Evaluation of Total Hip Arthroplasty Research in the US and Europe

by

Marcy J Simmons

A THESIS submitted to Oregon State University

in partial fulfillment of the requirements for the degree of Bachelor of Arts in International Studies in Mechanical Engineering

Presented May 25, 2007
Commencement June 2007
ABSTRACT
OF THE THESIS OF


Abstract approved:

________________________________________________
Brian Bay

Background. Hip replacements are considered routine operations which improve the quality of life in patients by reducing pain and restoring function. There are many variations in prostheses with varying benefits and disadvantages.

Methods. Literature from the United States and Europe was collected on total hip replacements between 2001 and the present and analyzed based on outcome measures and survivorship.

Results. Studies generally use revision as end point with few studies measuring patient satisfaction. The Charnley style metal-on-polyethylene implants still show the most consistent results long term. Metal on metal and ceramic on ceramic show reduced wear and osteolysis, the main indication for revision, so may prove better when more long term studies become available. There are so many implants that long term studies are not actually feasible due to options entering and leaving the market on a rapid basis.

Conclusions. Large-scale, long term, independent studies are needed to compare available choices. Immediate feedback is also necessary to keep up with new products.

The creation of a national registry in the US and standardized outcome measurements would allow large scale conclusions to be made on the benefits and risks of each implant design and approach, giving doctors, engineers, and patients an invaluable source of information for use in decision making.
Bachelor of Arts in International Studies in Mechanical Engineering
Thesis of Marcy J Simmons
Presented on May 25, 2007

Approved:

Brian Bay, Thesis Advisor (Mechanical Engineering)

Belinda Batten, Head or Chair (Mechanical Engineering)

Dr. Joseph G. Hoff, Academic Coordinator, International Degree Program

I understand that my thesis will become part of the collection of Oregon State University. My signature below authorizes release of my thesis to any reader upon request. I also affirm that the work represented in this thesis is my own work.

Marcy J Simmons, Author
Acknowledgements Page

I’d like to acknowledge the assistance of my major advisor, Dr. Brian Bay, as well as the help of Ms. Renee Stowell.
Table of Contents

ABSTRACT ................................................................. 2
Acknowledgements Page ............................................ 4
List of Figures ........................................................... 6
Background ............................................................... 7
Overview of Choices .................................................. 8
  Prostheses ............................................................. 8
  Surgery ................................................................. 8
  Post-operative ...................................................... 9
  Difficulties ........................................................... 9
  Goals ................................................................. 10
Definitions of Common Terms ......................................... 12
Abbreviations .......................................................... 13
Methods ................................................................. 13
Evaluation of Materials and Designs ................................. 13
  Metal on Polyethylene .............................................. 13
  Metal on Metal ...................................................... 17
  Ceramic on Ceramic .............................................. 20
Rehabilitation ............................................................ 22
  Patient Based Outcome Measures ................................ 25
Comments on Cultural Aspects of Joint Registries ............... 30
Appendix – Descriptions of Common Scoring Systems ........... 32
  Harris Hip Score .................................................... 32
  WOMAC ............................................................... 32
  SF-36 ................................................................. 32
  Oxford Hip Score ................................................ 32
Works Cited ............................................................. 33
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1:</td>
<td>Implant compared with natural hip bones. From Brown.</td>
<td>7</td>
</tr>
<tr>
<td>Figure 2:</td>
<td>Highly cross-linked polyethylene cup. Copyright Zimmer.</td>
<td>13</td>
</tr>
<tr>
<td>Figure 3:</td>
<td>Gross Loosening in Harvard Implant after 7 years (left) and well-fixed Charnley implant in same patient (right). From Datir and Wynn-Jones.</td>
<td>15</td>
</tr>
<tr>
<td>Figure 4:</td>
<td>Metal on metal hip prosthesis. Copyright Zimmer.</td>
<td>16</td>
</tr>
<tr>
<td>Figure 5:</td>
<td>TEM micrographs</td>
<td>17</td>
</tr>
<tr>
<td>Figure 6:</td>
<td>Survival curves based on revision for any reason of cup and stem</td>
<td>20</td>
</tr>
<tr>
<td>Figure 7:</td>
<td>Comparison of Cementless implants with Cemented Palacos with Gentamycin showing better survival in the latter at 8.5 years</td>
<td>29</td>
</tr>
</tbody>
</table>
Background

Because of the vast improvement in quality of life, the total hip replacement is regarded as the most successful and most cost effective procedure available. Over 234,000 hip replacements were performed in the US in 2004 (Kozak 40), but with popularity and success have come an abundance of new surgical procedures, new bearing surfaces, and new prostheses.

A hip joint is a ball-and-socket type joint. The cup, part of the pelvis, is known as the acetabulum. The ball is the femoral head, a large ball on the end of the femur. Hips receive forces of up to 8 times one’s body weight during actions such as running or walking up stairs because of their distance from the center line of the body.

Early hip implants were largely failed experiments until the 1960’s. That’s when Dr. John Charnley realized that a successful hip implant would have a minimum of friction. Therefore he coupled a polyethylene cup with a polished femoral head (Toledo-Pereyra 299). He also made the head as small as possible to further reduce the friction. A small head may reduce wear, but it means less stability and a smaller range of motion because if the ball comes to the edge of the cup, it can come out of the socket – an occurrence known as impingement. If it pops out fully, that is known as dislocation and frequent dislocation will require revision surgery. Reduction in wear means larger heads can be used so the patient will be allowed more activities and be less likely to experience dislocation.
Any surgery done on the hip after the initial surgery is known as revision. Revision surgeries have much higher rates of complications than the initial surgery and they are necessary to treat frequent dislocations, infections, thigh pain, loosening of the implant, and fractures. A big difficulty with revision surgery is that there is often less bone to work with the second time around, especially in the case of loosening. Loosening is often caused by the body’s immune system reaction to foreign particles, either polyethylene or cement. In attacking the particles, the bone is also absorbed – causing loosening and leaving a larger cavity for the revision surgery. This loss of bone is called osteolysis.

Bone loss also occurs as a result of stress shielding. Bone adapts itself to different conditions, increasing in density if more stress is placed on it and decreasing in density if it experiences less stress. If the implant is stiffer than the bone, it supports more than its share of forces on the hip then bone loss happens through resorption. This is known as stress shielding. Femoral stems today are made out of a material that has similar rigidity with natural bone to eliminate most stress shielding.

Many solutions for other revision issues have been proposed and tested, but it is difficult to make conclusions based on the small number of studies done and their varying methodologies.

**Overview of Choices**

Doctors and patients have a great deal of options and only a small amount of evidence on which to base their decisions. Most prostheses offer similar success rates over the short term and there are new options available every day.

**Prostheses**
The number one joint replacement couple in the world is metal on polyethylene. Other choices include ceramic on cross-linked or high density polyethylene, metal-on-metal, and ceramic-on-ceramic. Each of these prostheses can be cemented, uncemented, or hybrid – a hybrid implant is one in which one component is cemented and the other uncemented. Hybrid implants usually feature cemented cups and uncemented stems. There are choices about which type of cement to use and also different techniques for packing the cement.

**Surgery**

On the surgical side of things, a relatively new ‘minimally invasive’ surgical technique has been proposed in addition to the posterior, anterior, and lateral approaches already available. It promises shorter recovery time, less blood loss, and possible quicker return to activity, but studies have shown that only very experienced doctors see these benefits while the risk of complications is markedly increased.

Systemic antibiotics are also used by many doctors, but it has not been conclusively shown that they reduce the risk of infection.

**Post-operative**

Standard rehabilitation procedures include exercises for muscular strength, range of motion, aquatics, and walking. Other proposed postoperative care protocols involve intensive upper body strength training, proscribed limited weight bearing, use of abduction pillows, elevated toilets and chairs, restrictions from driving, and restrictions from side-sleeping.

**Difficulties**

The problems that still remain with artificial hip joints are the same problems faced when Charnley was performing implants in the 60’s: wear, impingement, osteolysis, dislocation, thigh pain, aseptic loosening, infection, and abnormal gait or limp. A new problem being faced is that of metallosis which comes from metal-on-metal wear causing high serum levels of
cobalt or chrome in the blood. In the nearly fifty years of artificial hip replacements, we have learned very little about how to treat or prevent these difficulties, despite thousands of studies and papers on the subject. That is because there is no organized or efficient way to compare different components or techniques.

Goals

In October of 2001, the Agence Nationale d’Accréditation et d’Evaluation en Santé (ANAES) published a report detailing current available information on the choice of implant and choice of surgical techniques for implantation. The analysis covered clinical literature from 1995 to 2000 as well as economic literature from the same time period. In total, the group of eleven professionals analyzed 151 papers in French and English.

They found that the methodology in the literature was ‘consistently inadequate’ in that comparisons were made in simple cohorts, end-points and goals were not well defined, and patient satisfaction mattered less in the evaluation than radiographic measurements, despite patient discontent being the impetus for the replacement.

They suggest that studies should include possible preoperative tests that would help doctors decide on a type of prosthesis, a procedure, and a material combination based on characteristics of the patients. Most doctors choose a prosthesis and procedure based on what they are taught in their hospitals. The report authors also recommend comparative research across facilities, with independent data collection, incorporating economic studies. The end-point should include patient’s quality of life using validated scoring systems, especially for cementless and hybrid implants, different material combinations, and patients under 50.

It is very important to collect enough data on the patients, implants, and conditions to control important factors when making comparisons. Comparing one study done on the elderly with another study that has a broad
age range does not yield valid results because young people are more active and demand more of their hips.

Another study written in 2006 for the Journal of Bone and Joint Surgery (Huo) attempts to make suggestions on prosthesis type, surgery approach, surface replacement, and practice management based on current papers, but there is still no study doing a large-scale comparison between prostheses or surgeries, no independent data collection, and the end-points still are based on revision or radiographic data rather than patient satisfaction. There are also very few long term studies or studies involving younger patients.

From the engineering point-of-view, the modern hip implant is considered an exceptional achievement but not a perfect success (Brown 29). It has improved enough in the last thirty years that doctors are now encouraging younger people to go ahead get hip replacements instead of putting them off as long as possible (Brown 32). This is now skewing the data on newer implants since Charnely originally performed total hip replacements only in older patients in very good health. The new and older patient populations are dissimilar so comparisons are difficult to draw. Advances in hip replacement technology have allowed arthritis sufferers to get treatment sooner and made revisions easier, but it is difficult to sort out which changes have been beneficial and how patients and doctors can make informed choices.

Despite advancements in engineering, the revision rate of hip prostheses has remained steady or increased over the years (in Maloney 1582). Randomized, comparative, clinical trials of implants are few, and non-existent with large-scale cohorts. Prospective trials are impractical because of the large number of new implants released each year, and retrospective studies are not timely enough, contain investigator bias, and are usually performed by experts so do not reflect average clinical practice.
Definitions of Common Terms

abduction – movement which separates the leg further from the center-line of the body

acetabulum – The cup-shaped cavity at the base of the hipbone into which the ball-shaped head of the femur fits forming the ball-and-socket joint of the hip

adduction – movement which brings the leg closer to the center-line of the body

aseptic loosening – loosening in the absence of clinical or microbiological evidence of infection

femoral head – the ball shape on the end of the femur that along with the acetabulum completes the ball-and-socket joint of the hip

Harris hip score – a standardized questionnaire that allows physicians and other caregivers to quickly rate the status of a patient following hip replacement

impingement – when the femoral head goes outside the limits of the acetabular cup, sometimes resulting in a dislocation

metallosis – toxicity related metal particles in the blood

osteolysis – degeneration of bone, often caused by the immune system response to wear particles

Oxford Hip score – a standardized questionnaire evaluating pain and function in a hip

stress shielding – resorption of bone caused by implant taking too much of the stress in the bone

subsidence – settling in or lowering

tribology – a study of friction, lubrication, and wear of interacting surfaces in relative motion

trochanter – the bony protrusions on the upper extremity of the femur to which muscles are attached
WOMAC score - Western Ontario MacMaster score that assesses lower extremity pain and function in patients with osteoarthritis of the knee or hip. A lower score is better.

**Abbreviations**

MOM – Metal on Metal  
MOP – Metal on Polyethylene  
COP – Ceramic on Polyethylene  
COC – Ceramic on Ceramic

**Methods**

I collected and analyzed documents relating to prosthesis design, surgical technique, and rehabilitation programs published between 2001 and the present with the studies taking place either in the United States or Europe. I focused on the most recent articles and gave preference to articles with longer than average studies.

**Evaluation of Materials and Designs**

**Metal on Polyethylene**

The most common choice for total hip replacement is a metal femoral head with a polyethylene socket. This metal on polyethylene (MOP) combination has very low friction, shock absorption, and is also inexpensive. To minimize friction, the femoral head is often very much smaller than that made by nature. The average size was 28mm before cross-linked polyethylene came into use; the average today is near 36mm compared to over 50mm for an average skeleton. This couple has high success rates but, long term, still experiences problems related to osteolysis caused by the polyethylene wear particles.
Callaghan et al (2004) looked at the thirty year survival rates of cemented Charnley total hip arthroplasty performed between 1970 and 1972 in 262 patients (330 hips). At the 30 year follow-up time, twenty-seven patients (34 hips) were available for clinical evaluation. The original implant was still in place and functioning at the time of death in 290 hips of the original 330, meaning a cumulative survival rate of 88%. For the hips in living patients thirty years after surgery, the survival rate was 68% with revision for any reason as the end point. The author collected WOMAC scores along with radiographs for patients available for follow-up and found probable or definite aseptic loosening in fourteen acetabular cups and in six femoral stems. The WOMAC scores had increased from normative values which the author supplies but he does not indicate at what time period they were established. They were still low, with only one patient reporting pain that required a cane for long walks. WOMAC scores are included, but unfortunately subjective patient satisfaction measurements are missing.

Dorr et al (2005) investigated the advantage of a Durasul highly cross-linked polyethylene acetabular liner over the traditional liner after five years. They found that after the bedding-in period, the wear rate of the cross-linked liner was 45% that of the conventional liner. They feel this warrants increased use of the cross-linked polyethylene liner. This study examined wear exclusively – creating its data pool from patients who had not had a revision or cup recall at the time of the study. They found no difference
between the Harris Hip scores of the two groups, and similar bedding-in penetration.

Their advocacy of continued use of this liner is based exclusively on linear wear rates after the bedding-in period in a group of thirty-two test hips and thirty-five control hips. It also would have been useful to compare the revision rates between the two populations. Both the size and the outcome measures of this study limit its applicability.

Buckwalter et al (2006) reported on the twenty-five year results of Charnley implants with improved cementing techniques compared to traditional hand-packing of cement. The new method involved the use of a cement gun and a distal cement plug. The original implant was still functioning after 25 years, or was in place at time of death in 88.2% of the original 357 hips. Of the forty-nine patients (fifty-two hips) available for the study, eight were seen to be loose radiographically and five of these required revision. This lowers the 25 year survival rate to 80% with or without the improved cementing technique. Of patients alive after 25 years, 77% had their original prosthesis.

The authors give detailed numbers relating to revisions and the reasons for them which will be useful when cementless prostheses are available for comparison after 25 years. The study found no statistically significant advantage to the improved cement packing technique over traditional hand packing, but they continue to believe that the contemporary

Figure 3: Gross Loosening in Harvard Implant after 7 years (left) and well-fixed Charnley implant in same patient (right)
technique provides more reliable cement filling and pressurization.

Grant and Nordsletten (2004) report on the results of 116 lord prosthesis total hip replacements conducted between 1981 and 1985. They were able to report on the final status or status at the time of the study on all 116 hips thanks in large part to the Norwegian Arthroplasty Register (2637). Seventy-one patients were still alive at the time of follow-up, although one patient refused to participate he told the phone interviewer that his hip was still in place and functioning well. The rate of survival for the femoral component, with revision for any reason or mechanical loosening as the end point, was 92% at 17.5 years. The rate of survival of the acetabular cup at 17.5 years with revision for any reason or radiographic loosening as the end point was 64%, but if only revision as a result of mechanical or radiographic loosening is considered, the rate is 86% at 15 years, 65% at 17.5 years, and 42% at 19 years.

Thus, the fixation of the acetabular cups, which were cementless and threaded, was very poor. The authors noted that many of the patients had substantial osteolytic destruction, even though they had no symptoms of loosening, and therefore need revision surgery before more destruction can be caused. So the beaded Lord femoral stem showed excellent fixation, but the threaded cup had worse than poor results. Since neither cup nor stem is currently available, these results have little value for current decision-making, but add to the bank of knowledge concerning implant fixation and threaded cups. If all threaded cups could be compared to all cementless unthreaded cups, we could see if this is a common problem.

Datir and Wynn-Jones (2005) reported on rough versus smooth surface femoral stems. Their study had a useful technique because they used patients undergoing bilateral replacement and implanted one rough and one smooth surfaced stem in each patient. This helps alleviate errors due to patient factors, although there may still be differences between left and right hips.
The literature has reported good results with either rough or smooth surfaced stems, and in 1988 and 1993, studies showed increased rates of aseptic loosening when a stem changed surface from smooth to matte. The operation was performed in 51 patients, but 3 were excluded from the study.

In this study, the smooth surfaced component performed significantly better after 10 years than the matte surface component (92% versus 70% measuring aseptic loosening with or without revision surgery).

This study attempted to show the difference in survivorship between rough and smooth surfaced femoral stems of almost similar design, but unfortunately the geometry was not exactly the same between the two stems and the authors noted many previous studies with conflicting results. They conclude that the surface finish isn’t important as long as it doesn’t interfere with the bond between the cement and the stem. Despite the obvious advantage shown in the group analyzed, in the face of other literature on the same subject, a conclusion cannot be drawn on the value of a rough versus a smooth surface.

**Metal on Metal**

Recently metal-on-metal (MOM) and ceramic-on-ceramic (COC) implants have gained popularity as a solution to the wear problem. With highly polished surfaces and very tight clearances between components, metal-on-metal couples promise better results for younger, active patients, although at a much higher cost. Metal-on-metal implants also present the problem of metallic wear particles which could potentially cause cancer or other problems.
Cobb et al (2006) investigated whether there was any clinical significance to metal ions released from cobalt-chromium implants by comparing and analyzing 81 different clinical studies. They noted that cobalt-chromium ions are elevated in patients with traditional ceramic-on-polyethylene implants, and even further elevated when the implant is metal-on-metal. The first challenge with determining clinical significance is measuring the amount of metal in the body. Measurements are easily contaminated and can vary widely. Different measures are also used by different researchers, making comparison between them even more difficult. Another issue is that the most common method of measurement, using serum, only measures the amount of metal freely circulating in the system and does not give an estimate of bound ions, which may be in the lymph nodes and therefore more significant. The measurements also vary patient to patient, increase over time, and it is not definitely known whether they vary by activity level of the patient.

Catelas et al (2006) compared particle characteristics in patients with older MOM implants to patients with newer MOM implants. They isolated particles from tissues near the implant site at three different times in the life of the implant. The tissues were collected during revision surgeries. This casts doubt on the validity of the measurements that came from short term implants. The need for revision in short term implants means they would have greatly different wear characteristics than well-functioning implants. The authors found that different implant styles or

Figure 5: TEM micrographs of isolated particles from each group. From top left, 1, 2, 3, and 4.
generations had no significant impact on the quantity or type of wear particles, whereas implant length had the greatest effect. The types of particles found suggest that the type of particles created changes over time, although the immune system response would be the same for any stage, it does present an interesting puzzle on what wear might occur and why in these different stages. More information gathering is definitely warranted.

They gave very great detail on their measurement methods, which they believed to be the most accurate available based on current research, so the measurements they took can be easily compared with future studies. They also gave a good discussion on the limitations of their study. The discussion on measurement technique I found to be very relevant because someone wanting to replicate their study or use it for comparison will be able to decide if they want to use the same methods or not and the reasoning behind it is clearly laid out.

Most evidence seems to suggest that the highest levels of cobalt and chromium are found during the settling in period of the implant, and then reach a lower, steady-state level for the life of the implant (in Cobb 387). The life of the implant also exceeds the life of the patient, with a lower risk of osteolysis than polyethylene combinations. The risks associated with the presence of metallic ions in the body are still much lower than the risks associated with revision surgery. Continued studies comparing the incidence of cancer and other maladies in patients with MOM implants compared with the population are necessary to fully quantify the risks associated with ion levels in the body. The levels can also be much higher if the implant is misaligned during surgery resulting in increased wear. Further developments in accurate implantation along with precision manufacturing methods may further decrease these risks.

Both studies investigating the clinical significance of wear particles and their relation to implant design or length were similarly thorough in
taking current research into account. If all of the data was centrally pooled, more conclusions could be drawn.

Ceramic on Ceramic

Ceramic-on-ceramic implants offer similar promises on long-term wear rates and osteolysis incidence, without concern of metallosis. “Alumina is very hard, highly oxidized, wettable, and very resistant” (Sedel 22). Medical grade alumina ceramic is a polycrystalline material and has been produced with high quality and tight standards since the 1980s (Sedel 22, Hamadouche 75). In laboratory settings, ceramic materials outperform metal by a factor of 4. In clinical conditions, researchers have seen at best factor of 2 improvements on development of wear particles. A recent study has shown the concentration of wear particles in periprosthetic tissue following a COC implant to be “two to twenty-two times lower than MOP” after ten years (in Hamadouche 75). Also, osteolysis is very seldom reported, with none found in the study performed by Hamadouche.

The design parameters that have the greatest significance on performance are sphericity and clearance. Like MOM implants, ceramics experience bedding-
in, where the wear rate is higher than in the eventual steady state period. However, even during the bedding-in period, wear rates for ceramics are about a thousand times lower (22). Improper alignment or unusual muscle contractions can result in abnormal wear characteristics and the wear rate can be increased by high factor, but will still be less than other options.

A major disadvantage to ceramic implants is their risk of fracture. Recent materials (since 1986) have fracture rates lower than 1/100 000. The author compares this to a 10 percent failure rate at 10 years in MOP prostheses, but a fractured implant is a much more grave concern than a loosening one (Sedel 24). Since 2001, over 1000 ceramic hips have been implanted with no incidence of fracture (Sedel 25). Hamadouche implanted 106 patients with 118 COC hips and there was no incidence of fracture in either the head or socket (75).

Another issue is that of ceramic fixation. Osteolysis may be avoided with the ceramic fixture, but if proper bone grow-in does not occur, the cement particles can lead to acetabular loosening. A recent study of patients under the age of 50 with press-fit sockets showed a 93 percent survival rate at 14 years with revision for any reason as the end point (Sedel 24). The current series performed by Hamadouche had a relatively low survival rate at twenty-years, especially compared to the traditional MOP Charnley hip, but 92% of the revisions were the result of aseptic loosening of the acetabular cup caused by acute debonding in cemented implants and by progressive migration in cementless (76). Obviously socket fixation needs to be improved before COC implants can be considered superior to traditional Charnley implants. The author suggests a titanium shell with alumina insert or bioactive material coating, either of which would enhance primary fixation (Hamadouche 76).

One bioactive material coating is hydroxyapatite (HA), which is a synthetic form of the crystalline portion of natural bone mineral. Goosen et
al (2005) studied the results of HA-coated femoral stems after a minimum of 6 years in 100 patients (106 hips). The implants were metal-on-polyethylene and the authors compared them to porous coated implants with similar geometry. Harris hip scores and radiographs were collected at 6 weeks, 3 months, 6 months, 1 year, and every year thereafter. At the time of follow-up, no femoral stems had been replaced and just one acetabular cup.

When compared with the porous coated stems, HA-coated had higher Harris hip scores and pain subscores, as well as a 48% increase in bone mineral density. The author did a very thorough job of compiling and comparing other 5-year studies of porous or HA-coated stems to increase the extent of validity, and there was no obvious advantage to HA-coated stems, despite the excellent survival rate in the current study. An even broader range comparison would be necessary to determine advantage either way.

Rehabilitation

Maire et al (2006) report on the results of a pilot 6 week arm exercise program that resulted in significantly better gait, walking distance in a timed test, and WOMAC scores. They theorize that normal rehabilitation programs including a low level of daily physical activity lead to decreased exercise capacity and physical condition. Activity is necessarily limited by the healing hip, but upper body exercises to improve cardio-respiratory fitness were ‘clinically well-tolerated’ and impacted not only the test outcome, but also the patients’ perception of exercise intensity and exertion. The authors speculate that the results can be attributed mostly to the improved cardio-respiratory fitness but speculate that there may have been some transfer effect where upper body exercises result in increased muscle mass in the lower body and vice versa. They also point out the possibility of psychological effects of strength training and achievement. This study, while the sample size is small, was completed in a randomized
fashion with independent, blinded outcome measurements and the study authors discuss at length the study limitations. The measurements were distance walked in a 6 minute test and WOMAC scores, both clearly defined and repeatable, allowing clear result comparison for other rehabilitation programs. The length of follow up was just one year, as the rehabilitation program addresses return to normal function.

Nehme et al (2003) address whether or not hip replacement affects the bone mineral density in the spine and contralateral hip. The authors measured bone mineral density in the contralateral hip and lumbar spine 1 month before surgery, and several times postoperatively, extending out to 2 years. The patients were treated with cemented modular straight stem implants which allow immediate post-operative weight bearing. The measurements found no bone loss in either the contralateral hip or the spine. They conclude that immediate full weight bearing avoids bone loss in elderly patients at risk for osteoporosis. Their study size was 52 male patients and they took measurements at multiple sites to confirm density.

The drawback to their study is that there was no control group to evaluate their conclusion. Based on their test they can conclude that bone mineral doesn’t significantly decrease after total hip replacement. Their study can be repeated in order to compare the bone density results to a population that is not allowed immediate full weight bearing, but would be more useful if other factors, such as patient pain and functionality were also measured. The two-year follow up length also seems needless, because if the patients haven’t experience bone loss during what would normally be the prescribed period of bed rest (6-weeks or less); further measurements report nothing of value.

Tveit and Karrholm (2001) and Peak et al (2005) both report on effectiveness and results of patient restrictions following total hip replacement. Tveit and Karrholm found that when patients were asked to use
crutches to support 30% of their body weight while navigating five different types of terrain, not a single patient managed to comply with the directive after specific and individual training and under continuous direct supervision. This study was done using a sample size of 15 and measurements were made using registering pressure-sensitive insoles.

This study indicates that prescribed patient restrictions such as this are probably ineffective, especially since five of the patients incorrectly thought they had complied with the directive. This study is easily repeatable, and so could be used to compare with different patient restrictions.

A useful follow-up would be to determine if early post-operative weight bearing has an effect on the implant fixation or on later loosening. The secondary goal of the study was to test a new type of load measuring insole that could be used in the patient’s home environment to collect data away from the laboratory setting to accomplish this analysis.

Peak et al tested a specific and well defined hypothesis that patients not placed on functional restrictions following total hip arthroplasty would be more likely to experience dislocations following surgery. Their sample size of 256 was randomly placed into either the restricted or unrestricted group and the surgeon was blinded as to the group until after wound closure to avoid any operative biases. Patients in the unrestricted group were allowed to use restrictive equipment and aids if desired and all patients kept logs of their compliance. Significant differences were noted in terms of compliance with range-of-motion restrictions and use of equipment except for side sleeping pillows after rehabilitation stays.

Many patients in both groups complied with the range of motion restrictions beyond the six weeks required. Patient satisfaction, and return to functionality were significantly better in the unrestricted group and there were no dislocations. The authors determined that additional
restrictions are not only unnecessary, but result in increased cost in terms of equipment and delay in return to work.

The outcome measurements were prevalence of dislocation, prevalence of limp, patient compliance with limited range-of-motion 6 weeks following surgery, patient satisfaction, and return to activities of daily living. These outcomes are not defined well enough to compare with other rehabilitation protocols, but they do include subjective, patient-define outcomes which are more useful than radiographic findings or other objective measurements. The study was also the only one able to draw a definitive conclusion on the value or harm of patient restrictions, effecting policy change in that hospital.

The rehabilitation protocol studies covered a broad range for both US and European study authors, but the European authors I read had a stronger tendency toward well-defined methodology and concrete outcome measurements.

Patient Based Outcome Measures

Marx et al (2005) reported on the use of patient-based measures of outcome in rating improvement following total hip and knee arthroplasty. They found that all three scales they tested (MODEMS, Oxford, and WOMAC) were useful in determining patient pain and function levels, although the Physical Component Summary of the SF-36 is more sensitive to small improvements. They feel this is important for future comparisons of surgical methods and devices when final outcomes are going to be extremely similar. The authors have found a way to compare two ‘successful’ operations to determine whether there is any advantage between multiple methods, especially in terms of patient satisfaction.

They did find that, at six months, most hip replacement patients feel that their hip is perfect or near-perfect and the same can be said of knee replacement patients. The study needs to be extended further beyond surgery
Simmons 26

(this study went out one year) to surpass the effects of the considerable pain relief and increased function compared to before surgery. Their SF-36 scale has the added benefit of measuring both mental and physical components of recovery. Its drawback is that it is not joint specific.

Feeny et al (2004) tested the stability of utility scores along with their interpretation. One of the methods they used, called the standard gamble method, seemed unnecessarily complicated and the authors found that in terms of results, a visual scale and marker-state descriptions were found to be much more intuitive to both patients and physicians when they tried to compare and interpret changes in states.

They found that the test-retest reliability of the marker-state scores was similar to the reliability of many other widely used measures (WOMAC, Harris Hip Scale, MACTAR, SF-36). They conclude that marker-state scores help with interpretation of utility scores. This study was completed by repeatedly asking cohorts of patients the same questions to determine test-retest reliability. Judging by the straightforward nature of the questionnaire, I am not sure that retesting was necessary even once, and they performed it twice in addition to a battery of health status and health-related quality of life tests at the beginning of the study. The goal in this study (to determine in marker-state scales help interpret utility scores and whether they are repeatable) seems of questionable worth, although it is easily repeatable, the merits of this new scoring system are the same as the large number currently in use.

Gosens et al (2005) translated the Oxford Hip scoring system into Dutch to allow the combination of the patients’ opinion, physicians’ opinion, and radiographic results. The questionnaire is just 12 items long, but two items specific to hip function were added for this study. The score proved very sensitive to change in a multi-center study and provides valuable insight into whether the patient considers the operation a success. The testing of
the scoring method involved 150 patients and a test retest method, along with comparisons to the two available Dutch-language questionnaires. The OHS was found to be as effective in measuring pain as the two available questionnaires, but was found to be more sensitive to changes and also had no floor or ceiling effects during any of the time periods studied.

The one drawback was found to be low sensitivity to changes in mobility following the surgery. This question asks about mobility with regards to daily activities, which at 7 weeks would be low despite differences in actual mobility level. A new scoring method that is easy to fill out and can be used in conjunction with physician and radiographic assessment can be very useful in determining hip replacement success and encourages the inclusion of patient satisfaction along with other, subjective outcomes.

Measures of outcome that include the patients’ subjective satisfaction as well as pain and functionality are important to really establish joint replacement success. The introduction and increased use of such scales is happening both in the US and Europe. The inclusion of this information in national registries is only happening in New Zealand however.

**Analysis and Conclusions**

The final result of all of these studies is that little more is known about prostheses today than when Charnley introduced the theories of tribology into the design of implants. Metal on metal and ceramic on ceramic implants seem to be showing promise thanks to better manufacturing processes and tighter tolerances, but long term we do not know if they are performing any better than traditional implants – at twice the cost.

The area of improvement remaining is implant life past 20 years where the number of revisions starts to increase dramatically. Right now the group of patients requiring high activity, long-life implants is relatively low, but it is increasing everyday and these patients are willing to pay more for
a higher performance implant. It would be nice if we could say for sure which implant performs best.

Engineers are working to create less and less friction and fewer wear particles – and this pursuit may result in longer lasting implants. But many studies show that wear particles are just one of many variables affecting aseptic loosening (Sundfeldt). If we are basing innovations on empirical and anecdotal evidence, why don’t we increase the amount of evidence available for study? Pooling all available trials and standardizing outcome measures would greatly aid researchers. Even in small prospective trials, a baseline for comparison with current implants would make the trials much more valuable. Many trials use the traditional Charnley for comparison, but as stated before, Charnley had a more ideal and elderly patient group so the results do not take patient factors into account.

In the meantime, we have studies comparing cross-linked polyethylene with normal high density polyethylene, studies comparing minimally invasive procedures with standard incisions, coated and un-coated stems, cemented or cementless or hybrid implant comparisons, etc.

After almost 50 years of successful implants, doctors and patients should be able to make informed decisions based on the results of all of these trials. What is needed is a national registry in the US, where the majority of hip replacements are performed. The registry would need to include standardized outcome measures – both objective clinical and subjective patient scores, complete information on implant design, dimensions, surgical technique, background of patient’s health, and any other factors currently being compared in clinical studies.

There are limitations to such registries. Selection of patients still lacks randomization, but this is already the case, and the larger data pools provided by a registry could outweigh the inherent biases. The registry in Sweden, started in 1979, updates every 24 hours with the results of patient
follow-up and one can see long term statistics for patient factors, implant factors, and surgical factors displayed directly on the site. More detailed information is available to doctors and researchers and an annual report is prepared by the Orthopedic Department of Sahlgrenska University Hospital.

Similar registries exist in England, Scotland, Africa, Australia, New Zealand, and Canada. Not all registries are equal, and even those with high participation do not always hold influence over practice. New Zealand’s registry is the only one to include patient outcomes, while England’s annual report was just on increasing participation – they did not include any information on best practices.

After providing a baseline for researchers (number of implants and number of revisions), providing information on success rates to patients should be a registry’s number one concern. Patients are an important means for change in practice when doctors prefer to stick with what they know or don’t have the time to investigate different procedures. Most surgeons perform less than 10 implants a year.

![Figure 7: Comparison of Cementless implants with Cemented Palacos with Gentamycin showing better survival in the latter at 8.5 years](image)

Through the Swedish registry, a specific type of cement and method has proven itself significantly more successful than uncemented versions. Even better, clinicians have moved toward procedures that show more success and the revision rate in Sweden has dropped from the worldwide and US average of
18% to around 8% (Maloney 1583). The annual reports in these countries are also able to compare survival rates of implants not just on a national basis, but also controlling for patient factors.

These are the kinds of comparisons that are necessary to evaluate design and method differences. These are the kinds of comparisons that should be required of all innovations. The impression given by the short ranges in these studies is that doctors and engineers want to try something new for the sake of trying something new, because they follow it for five to ten years and then move on. Long term studies are the exception, and they all seem to show similar results to the original Charnley implant. For the sake of patients undergoing an increasingly frequent procedure, we need to get our collective act together. We need to invest our research time and effort in finding real improvements and ways to measure them.

**Comments on Cultural Aspects of Joint Registries**

My own opinion, not backed by research, is that the reason the US does not already have a registry is actually a cultural issue more than anything. Many orthopedic surgeons have been calling for a registry for a number of years and the arguments against one are far weaker than the arguments for.

In New Zealand, the doctors were tired of depending on studies performed in the US and Europe for their practices, so they decided to make a registry and they pay for it themselves. Private doctors pay $10 per case entry and the public hospitals pay a flat $32,000 a year. It’s a self-funding instrument, and the people who created it did so out of a desire to help each other and help their patients. This kind of non-competitive, close-knit community of doctors is not likely to be found in the United States.

Private practice doctors are extremely competitive and the best evidence is their reports that do not give enough information to aid small clinicians. It seems they are not looking to help, just to brag. Also, the
United States is very litigious. Even satisfied patients will sue if they think they might have a case, because most malpractice insurance companies will settle to avoid paying court fees or attracting negative attention.

The directory in Sweden is admittedly easier to maintain than one would be in the US because of population mobility, but thanks to the internet, patient tracking could be relatively simple. Most patients in the US are on Medicare, so that database would be used to make sure all patient data was being collected, which is how the registries in most of the other countries do it.

The attitudes that come with being socialist versus the attitudes that come with being capitalist are why it makes sad sense that there exist registries in Europe while we have none here in the US.

Lowering the revision rate, even by a few percentage points, would save enough money to pay for the registry. Volunteer information would be incomplete, but would still help more than the studies being performed now, and real advancements could be made with controlled comparisons, subjective patient outcomes, and a sharing attitude.
Appendix – Descriptions of Commons Scoring Systems

Harris Hip Score

Scoring developed by Dr. William Harris, an orthopedist from Massachusetts, and is used to gauge pain and functionality after hip replacement. It has a possible 100 point total with 44 points for pain (44 for no pain down to 0 for disabling pain), 33 points for function, 14 points for functional activities, and 9 points based on a physical exam. A score of 90–100 is considered excellent, 80–89 good, 70–79 fair, 60–69 poor, and below 60 is a failed result. (Site for calculating the Harris Hip Score: http://exper.ural.ru/trauma/harris_e.phtml)

WOMAC

Western Ontario MacMaster score that assesses lower extremity pain and function in patients with osteoarthritis of the knee or hip. Score is based on pain, disability, and joint stiffness.

SF-36

A thirty-six item short form questionnaire, for self-assessment or assessment over the phone by a trained interviewer. It addresses eight health concepts in order to evaluate health status for medical outcome studies:

1) limitations in physical activities because of health problems;
2) limitations in social activities because of physical or emotional problems;
3) limitations in usual role activities because of physical health problems;
4) bodily pain;
5) general mental health (psychological distress and well-being);
6) limitations in usual role activities because of emotional problems;
7) vitality (energy and fatigue);
8) general health perception

Oxford Hip Score

A joint specific outcome measure tool designed to assess disability in patients undergoing total hip replacement (THR). It is highly sensitive to change, correlates well with patient satisfaction, and has been found to be more responsive than the SF-36 or WOMAC.
Works Cited


