ISPOR Good Research Practices Task Force Report. *Value in Health.* 2015;18(2):161-172.

69. Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the panel on cost-effectiveness in health and medicine. *JAMA.* 1996;276(15):1253-1258.

70. Braithwaite RS, Meltzer DO, King JTJ, Leslie D, Roberts MS. What Does the Value of Modern Medicine Say About the $50,000 per Quality-Adjusted Life-Year Decision Rule? *Medical Care.* 2008;46(4):349-356.

71. Robinson JC, MacPherson K. Payers Test Reference Pricing and Centers of Excellence to Steer Patients to Low-price and High-quality Providers. *Health Affairs.* 2012;31(9):2028-2036.

72. Bozic KJ, Ward L, Vail TP, Maze M. Bundled Payments in Total Joint Arthroplasty: Targeting Opportunities for Quality Improvement and Cost Reduction. *Clinical Orthopaedics and Related Research®.* 2013;472(1):188-193.

APPENDIX A: DETERMINATION OF IMPLANT AND TOTAL HOSPITALIZATION COSTS

The PRD contains line item charges, each of which is noted by a hospital-provided charge description (HCD). In consultation with the hospitals, Premier developed a standardized charge description (SCD) which groups similar HCD items into SCD values that create taxonomy where like HCD items are collected into appropriately described SCD values. These 2 steps were used in tandem to identify the charge items attributable to hip implant hardware and to segregate items not related to research interests. My categorization here was conducted on the charge items as received in the PRD, with the appropriate ICD-9-CM procedure codes as already detailed in the methods applied to establish inclusion in the study and to determine which bearing surface was used.

1. The SAS Prxmatch procedure was used to search the HCD variable for presence of any of the following list of hip arthroplasty related terms: ortho, hip, implant, implnt, liner, poly, acetab, femoral, femur, head, stem, shell, cup, stryker, depuy, insert, zimmer, stem, ball, biomet, sleeve, smith, nephew, s&n, neck, biomet, wright, exactech, encore, spacer, extension, austin moore, bipolar, cobalt, chrome, ceramic, steel, metal, biolox, endo, modular, articulation, alumina, zirconia.
2. Review of the generated list of HCD values included many appearances of search terms in descriptions that are unrelated to the THA procedure. Detailed scrutiny of the list of items revealed that the word “implant” appeared in the SCD value for HCD items that were implant related although this list also included many items that were not THA related.
3. The SCD values were then searched for the word “implant.” This considerably shortened the list of items with the vast majority of items related to THA hardware, but other body parts were represented as well. This observation was interpreted to mean that Premier used the term “implant” in the SCD to characterize devices implanted somewhere in the body. The SCD items containing the word “implant” were treated as follows:

   a. SCD values that included specific hip arthroplasty terms (brand names, specific hip/upper leg parts) were noted to be hip implant cost items and characterized as “hi.”
   b. SCD values that included “implant” and body parts other than hips, with the exception of cardiac-related terms, were characterized as “no” and excluded from the study.
   c. Since one cannot tell if a cardiac procedure was planned or arose due to a THA complication, these items were kept in the study, but put into the non-implant hospitalization cost category, characterized as “h.”

4. Two SCD descriptions, “implant ortho” and “implant misc,” were too general to make a decision on inclusion. Since these accounted for over 25% of all charge items with “implant” in the SCD, simply excluding them would not have been desirable. Thus, the HCD values for these items were reviewed individually as described in step 1 above for inclusion in the “hi” cost designation.
5. Noting the presence of multiple other body parts in the SCD values that included the word “implant,” the “implant ortho,” and “implant misc,” the HCD values were searched for a list of non-hip body parts that were observed while reviewing the data. These follow: knee, patel, spin, mitek, vascular, should, verte, heart, tibia, rotat, aort, introaoc, tendon, elb, mamm, eye, ear, hernia, lens, finger, toe, tutoplast, valve, vein, write, ulnar, incus, humeral, trap, aneurism, orbit, scleral, tube, cochl, vasc, ankl. Items including any of these terms were reviewed and noted for exclusion by characterizing them as “no.”
6. HCD values for the remaining items from “implant ortho” and “implant misc” were then searched for the list of hip terms in step 1. By now, the remaining items were heavily concentrated in hip related items and thus review of these items produced a large volume of items characterized as “hi.” Manual review of the remaining items resulted in each either being characterized as “h” or “no” depending on the judgment of the author.
7. All remaining cost items not characterized as “hi” or “no” were characterized as “h.” Hip implant costs for each discharge were determined by summing all hi items for each discharge. Total hospitalization costs were determined by summing all h and hi items for each discharge.

APPENDIX B: TOTAL HOSPITALIZATION COSTS BY SURFACE AND YEAR

Year	MoP	MoM	CoC	CoP
2007	18228, 5471; 17202	19452,* 5525; 18662	19376,* 5301; 18344	18767,* 5520; 17443
2008	17856, 5530; 16533	19067,* 5738; 18447	18304, 4493; 17774	18330,* 5015; 17437
2009	18163, 5357; 17151	19293,* 5413; 18836	18467, 4524; 18085	18619,* 5545; 17476
2010	17812, 5292; 17011	19248,* 5457; 18653	18579,* 4759; 17703	18288,* 5222; 17331
2011	17722, 5578; 16671	18824,* 6187; 17483	18057,* 5388; 16694	18383,* 5861; 17115
2012	17330, 5159; 16558	18869,* 6782; 17218	17425,* 5413; 16321	17699,* 5514; 16747

Values reported are mean, standard deviation; median.
*Indicates significantly different from MoP cost in same year at $p<0.01$.
2015$

APPENDIX C: IMPLANT COSTS BY SURFACE AND YEAR

Year	MoP	MoM	CoC	CoP
2007	8048, 3710; 7252	9029,* 3754; 8435	9013,* 3753; 8388	8223, 3679; 7245
2008	7702, 3630; 6927	8621,* 3646; 8248	8275,* 3288; 7435	7586, 3282; 6702
2009	7604, 3305; 7034	8766,* 3789; 8095	7696, 2724; 7098	7931,* 3433; 6924
2010	7553, 3242; 6995	8683,* 3912; 7795	8593,* 3208; 7730	7903,* 3422; 7003
2011	7330, 3333; 6526	7958,* 3874; 7180	7840,* 3498; 6894	8206,* 4080; 7037
2012	6847, 3130; 6080	7498,* 3915; 6516	6940,* 3222; 6628	7465,* 3657; 6360

Values reported are mean, standard deviation; median.

*Indicates significantly different from MoP cost in same year at $p<0.01$.

2015$

APPENDIX D: IMPLANT COSTS BY SURFACE AND HOSPITAL TYPE

Hospital Type	MoP	MoM	CoC	CoP
Urban nonteaching	8293, 3685; 7622	9238, 4018; 8832	9007, 3562; 8606	8183, 3888; 7069
Urban teaching	6679,* 2823; 6045	7950,* 3322; 7260	7403,* 3006; 6886	7418,* 3386; 6476
Rural	7151,* 3441; 6569	7965,* 4063; 7377	7794,* 3905; 6918	8702,* 3944; 8104

Values reported are mean, std deviation; median.
*Indicates significantly different from Urban nonteaching in each surface column at $p<0.01$.
2015$

APPENDIX E: IMPLANT COST BY SURFACE AND HOSPITAL SIZE

Hospital Size	MoP	MoM	CoC	CoP
Large	7845, 3486; 7291	8883, 3955; 8200	8737, 3745; 8063	8177, 3793; 7153
Medium	5914,* 2632; 5336	7848,* 3246; 7707	7126,* 3172; 6210	7062,* 3408; 5974
Small	6196,* 2233; 5617	7090,* 2657; 6575	6988,* 1468; 6894	6450,* 2392; 6035

Values reported are mean, std deviation; median.
*Indicates significantly different from Large in each surface column at $p<0.01$.
2015$

APPENDIX F: IMPLANT COST BY SURFACE AND CENSUS REGION

US census region	MoP	MoM	CoC	CoP
South	7708, 3364; 7200	8861, 3947; 8311	8994, 3639; 8918	7997, 3542; 7049
Midwest	8057,* 3444; 7421	8542,* 3587; 7863	8514, 3434; 8055	8509,* 3606; 7650
Northeast	6556,* 3126; 5785	7459,* 3685; 6489	7010,* 2693; 6807	6706,* 3541; 5800
West	8019,* 3442; 7761	8925, 3709; 8550	8912, 3635; 8481	8704,* 3711; 9003

Values reported are mean, std deviation; median.

*Indicates significantly different from South in each surface grouping at $p<0.01$.

2015$