Washington State Health Care Authority

Health Technology Assessment

HTA Final Report
Hip Resurfacing

Date: Friday, October 23rd, 2009

Health Technology Assessment Program
676 Woodland Square Loop SE
P.O. Box 42712
Olympia, WA 98504-2712
http://www.hta.hca.wa.gov

Hip Resurfacing

Provided by:

Spectrum Research, Inc.

Prepared by:

Robin Hashimoto, PhD
Joseph R. Dettori, PhD, MPH
Nora B. Henrikson, PhD, MPH
Erika Ecker, BS
Jeff Hermsmeyer, BS
September 23, 2009

This technology assessment report is based on research conducted by a contracted technology assessment center, with updates as contracted by the Washington State Health Care Authority. This report is an independent assessment of the technology question(s) described based on accepted methodological principles. The findings and conclusions contained herein are those of the investigators and authors who are responsible for the content. These findings and conclusions may not necessarily represent the views of the HCA/Agency and thus, no statement in this report shall be construed as an official position or policy of the HCA/Agency.

The information in this assessment is intended to assist health care decision makers, clinicians, patients and policy makers in making sound evidence-based decisions that may improve the quality and cost-effectiveness of health care services. Information in this report is not a substitute for sound clinical judgment. Those making decisions regarding the provision of health care services should consider this report in a manner similar to any other medical reference, integrating the information with all other pertinent information to make decisions within the context of individual patient circumstances and resource availability.

TABLE OF CONTENTS

TABLES

FIGURES

APPENDICES

EXECUTIVE SUMMARY

Introduction

Total hip arthroplasty (THA) is a well-established and effective treatment for severe degenerative diseases of the hip that has historically been performed in older, relatively inactive patients between 60 and 80 years of age. Over the past decade, however, THA has become increasingly common in patients under 65 years of age. As the growth of joint replacement continues in younger patients, the demand for THA among patients under 65 years is expected to exceed 50% of all THAs by 2011, up from 44% in 2005. Younger patients receiving hip replacement often have more active life styles than those who are older, causing concern about the longevity of the implant. Evidence suggests that higher rates of implant failure occur as the age of patients receiving the implant gets younger.

Total hip resurfacing (HR) is proposed as a bone-conserving alternative to the conventional THA for young and active patients after optimal medical therapy fails. In contrast to THA, total HR preserves the femoral head and neck, which may facilitate future revision surgery should it be necessary, and additionally, enable the patient to take advantage of newer technology or treatments in the future. Furthermore, hip resurfacing was designed to more closely mimic normal joint biomechanics and load transfer, and may be associated with a lower morbidity rate at the time of revision surgery.

Although two such hip resurfacing devices have been approved for use by the FDA, and a variety of devices have been used both within and outside the U.S. for several years, questions remain regarding a number of important issues.

1. How does the effectiveness of hip resurfacing compare with conventional total hip arthroplasty (THA) with respect to patient functional and activity outcomes and pain relief as well as other outcomes including those related to quality of life?
2. What is the safety of hip resurfacing both over the short- and mid-term (e.g., revision, femoral neck fracture, avascular necrosis, device loosening or migration, osteolysis, heterotopic ossification, etc.), and how does it compare with that of conventional THA?
3. Might specific patient populations in particular benefit from hip resurfacing or have increased risks for complications from its use?
4. How might the substitution of hip resurfacing for THA in a proportion of patients with the appropriate indications impact health care systems and costs?

In light of the possible benefits of total hip resurfacing, the potential impact of its use on health care costs and uncertainties regarding the evidence of effectiveness and safety in the short term and longer time horizons, patients, clinicians, and payers will benefit from a structured, systematic appraisal of the comparative effectiveness, safety, and economic impact of hip resurfacing. Thus, the objective of this Health Technology Assessment is to critically appraise and analyze research evidence on the effectiveness of and complications related to the use of hip

resurfacing in patients with degenerative hip disease and to the extent possible, consider the potential financial impact.

Methods for evaluating comparative effectiveness

Spectrum Research, Inc.'s (SRI) method for technology assessment involves formal, structured systematic search of the peer-reviewed literature across a number of databases in addition to searches of pertinent databases related to clinical guidelines and previously performed assessments. Each included study is critically appraised using SRI's Level of Evidence (LoE) system which evaluates the methodological quality based on study design as well as factors which may bias studies. An overall Strength of Evidence (SoE) combines the LoE with consideration of the number of studies and consistency of the findings to describe an overall confidence regarding the stability of estimates as further research is available. Included economic studies were also formally appraised based on criteria for quality of economic studies and pertinent epidemiological precepts.

Throughout the process, SRI sought clinical review to assure that the clinical components are accurately represented and relevant. In addition, peer-review by clinical experts, health services researchers and those with expertise in economic and outcomes evaluation provide an assessment of the systematic review methodology, analyses and report conclusions.

Inclusion of non FDA-approved devices

We included data from studies that used FDA-approved, FDA-510k investigational, and non FDA-approved total HR devices that otherwise met our inclusion/exclusion criteria. Our clinical experts believed that total HR devices are similar enough that including all devices in this review was appropriate, and that the results using one device could be reasonably generalized to other devices as well. Including all devices in this review provides more information to inform readers of this report on efficacy, effectiveness and safety of the procedure of hip resurfacing. Nevertheless, in our results, we attempt to identify whether an FDA-approved device was used in each study.

Results

For key question 1, we identified a total of 13 studies, three of which used FDA-approved devices, four used an FDA-510k investigational device and were conducted as part of an FDA investigational device exemption (IDE) trial, and five used devices not regulated by the FDA. For key question 2, we identified six additional case series: four used FDA-approved devices, one used an FDA-approved device in only 75% of patients, and one used an FDA-510k investigational device. For key question 3, we identified six retrospective cohort studies: two used FDA-approved devices, and four used an FDA-510k investigational device.

Key question 1

What is the evidence of efficacy and effectiveness of total hip resurfacing (HR) compared with conventional total hip arthroplasty (THA)?

Efficacy

There were three RCTs that provided one-year data regarding the efficacy of total HR compared with THA. In general, there were no significant differences in functional, quality of life and activity scores between groups after one year. We found no mid- or long-term comparative data on efficacy.

WOMAC scores
No significant differences were identified in WOMAC scores between total HR and THA groups at one-year follow-up as reported in three RCTs. Normalized postoperative WOMAC scores in the THA group ranged from 87.8 to 97.2, and from 90.4 to 96.9 in the total HR group.

SF-36 scores
No significant differences were identified in SF-36 quality of life scores between treatment groups at one-year follow-up as reported in two RCTs. Physical postoperative SF-36 scores ranged from 51.3 to 53.3 in the THA group and from 51.2 to 55.2 in the total HR group; mental postoperative SF-36 scores ranged from 52.1 to 55.1 in the THA group and from 51.9 to 53.9 in the total HR group.

UCLA activity scores
There was conflicting evidence regarding one-year UCLA activity score. One study reported significantly higher scores in the total HR group, while the other two RCTs found no significant difference between groups. In the THA group, postoperative scores ranged from 6.3 to 8.3; in the total HR group, postoperative scores ranged from 6.8 to 8.0.

Merle D'Aubigné (MA) scores
There were no significant differences in mean one-year MA scores between total HR and THA cohorts as reported by two RCTs. Postoperative scores ranged from 16.6 to 18.0 in the THA group, and from 16.7 to 17.9 in the total HR group.

Pain scores
There was no significant difference in WOMAC pain scores at one-year follow-up between groups as reported by one RCT.

Walking speed
One RCT reported no significant differences in normal and fast walking speeds between treatment groups at one-year follow-up.

Effectiveness
There were eight cohort studies (one prospective and seven retrospective) that provided short-term (< 5 years) and one retrospective cohort study that reported mid-term (5.9 years) data regarding the efficacy of total HR compared with THA:

Harris hip scores (HHS)
No significant differences were identified in short-term postoperative Harris Hip scores as reported by one prospective and seven retrospective cohort studies. The mean postoperative HHS ranged from 89.7 to 96.2 in the THA group and from 90 to 98.1 in the total HR group.

Postoperative scores were slightly higher in the total HR group in six of the seven studies, although no differences between groups were statistically significant.

Oxford score
No significant differences were found in 5.9-year Oxford scores as reported by one retrospective study (THA: 18.5, total HR: 15.9).
Quality of life scores (EQ-5D, SF-12)
Data suggests that quality of life scores may be higher following total HR compared to THA. Postoperative EQ-5D scores were significantly higher in the total HR group (0.9) compared to the THA group (0.78) as reported by one study with 5.9-year follow-up. Another study with short-term follow-up reported significantly higher postoperative SF-12 physical scores in the total HR (53.6) group versus the THA (47.0) group, but found no significant difference in the postoperative SF-12 mental scores between cohorts.

Activity scores (UCLA, Mont)
Data suggests that patients treated with total HR may have significantly higher postoperative activity levels than those who received THA. Postoperative UCLA activity scores were higher in the total HR group in all three studies that reported this outcome; this difference was statistically significant in the two studies in which the p-value was reported, one of which had mid-term follow-up with a mean of 5.9 years. Postoperative UCLA activity scores ranged from 3.6 to 6.8 in the THA group, and from 6.1 to 8.4 in the total HR group. Two studies measured short-term patient activity using a scoring system devised by Mont (2009). Both reported significantly higher activity scores for the total HR group (ranged from 10.0–11.5) compared to the THA group (ranged from 5.3–7.0).

Pain
No significant differences were found in the postoperative level of pain between total HR and THA treatment groups as reported by five retrospective cohort studies with short-term follow-up. Postoperative pain scores ranged from 0.7 to less than 2 in the THA group, and from 0.9 to less than 2 in the total HR group. All but one study reported pain scores from the VAS (visual analogue scale); one reported Harris hip pain component scores, which were normalized here.

Key question 2
What is the evidence related to the safety profile of hip resurfacing?

Short-term (< 5 years) safety data were reported by three national registry studies, two RCTs, and eight cohort studies (one prospective and seven retrospective), while mid-term (5–10 years) safety data was reported by one retrospective cohort study, six case-series. No long-term safety data were available.

Revision
Short-term:
Data from three national registry studies suggest that revision rates are statistically higher in those receiving total HR compared with THA after three years of follow-up.

- From the National Joint Replacement Registry, Australia, the cumulative 3-year revision rate is 3.1% for total HR (95% CI = 2.8%, 3.5%) and 2.5% for conventional THA (95% CI = 2.5%, 2.6%).
- The three year risk for revision in the Swedish Registry is three times higher in those with hip resurfacing compared with THA, RR = 3.33 (95% CI = 2.04, 5.43).[1]
- The National Joint Registry, UK overall 3-year revision rates were lowest in patients who received cemented prostheses (0.7%, 95% CI: 0.6%, 0.8%) and highest after hip resurfacing (2.8%, 95% CI: 2.3%, 3.5%; hazard ratio = 3.6, 95% CI: 2.9%, 4.6%) in patients ≤65 years.

Among comparative studies, the short-term revision rates ranged from 0% to 4.3% of hips in the THA group, and from 0% to 7.8% of hips in the total HR group.

- The difference in 1-year revision rates in one RCT is 0.9% in favor of THA. The difference in short-term revision rates between total HR and THA in eight cohort studies varied: 4 favored THA, 2 favored total HR and 2 reported equal rates.

There is some data that suggest revision rates vary depending on the total HR system.

- In the Australian National Joint Replacement Registry, the Birmingham total HR system has a statistically lower 3-year cumulative percent revision rate (2.5%) compared with the Articular Surface Replacement (ASR) (6.0%) and the Durom (5.8%) systems.

Mid-term:

The Australian National Joint Replacement Registry has 7-year follow-up data for 10,623 HRs.

- A comparison of time to revision revealed a significantly higher revision rate for total HR compared with conventional THA, adjusted hazard ratio = 1.42 (1.24, 1.63), P < .001.
- The cumulative 7-year revision rate for total HR is 4.6% for total HR (95% CI: 3.9%, 5.4%) and 3.4% for conventional THA (95% CI: 3.2%, 3.7%).

One small retrospective cohort study with 5.9-year follow-up reported similar revision rates in hips treated with THA (7.8% planned) compared to those who underwent hip resurfacing (7.1%). Revision rates for resurfaced hips ranged from 0% to 7.7% as reported by six case-series.

Complications

Reported risks of other complications in the short-term for total HR are generally low except for heterotopic ossification. Revision risk:

- Femoral neck fractures range from 0.4% to 2.6%
- Avascular necrosis range from 0.4% to 2%
- Femoral component loosening from 0% to 3.6%
- Acetabular component loosening from 0% to 1.8%
- Acetabular component migration from 0% to 1.9%
- Femoral component migration was not detected in any hips
- Heterotopic ossification rates ranged from 0% to 42.7%.

Metal ion safety

Patients with metal-on-metal total HR or THA are likely to experience elevated metal serum levels (Co and Cr). Concerns have been raised regarding the safety of and risks associated with

prolonged exposure to metal ions, and whether such exposure may increase the risk of cancers or metabolic disorders. However, an association between MoM prostheses and cancer or metabolic disorders has not been reported with the current length of follow-up. The results from long-term monitoring will be needed to assess the risk of metal ion exposure. Since metal ions have been shown to cross the placental barrier, metal-on-metal hip arthroplasty is contraindicated for females of child-bearing age.

Key question 3
Is there evidence of differential efficacy or safety issues with use of hip resurfacing?

Preoperative diagnosis
Three diagnoses were compared to evaluate whether total HR results vary by diagnoses in the Australian National Joint Replacement Registry.

- The Registry reports a significantly higher risk of revision with resurfacing procedures undertaken for dysplasia (hazard ratio = 2.08; 95% CI 1.35, 3.19) compared with osteoarthritis. The five-year cumulative percent revision is four times greater in those receiving total HR compared with THA for dysplasia, 12% versus 3%.
- The Registry reports a significantly higher risk of revision with resurfacing procedures undertaken for AVN (hazard ratio = 1.89; 95% CI 1.03, 3.5) compared with osteoarthritis. The five-year cumulative percent revision for AVN is approximately 2-times greater in those receiving total HR compared with THA, 6.1% versus 3%.

Sex and component head size
Three registries agree that the risk of revision is higher among female patients receiving total HR compared with males.

- The Swedish Registry reported a two-fold increased risk of revision for females.
- The National Joint Registry from the UK reported a 57% increased risk of revision after three years, 3.6% (95% CI 2.7%, 4.8%) for females and 2.3% (95% CI 1.9%, 3.3%) for males.
- The Australian National Joint Replacement Registry reported that the 5-year cumulative percent revision risk for females is 2.5 times higher than males, 6.5% versus 2.6%. The Registry recently reported an inverse relationship between femoral component head size and the risk of revision. As the head size increases, the five year cumulative percent revision decreases. After adjusting for femoral component head size, the Registry found no significant difference in the risk of revision between males and females concluding that the higher risk in females is probably due to the size of the femoral component head size.
- One cohort study reports a higher risk for females than males when the outcome is stem radiolucencies, hazard ratio = 3.1 (95% CI: 109, 1.59). Females were shown to have a significantly higher risk of metallosis than their male counterparts in a second cohort study, relative risk = 4.94 (95% CI: 1.3, 18.3).

Surface Arthroplasty Risk Index (SARI)
The Surface Arthroplasty Risk Index (SARI) is a severity measure that evaluates a patient's risk for radiologic changes or revision surgery following modern hybrid hip resurfacing. The SARI assigns risk based on the following: UCLA activity score > 6 (1 point); previous arthroplasty (1 point); weight less than 180 pounds (2 points); and femoral cysts with a diameter of more than 1 cm (2 points).

- In one cohort study, patients with problematic hips had a significantly higher mean SARI score (4.7, all ≥ 3) than those without hip problems (mean SARI score: 2.6) (P = .001). A SARI score of four or higher was associated with a twelve-fold increase in early complications compared to patients with a score of three or less.
- A second cohort study noted that a SARI score greater than 3 was associated with earlier time to revision (P = .004). Patients with a SARI score > 3 had a lower four-year survival (88.8%; 95% CI: 80.0, 97.6%) than those with a SARI score ≤ 3 (97%; 95% CI; 94, 100%). The risk of femoral radiolucency was also 4.2-fold higher in patients with a SARI score higher than 3 than in patients with a lower SARI score.

Obesity
Two low quality studies evaluated the effect of obesity on total HR with conflicting results. One reported lower revision risk with increasing obesity, and one reported higher.

Key question 4
What is the evidence of cost implications and cost effectiveness of hip resurfacing?

There is limited evidence on the economic implications of hip resurfacing from two published articles and one HTA. Revision rates are important input factors in the prediction models, and no study estimated the revision rates using current data. Although further study is necessary to include more current data, there is currently insufficient evidence to warrant a conclusion about the economic value of hip resurfacing in a US setting.

Summary

Key Question 1: What is the evidence of efficacy and effectiveness of hip resurfacing?

HR vs. THA	Strength of evidence	Conclusions/Comments
1. Efficacy (≤1 year)	**Moderate**	• There is moderate evidence from three small randomized controlled trials that total HR is similar to THA with respect to short-term (1 year) functional, quality of life and activity outcomes.
> 1 year	**No evidence**	• There are no data available to assess efficacy beyond one-year follow-up.
2. Effectiveness Short-term (<5 years)	**Low evidence**	• There is low evidence from studies directly comparing total HR with THA to suggest that short-term (≤5 years) patient-reported outcomes, clinician-based outcomes, and pain are similar comparing total HR and THA. Activity scores tend to be slightly higher (better) in total HR patients.
Mid-term (5–10 years)	**Very low evidence**	• There is very low evidence from one cohort study to suggest that at an average of 5.9 years follow-up, patients treated with total HR may have better quality of life and activity outcome scores, but similar functional scores, compared with those treated with THA.

Key Question 2: What is the evidence about the safety profile for hip resurfacing?

	Strength of evidence	Conclusions/Comments
1. Revision Short-term (<5 years)	**Moderate evidence**	• There is moderate evidence that short term revision rates are slightly higher in patients treated with total HR compared with those treated with THA. The difference in 3-year revision rates between total HR and THA in 3 registry studies range from 0.6% to 2.5% in favor of THA. The difference in 1-year revision rates in one RCT is 0.9% in favor of THA. The difference in short-term revision rates between total HR and THA in eight cohort studies varied: 4 favored THA, 2 favored total HR and 2 reported equal rates.
Mid-term (5-10 years)	**Low evidence**	• There is low evidence from one large registry study that 7-year revision rates are higher in patients receiving total HR versus THA (hazard ratio = 1.42, rate difference = 1.3%). Data from one small cohort study with a mean follow-up of 5.9 years reports revision rates that are similar between total HR and THA.
Long-term (10+ years)	**No evidence**	• There is no evidence comparing long-term revision rates between total HR and THA.
2. Other complications	**Low evidence**	• Reported risks of other complications in the short-term for total HR are generally low except for heterotopic ossification; the risk of femoral neck fractures range from 0.4–2.6%, avascular necrosis from 0.4-2%, femoral component loosening from 0-3.6%, acetabular component loosening from 0-1.8%, acetabular component migration from 0–1.9%, and femoral component migration was not detected in any hips. Heterotopic ossification rates ranged from 0-42.7%.
3. Learning curve threshold	**Very low evidence**	• A number of studies identified that the rate of major complications (including femoral neck fracture and revisions) decrease as surgeons gain experience performing total HR. The studies suggested that experience is associated with improved surgical technique and patient selection. However, with respect to identifying the number of procedures necessary for improved outcome, no consistent threshold was identified.
4. Metal ion safety	**Very low evidence**	• Patients with metal-on-metal total HR are likely to experience elevated metal serum levels (Co and Cr). Concerns have been raised regarding the safety of and risks associated with prolonged exposure to metal ions, and whether such exposure may increase the risk of cancers or metabolic disorders. However, an association between total HR and cancer or metabolic disorders has not been reported with the current length of follow-up. The results from long-term monitoring will be needed to assess the risk of metal ion exposure.

Key Question 3: Is there evidence of differential efficacy or safety issues with use of hip resurfacing?

	Strength of evidence	Conclusions/Comments
1. Dysplasia vs. other arthritic conditions	**Low evidence**	• There is low evidence to suggest that short-term revision rates are twice as high in patients who receive total HR for a primary diagnosis of dysplasia compared with patients of primary osteoarthritis. The 5-year cumulative revision percent for dysplasia is four times greater in those receiving total HR compared with THA (12% vs. 3%) in one registry study. One small prognostic study supported this data, with 5.2% revision rates in dysplasia patients compared with 0% revision rates in osteoarthritic patients.
2.Osteonecrosis (AVN) vs. other arthritic conditions	**Low evidence**	• There is low evidence to suggest that short-term revision rates are slightly higher in patients who receive total HR for a primary diagnosis of ostenecrosis (AVN) compared with patients of primary osteoarthritis. The 5-year cumulative revision percent for dysplasia is two times greater in those receiving total HR compared with THA (6% vs. 3%) in one registry study and rates are the same in one small

		prognostic study.
3. Gender	**Moderate evidence**	• There is moderate evidence from three registries that 3- and 5-year revision rates are higher in females than in males (hazard ratios range from 1.57 to 2.5). Much of the difference in rates between sexes disappeared in one study when controlling for femoral component head size; the smaller the head, the higher the failure rate.
4. Obesity	**Very low evidence**	• Two low quality studies evaluated the effect of obesity on total HR with conflicting results. One reported lower revision risk with increasing obesity, and one reported higher.
5. SARI index	**Very low evidence**	• Two low quality studies evaluated the effect of the SARI index on total HR. Both suggest a SARI score $\geq$ 3 preoperatively results in an increased risk of early complications and revision.

Key Question 4: What is the evidence of cost implications and cost effectiveness of hip resurfacing?

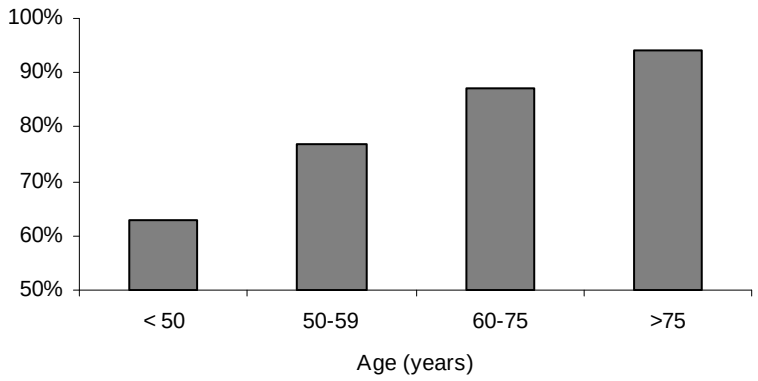
	Strength of evidence	Conclusions/Comments
5. SARI index (cont.)	**Very low evidence**	• There is limited evidence on the economic implications of hip resurfacing from two published articles and one HTA. Revision rates are important input factors in the prediction models, and no study estimated the revision rates using current data.

1. Appraisal

1.1. Rationale

Total hip arthroplasty (THA) has proven to be effective for elderly patients with hip pain and dysfunction from non-inflammatory arthritis such as osteoarthritis, traumatic arthritis, avascular necrosis, dysplasia, or inflammatory arthritis such as rheumatoid arthritis. Over the last decade, the prevalence of THA in younger patients (those under 65 years) has increased. As the growth of joint replacement continues in younger patients, the demand for THA among patients under 65 years is expected to exceed 50% of all THAs by 2011, up from 44% in 2005.[2] Younger patients receiving hip replacement often have more active life styles than those who are older, causing concern about the longevity of the implant. Evidence suggests that higher rates of implant failure occur as the age of patients receiving the implant get younger, Figure 1.

Figure 1. Implant Survival (%) after 16 Years*



* Estimated from Swedish Total Hip Replacement Register; Annual report 2007: www.jru.orthop.gu.se

Hip resurfacing is proposed as a bone conserving alternative to the conventional THA after optimal medical therapy fails. Unlike THA, hip resurfacing does not involve the removal of the femoral head and neck or removal of bone from the femur. Rather, the head, neck and femur bone is preserved in an effort to facilitate future surgery should it be necessary and to enable the patient to take advantage of newer technology or treatments in the future. Hip resurfacing is anatomically and biomechanically more similar to the natural hip joint.

Proposed benefits of hip resurfacing include: increased stability, flexibility and range of motion. Younger patients needing full joint replacement that are expected to outlive the full replacement may benefit from symptom relief and increased bone preservation to tolerate a subsequent replacement surgery later. With hip resurfacing, the risk of dislocation is lower and the possible activity level is higher with less risk than THA.

However questions remain about the unknown longevity and durability of the procedure; the magnitude of the reported failure rates; the appropriate patient selection criteria (e.g., age, gender, tried and failed therapies); impact on long term health outcomes; additional surgical risks and complications from multiple surgeries, and the health system impacts of a surgery designed to delay but not eliminate need for later surgery.

1.2. Key Questions

Key questions are developed by the Washington State Health Technology Assessment Program.

When used as an alternative in patients where total hip replacement is indicated:

Key Question 1:
What is the evidence of efficacy and effectiveness of hip resurfacing?

Key Question 2:
What is the evidence about the safety profile for hip resurfacing?

Key Question 3:
Is there evidence of differential efficacy or safety issues with use of hip resurfacing?

Key Question 4:
What is the evidence of cost implications and cost effectiveness of hip resurfacing?

1.3. Outcomes Assessed

1.3.1. Efficacy and effectiveness measures

Studies reported functional and activity scores from generic quality of life, disease specific clinician-based or patient-reported outcomes, and pain, Table 1.

- Three quality of life measures were used: the EQ-5D, SF-36 and SF-12 outcomes measures. Domains assessed by the EQ-5D include patient mobility, self-care, usual activity, pain and anxiety/depression.[3] SF-36 and SF-12 include 8 subscales that assess physical function, role limitations due to physical health problems, pain, general health, vitality, limitations due to emotional problems, and mental health.[4] The SF-12 measures the same subscales as the SF-36 with fewer items.[5]
- Two patient-reported disease specific outcomes measures were used, the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), and the Oxford Hip Score. The WOMAC assesses the patient's pain, stiffness and physical function.[6] The Oxford Hip Score uses 12 questions to assess perception of pain and function.[7]
- Two different clinician based outcomes, the Harris Hip Score (HHS)[8] and the Merle D'Aubigne Hip Score,[9] were also reported frequently; both combine a component of patient symptoms with physician assessment.

- Two activity scores were used. The activity score of Mont attempts to assess the frequency and duration of activity in which each patient regularly participates.[10] The UCLA activity scale seeks to determine how active a patient is on a 1-10 scale with one representing a person who is wholly inactive and dependent on others, and 10 representing a person who regularly participates in impact sports.[11]
- Pain was assessed by some studies using a visual analog scale.[12]

Table 1. Outcome measures

Outcome measure	Clinician or patient reported	Instrument type	Components	Score range	Interpretation
Mont Activity[10]	Patient	Disease specific	Each activity that the patient regularly participates in is assessed: Score = frequency (# times per week) x duration (hours) x weighed points (1–3; based on competitiveness)	0–?*	Low activity patients: 0–8 High-activity patients: ≥9*
EQ-5D (European Quality of Life)[3]	Patient	Generic	Mobility (1–3) Self-care (1–3) Usual activity (1–3) Pain (1–3) Anxiety/depression (1–3)	0–1†	Optimal health: 1 Death: 0
HHS (Harris hip score)[8]	Clinician	Disease specific	Pain (44) Function (47) Deformity (4) Range of motion (5)	0–100	Excellent: 90–100 Good: 80–89 Fair: 70–79 Poor: <70
Merle D'Aubigne hip score[9]	Clinician	Disease specific	Pain (6) Mobility (6) Walking ability (6)	0–12‡	‡Very good: 11–12 Good: 10 Medium: 9 Fair: 8 Poor: <7
Oxford hip score[7]	Patient	Disease specific	12 questions concerning the perception of pain and function (1–5 each)	12–60	Higher score = lower function
SF-36 (Short Form 36 health survey questionnaire)[4]	Patient	Generic	8 subscales (# items) Physical functioning (10) Role limitations due to physical health problems (4) Bodily pain (2) General health (5) Vitality (4) Social functioning (2) Role limitations due to emotional problems (3) Mental health (5)	0–100 for each subscale (total score not used)	Lower score = greater disability
SF-12 (Short Form 12 health survey questionnaire)[5]	Patient	Generic	2 subscales (# items) *Physical health* General health (1) Physical functioning (2) Physical role limitations (2) Bodily pain (1) *Mental health* Emotional role limitations (2) Social functioning (1) Vitality/mental health (3)	0–100 for each subscale	Lower score = greater disability
UCLA activity scale[11]	Patient	Disease specific	Activity (10)	1–10	Unrestricted activity: 10 Bedridden: 1
VAS pain (Visual Analogue Scale)[12]	Patient	Generic	Pain	0–10	No pain: 0 Worst pain imaginable: 10
WOMAC (Western Ontario and McMaster Universities OA index)[6]	Patient	Disease specific	Pain (20) Stiffness (8) Physical function (68)	0–96	Higher score = greater disability

* Mont (2009): the maximum possible score was not reported.

† EQ-5D: final score is a 5-digit descriptor that corresponds to the level of disability in each subcomponent and ranges from 11111–33333; each score is assigned a preferential weight (e.g., 21111 = 0.85) to obtain a final score of 0 to 1.

‡ MA final score: the pain and walking ability scores are summed and then adjusted down by 1–2 grades based on the mobility score for the final clinical grade.

1.4. Key considerations highlighted by clinical experts:

1.4.1. Intervention

The issue the literature addresses rightly focuses on a tension between slightly higher short-term complication and reoperation rates with total hip resurfacing (HR) (in the higher quality studies) versus the potential benefits of a more bone-sparing approach and the possible increased durability of a metal-on-metal bearing couple. Issues which are unclear include the very long-term durability with conventional total hip arthroplasty (THA) using metal-on-polyethylene bearings in the very young/active/male population, which presumably will bring a very high revision, reoperation, and complication rate, and the safety of metal-on-metal articulations pertaining to metal allergy and systemic deposition of metal ion species and corrosion products. Much of the literature on total HR and THA is written by "advocates" of the procedures, and this potential bias needs to be taken into account.

As with any orthopaedic intervention, the decision to proceed with surgery needs to be individualized. In the context of total hip resurfacing versus conventional total hip arthroplasty, one needs to consider the pros and cons of both. Hip resurfacing is intended for patients with high functional demands for whom traditional total hip arthroplasty would be a poor option because of anticipated future failure and subsequent revision surgery. Many clinical experts believe that total hip resurfacing is a bone sparing procedure best done in males under the age of 55 years with good bone stock, good health, an active lifestyle, and minimal femoral deformity or leg length discrepancy. There should be no history of renal disease or metal sensitivity. Patients with significant avascular necrosis, a history of infection, or women of childbearing age may not be suitable candidates for this procedure.

Patients should consider hip resurfacing when: arthritis has been resistant to conservative measures; the patient is sufficiently healthy to undergo the procedure; the patient understands the risks and alternatives; the surgeon is trained and experienced in hip-resurfacing surgery; and no medical or surgical contraindication to hip resurfacing exists.

The most common risk of hip resurfacing is fracture of the femoral neck. Infections are a rare but potentially catastrophic problem. Component loosening is an infrequent complication. Potential risks from the production of ions (cobalt and chromium) have yet to be clearly documented in the clinical literature. There is a theoretical concern that metal ions may pose a cancerous risk. A patient who has kidney disease may have difficulty filtering these ions from the blood. Hip resurfacing is not recommended for women of childbearing age because of the uncertainty regarding the effects of metal ions on the developing fetus. Hypersensitivity to metal ions is a risk that is being increasingly

recognized, and therefore patients with a history of metal allergy should not undergo this procedure.

1.4.2. Costs

Cost data is somewhat controversial and hard to decipher, but the metal-on-metal bearings (whether in a resurfacing or a replacement) will typically be more expensive. Since the hard-on-hard surfaces will be used in younger active patients (both in resurfacing and in replacement), comparing the average cost of a resurfacing to the average cost of a replacement will artificially bias the results: the cost of a resurfacing in a 50 year old male is being compared to a the cost of a cheaper cemented, metal-on-polyethylene replacement in a 75 year old female.

1.4.3. Patient considerations

In all women, the revision rate is higher for total HR compared to THA, but the rates are nearly similar for those under 55 years. In men over 65 years of age, revision rates for total HR are higher than for THA, while revision rates for resurfacing are LOWER than for THA in men under 65 years. Therefore, resurfacing can be recommended for most young men (those with good bone quality, minimal hip deformity and DJD from a source other than post infectious arthritis), and some young women. Resurfacing at this time can be recommended for selected older patients.

1.4.4. Professional considerations

Because hip resurfacing devices have received approval only recently in the United States, many communities do not have surgeons trained in this procedure. For FDA-approved devices, the device manufacturers require that surgeons who implant their devices be properly trained for technique. There is a definite learning curve for this procedure. To reduce complications, this procedure should be performed by surgeons with extensive experience in this surgery. This has been well documented.[13]

1.5. Inclusion of non FDA-approved devices

We included data from studies that used both FDA-approved and non FDA-approved total HR devices that otherwise met our inclusion/exclusion criteria. Our clinical experts believed that total HR devices are similar enough that including all devices in this review was appropriate, and that the results using one device could be reasonably generalized to other devices as well. Including all devices in this review provides more information to inform readers of this report on efficacy, effectiveness and safety of the procedure of hip resurfacing. Nevertheless, in our results, we attempt to identify which devices were used in each study.

1.6. Washington State utilization and cost data

The following data were provided from the Washington State Health Care Authority and represent estimates for costs and utilization from the Uniform Medical Plan, Labor and Industry and Medicaid.

Table 2. Count of Procedures by Year, Washington State

UMP, L&I, & Medicaid

ICD-9 Procedure Codes	2005	2006	2007	2008	Total
00.85 (total hip resurfacing)	0	3	20	22	**45**
00.86 (resurfacing, femoral head)	0	1	2	2	**5**
00.87 (resurfacing, acetabulum)	0	0	0	0	**0**
81.51 (total hip replacement)	432	471	487	614	**2004**
81.52 (partial hip replacement)	108	100	82	102	**392**
Total	**540**	**575**	**591**	**740**	**2446**

Table 3. Amount Paid* by Procedure by Year, Washington State

UMP, L&I, & Medicaid

ICD-9 Procedure Codes	2005	2006	2007	2008	Total
00.85 (total hip resurfacing)	$0	$63,532	$369,675	$413,648	**$846,855**
00.86 (resurfacing, femoral head)	$0	$17,130	$33,443	$56,845	**$107,418**
00.87 (resurfacing, acetabulum)	$0	$0	$0	$0	**$0**
81.51 (total hip replacement)	$2,939,023	$6,121,243	$3,747,832	$5,266,802	**$18,074,900**
81.52 (partial hip replacement)	$288,676	$300,097	$986,672	$1,255,759	**$2,831,204**
Total	**$3,227,699**	**$6,502,002**	**$5,137,623**	**$6,993,053**	**$21,860,377**

* includes facility, professional and other payments

Table 4. Amount Paid* per Procedure by Year (NonMedicare)

UMP, L&I, & Medicaid

ICD-9 Procedure Codes	2005	2006	2007	2008
00.85 (total hip resurfacing)	0	$23,135	$22,451	$20,638
00.86 (resurfacing, femoral head)	0	$19,991	$18,172	$30,229
00.87 (resurfacing, acetabulum)	0	0	0	0
81.51 (total hip replacement)	$17,902	$18,650	$18,361	$20,037
81.52 (partial hip replacement)	$20,071	$17,102	$21,750	$21,487

* includes facility, professional and other payments. Amount paid divided by procedure count.

Table 5. Age and Sex by Procedure

UMP, L&I, & Medicaid		Procedure Code				
Age	Gender	00.85	00.86	81.51	81.52	**Total**
0-19	F	0	0	1	3	**4**
	M	0	0	0	0	**0**
20-44	F	3	0	66	9	**78**
	M	6	1	116	11	**134**
45-64	F	7	2	579	74	**662**
	M	27	2	588	53	**670**
65-74	F	1	0	243	37	**281**
	M	1	0	193	10	**204**
75-84	F	0	0	115	64	**179**
	M	0	0	67	31	**98**
85+	F	0	0	26	76	**102**
	M	0	0	8	24	**32**
Total		**45**	**5**	**2002**	**392**	**2444**

Data Notes:
The data for UMP in 2008 also includes Public Employees Health Plan (formerly PEBB) members being served by Aetna. This adds approximately 25,000 people to the analysis.

Table 4 does not include UMP and Aetna Medicare patients in the analysis because Medicare is the primary payer and this skews the cost data.

2. Background

2.1. History of Hip Resurfacing

Total hip arthroplasty (THA) is a well-established and effective treatment for severe degenerative diseases of the hip. Originally designed for elderly patients,[2, 14, 15] most THAs have historically been performed in patients between 60 and 80 years of age (late middle-aged and elderly) who are relatively inactive. Over the past few years, however, THA has become increasingly common in young, active patients. Data from the Swedish National Hip Arthroplasty Register suggests that younger patients are more likely to need revision surgery following THA than their older counterparts. Survival rates increase with age (at the time of the primary THA): while patients aged 60 years or older have a survival rate greater than 82% 16 years following surgery, those aged 50 to 59 years have a survival rate of ~76%, and patients younger than 50 years of age have a 16-year survival rate of only ~65%.[1] Therefore, the longevity of THA for this patient population is of concern. The need for hip prostheses in younger patients is only expected to increase: by 2011, patients under 65 years of age are projected to account for over half of THAs, and by 2030 the number of these procedures in the United States is expected to triple.[2]

Total HR is an alternative to THA for young and active patients. In contrast to THA, total HR conserves femoral bone, attempts to maintain normal joint biomechanics and load transfer, and may be associated with a lower morbidity rate at the time of revision surgery.

Total HR prostheses were initially introduced to the medical community in the late 1970s, but most surgeons abandoned the technique in the 1980s due to high failure rates. These first generation total HR prostheses failed largely due to the wear created by the metal-on-polyethylene design, which resulted in early component loosening, osteolysis, and subsequent femoral neck fractures.[16, 17]

The redesign of total HR prostheses in the 1990s spurred renewed interest in this procedure for the treatment of younger, active patients. With correct patient selection, surgeon education, and operative technique, survivorship at five years has been shown to be comparable with that of conventional hip replacements.[18, 19] Modern total HR components consist of high-carbide cobalt chrome metal-on-metal bearings that articulate against an intermediate synovial fluid film, a design that results in low surface wear. Hybrid fixation, which utilizes cementless fixation of acetabular components, has been correlated with a lower incidence of early acetabular component failure. Modern total HR has been associated with promising early and mid-term results. Survivorship in younger patients has been reported to be 99.8% at a mean of 3.3 years in 446 osteoarthritic hips[20]; 94.4% at 4 years in 400 hips[21]; and 99.1% at a mean of 5 years follow-up in a prospective study of 230 resurfaced hips.[22]

2.2. Advantages of Hip Resurfacing versus Total Hip Replacement

During a total hip arthroplasty procedure, the entire femoral head is removed and replaced with a metal prosthetic ball. In contrast, during a hip resurfacing procedure, only the surface of the femoral head is removed and replaced with a hollow cap inserted into the hollow part of the femoral shaft. On the pelvic side, both procedures involve replacing the acetabulum with a metal cup, which functions as the socket of the new hip joint.

One of the major advantages of total HR is the preservation of femoral bone stock.[11, 23, 24] Following a conventional THA, osteolysis of the periprosthetic bone may occur because the load has been transferred to the implant. As a result of this stress shielding, the bone becomes thin, weak, and prone to fracture. In contrast, normal femoral loads are maintained following total HR, which helps to maintain normal bone density and quality.[25] Preservation of the femoral head also improves function and allows for conversion to a THA in the future if needed.[26, 27] Total HR is associated with lower morbidity at the time of revision surgery than THA.[28, 29] In addition, patients tend to recover more quickly following total HR surgery than they do after THA. Other theoretical advantages of hip resurfacing over THA are a decreased risk of dislocation due to the larger femoral head and better replication of normal anatomy,[30, 31] a greater range of motion, and avoidance of wear debris-induced osteolysis.[32-35]

2.3. Target Population

The ideal candidates for metal-on-metal hip resurfacing are younger, active adults with isolated degenerative diseases of the hip, good proximal femoral bone quality and morphology, and normal kidney function.[18] The aim of hip resurfacing is to allow the patient to resume a physically active lifestyle after pain relief is achieved.

2.4. Indications for Hip Resurfacing

Total HR is intended for reduction or relief of pain and improved hip function in skeletally mature patients with the following conditions: (1) non-inflammatory degenerative arthritis such as osteoarthritis, traumatic arthritis, dysplasia or avascular necrosis; or (2) inflammatory arthritis such as rheumatoid arthritis.

Primary or secondary osteoarthritis (OA), or degenerative arthritis, is the most common form of arthritis and typically occurs after middle age. OA is characterized by the chronic breakdown of articular cartilage and underlying subchondral bone in the joints due to the combination of wear and tear with a variety of hereditary, developmental, and metabolic factors. The hip and knee are the most commonly affected joints. Clinical symptoms of OA may include joint pain, tenderness, stiffness, inflammation, creaking, locking of joints, and disability.[36, 37]

Rheumatoid arthritis (RA) is a chronic, systemic inflammatory disorder that primarily affects the joints. Inflammation of the synovial membrane lining the joint, or synovitis, can lead to the destruction of articular cartilage and ankylosis of the joints. Women are three times more likely than men to have RA and onset occurs most often between ages 40 and 60 years.[38]

Avascular necrosis (AVN) or osteonecrosis, results from temporary or permanent loss of blood supply to an area of bone. This debilitating disorder primarily affects the joints at the shoulder, knee, and hip, and can lead to the destruction of the articular surface of the joint. Total HR is contraindicated if osteonecrosis affects more than half of the femoral head. Osteonecrosis following a hip fracture may occur if the fracture interrupts blood flow to the femoral head, resulting in slow and incomplete healing or even bone death. If severe enough and damage to the hip socket has occurred, replacement surgery becomes necessary.

Ankylosing spondylitis (AS) is a chronic, inflammatory arthritis that affects the joints of spine and the sacroiliac joint of the pelvis and causes eventual fusion of the spine. Medication and physical therapy are common treatments for AS but in severe cases joint replacement may be necessary, particularly in the knees and hips.

Perthes disease occurs only in children, usually between 4 and 10 years of age, and is characterized by a temporary loss of blood to the femoral head resulting in bone death, inflammation, and irritation around the hip joint.

Hip dysplasia is a congenital or acquired deformity of the hip joint and often causes osteoarthritis of the hip at a relatively young age. Arthroplasty, in conjunction with osteotomy, is sometime used to correct the misalignment.

2.5. Contraindications for hip resurfacing

Contraindications for hip resurfacing include the following[39, 40]:

- Severe osteoporosis or osteopenia
- Skeletal immaturity
- Multiple femoral neck cysts greater than 1 cm in diameter
- Infection or sepsis
- Vascular insufficiency, muscular atrophy, or neuromuscular disease severe enough to compromise implant stability or postoperative recovery
- Avascular necrosis involving more than 50% of the femoral head
- Moderate to severe renal insufficiency
- Immunosuppression (ie, AIDS) or high doses of corticosteroids
- Females of child-bearing age due to the unknown effect of metal ion release on the fetus
- Severely overweight (BMI > 35)
- Known or suspected metal sensitivity

2.6. Potential complications/harms of hip resurfacing

Femoral neck fracture
Short-term failure of hip resurfacing prostheses is most commonly due to periprosthetic femoral neck fractures, which account for 37% to 47 % of revisions.[41] Risk factors include a combination of patient-associated (i.e., poor bone quality, obesity, and female gender), technique-related (i.e., notching of the superior part of the femoral neck, varus femoral placement relative to the anatomical neck, poor exposure, incomplete seating of the femoral implant), and post-operative factors.[42-44]

Osteonecrosis
Osteonecrosis has been a common histological finding in failed resurfaced hips. Extensive involvement has been noted in femoral heads that failed by fracture, thus the role of osteonecrosis in the causation of these fractures has been questioned.[45]

Osteolysis
In hip resurfacing, osteolysis, or active bone resorption, arises from an inflammatory reaction to wear debris, primarily to polyethylene particulates, but also to metal debris. Osteolysis has been reported as a common complication in hip arthroplasty and a major cause of component loosening and failure.[46]

Prosthetic loosening
Aseptic loosening of the implant over time is a potential complication of hip resurfacing and is most likely related to inadequate initial fixation of the femoral component. Wear failure of the underlying cement-bone interface is another possible cause of loosening.[47]

Heterotopic ossification
Heterotopic ossification has been noted around total hip arthoplasty in numerous studies. Since hip resurfacing may require a somewhat larger incision and exposure than that of total hip arthroplasty, the rate of heterotopic ossification and its effect on function following resurfacing remains a concern.[48]

Metal wear debris
Concerns have been raised regarding the effect of metal wear debris from the metal-on-metal (MoM) bearing surfaces. The main metals used in MoM bearings are cobalt (Co) and chromium (Cr), however, the long-term biological consequences of the exposure to the particles and ions remain largely unknown and are poorly defined. Possible adverse consequences include local soft tissue toxicity, hypersensitivity reactions, bone loss, and carcinogenesis, as well as possible chromosomal aberrations and the risk of passing on those genetic abnormalities.[49] Compared to traditional metal-on-polyethylene (MoP) bearings, published reports on second generation MoM bearings have consistently revealed higher levels of serum Co and Cr when compared to preoperative values[50-54] and there is evidence to suggest that Co and Cr levels are influenced by factors such as the type, design, size, and positioning of the prosthesis.[49] Osteolysis has also been correlated with wear rate[55] and has

been reported in association with MoM total hip arthroplasty in a small number of cases.[49] Though the wear rate for MoM bearings has been reported to be substantially lower than for MoP bearing, the number of particles generated can be up to 500 times greater.[56]

Lack of long term follow-up
The current implants have only been used for about 10 years so the only data available is on short-term (1-5 years) and mid-term (5-10 years) follow-up. Data on longer-term (≥ 10 years) is needed.

2.7. Implant designs

The majority of total HR systems available today employ hybrid fixation, which includes cementless press-fit fixation of the acetabular component and cemented fixation of the femoral component. Although a few femoral components are designed for cementless applications, the majority of currently available femoral component designs employ cement as the means of implant fixation.[57]

Cemented femoral implants
There is variability in the amount and distribution of bone cement employed into the dome portion of the femoral implant. In one study, the proportion of retrieved femoral-head sections that were filled with cement ranged from 11% to 89%.[42] Another study reported that greater amounts of cement were measured in loosened femoral heads (51% versus 36% of fractured heads and 40% of nonfemoral failures), indicating cement volume as a possible cause for implant failure.[41] Excessive cementing may also lead to decreased bone-loading and thus increased stress-shielding in the proximal portion of the femur[58] and also may promote thermal necrosis.[42]

Cementless femoral implants
Cementless femoral design raises concerns regarding thermal necrosis from cementing, cement penetration, and controlling uniform mantle thickness. Excellent survival rates and significant improvements in pain and function at 2 years follow-up have been shown with a hydroxyapatite-coated (cementless) femoral design.[59] Many different aspects related to this type of design will need to be considered and addressed but as contemporary implants continue to mature, there is definite potential for cementless fixation in hip resurfacing.

2.8. Common hip resurfacing devices

Two total HR devices are approved currently by the FDA:
1. Birmingham Hip Resurfacing (BHR) System (MoM, high carbon, cobalt-chromium alloy; acetabular component coated with hydroxyapatite) – Approved by the FDA in May 2006. Manufactured by Smith & Nephew, Inc., Memphis TN

2. Cormet Hip Resurfacing System (MoM, cobalt chromium alloy; acetabular component has a bi-coating of plasma sprayed titanium and hydroxyapatite) - Approved by the FDA in July 2007. Corin USA, Tampa FL.
One device is undergoing 510K investigation, Conserve Plus (Wright Medical Technology) (MoM). Other common devices are found in Table 6.

Table 6. Common current devices

Device name	Company	FDA-approved	Where used
ASR	Depuy (J & J)	No	Canada, Europe, India, Australia
Birmingham	Smith and Nephew	Yes (May 06)	Globally
Conserve Plus	Wright Medical Technology	No (510-k investigational)	Europe and Asia
Cormet	Styker/Corin Medical	Yes (Jul 07)	USA
Durom	Zimmer	No	UK, North America outside USA

ASR: Articular Surface Replacement

2.9. Operative approach

Briefly, an incision is made in the side of thigh, allowing the surgeon to see both the femoral head and the acetabulum. The femoral head is dislocated out of the socket and its surface smoothed down and shaped so that the new metal cap will fit snugly on top of the bone. The femoral cap is then cemented into place; a small peg is also inserted down into the bone. The hip socket may remain unchanged but more often it is replaced by a thin metal cup. Friction holds the metal liner in place until bone grows into the holes in the surface and attaches the metal to the bone.

Total HR is substantially more technically demanding than a standard hip replacement and the optimal operative approach is controversial. Most current generation metal-on-metal hip resurfacing procedures are performed using a posterior approach. This approach employs circumferential capsulotomy at the acetabulum, which maximizes acetabular exposure and minimizes vascular insult. In addition, the short external rotator muscles are released in order to access the femoral head, which sacrifice the primary blood supply to the femoral head (the ascending branch of the medial circumflex artery). While the posterior approach provides excellent exposure, preserves the hip abductor muscles, keeps the joint capsule intact, and is easily reproducible, the possible intraoperative restriction of blood flow to the femoral head can lead to osteonecrosis and femoral nerve palsy.[47]

The lateral and anterior approaches are also performed in the US, but are much less common. The lateral approach provides very good exposure and also preserves femoral head blood flow but can be associated with abductor muscle weakness. The anterior approach does preserve blood flow to the femoral head but makes visualization of the socket difficult. This technique is also unfamiliar to may surgeons and may require a specialized surgical table.[47]

Without prior femoral neck resection, exposure of the acetabulum can present technical challenges. Accurate placement of a guide pin in the femoral neck is necessary to avoid varus positioning of the component and notching of the femoral neck. The use of a computer navigation system has been suggested as a possible way to improve the performance and accuracy of procedure. There is a high learning curve associated with total hip resurfacing[60] and surgeons are strongly encouraged to undergo additional training to properly prepare themselves for the technical challenges.

2.10. Clinical Guidelines

2.10.1. National Guideline Clearinghouse

No clinical guidelines related to hip resurfacing procedures were found when the NGC database was searched. Additional searching of the American Academy of Orthopaedic Surgeon's (AAOS) web site did not yield any guidelines specific to hip resurfacing.

2.10.2. National Institute for Health and Clinical Excellence

The National Institute for Health and Clinical Excellence (NICE), (which provides guidance on health technologies and clinical practice for the National Health Service in England and Wales) concluded in 2005 that "metal-on-metal (MoM) hip resurfacing arthroplasty is recommended as one option for people with advanced hip disease who would otherwise receive and are likely to outlive a conventional primary total hip replacement." Although there is sufficient short-term evidence to conclude that MoM hip resurfacing can be as effective as total hip replacement (THR) in patients less than 55 years, NICE acknowledges that there are no randomized controlled trials comparing MoM hip resurfacing arthroplasty with conventional THA. There are also no long-term (>10 years) observational data on the outcomes associated with MoM hip resurfacing devices.

2.11. Previous Systematic Reviews/Technology Assessments

Previously conducted reviews and assessments have not reached definitive conclusions regarding the safety and efficacy of hip resurfacing procedures. There is limited long-term data available. Table 7 summarizes the previous assessments.

Table 7. Overview of previous technology assessments of hip resurfacing.

Assessment (year)	Lit search dates	Prosthesis evaluated	Evidence base available*†	Critical Appraisal‡	Comments	Primary Conclusions
California Technology Assessment Forum (2007) *Metal-on-Metal total hip resurfacing as an alternative to total hip arthroplasty*	through 2007	BHR, Cormet 2000	• No RCTs comparing FDA-approved BHR and Cormet 2000 • 2 RCTs (90% f/u, 12 months); N = 234; compared earlier MoM non-FDA-approved devices • 7 case series (f/u NR; N = 1,150)	yes	Assessed only FDA-approved devices.	**Efficacy**: Because no RCTs with FDA-approved devices are available, MoM hip resurfacing has not been shown to improve health outcomes in an investigational setting. **Safety**: A national review of femoral neck fractures associated with BHR report an incidence of 1.46%. Chronic exposure to metal ions is a concern. **Economic**: not addressed
BlueCross BlueShield Technology Evaluation Center Assessment (2007) *Metal-on-metal total hip resurfacing*	through 01/2007	BHR, Cormet, Conserve plus	• 1 RCT (%f/u NR, 12 months); N = 210 • 12 case series (% f/u NR, 3 years); N = 2,076	no		**Efficacy**: Hip resurfacing represents a safe and effective means to defer a first THA in properly selected patients who require a total hip arthroplasty. **Safety**: See **Efficacy**. **Economic**: Not addressed
Alberta Bone and Joint Health Institute (2006) *Metal-on-metal hip resurfacing for young, active adults with degenerative hip disease.*	through 2006	BHR, Conserve Plus, Cormet 2000, ReCap, Durum, ASR	• 6 SR/HTA • 2 RCTs (90% f/u, 12 months); N = 234; compared THA with HRS • 13 Case series/case control (95% f/u, 5 years); N = 2,209	no	Most studies had limited follow-up (less than 2 years), thus it was difficult to assess long-term device performance.	**Efficacy**: Based on two RCTs, HRS and THA confer similar satisfaction rates in younger patients, but HRS may offer better functional performance. **Safety**: No significant differences were found for revision rates due to complications, although long-term (> 5 years) safety is unknown. **Economic**: MoM hip resurfacing could be more cost-effective than THA after year 1, but long-term revision rates are unknown.

Table 7. Overview of previous technology assessments of hip resurfacing.

Assessment (year)	Lit search dates	Prosthesis evaluated	Evidence base available*†	Critical Appraisal‡	Comments	Primary Conclusions
Ontario Health Technology Assessment Series (2006) *Metal-on-metal total hip resurfacing arthroplasty*	Jan 1, 1997 through October 27, 2005	BHR, Conserve Plus, Cormet 2000, ReCap, Durum, ASR	• 1 RCT (100% f/u, 8.5 years); N = 24; compared THA with HRS • 8 Case series (96% f/u, 4 years); N = 1,539	yes	RCT not used for assessment because newer generation of implants are now used.	**Efficacy**: MoM hip resurfacing arthroplasty has been shown to be effective as tested in younger patients. However, there are no RCTs that compare MoM hip resurfacing with THR. **Safety**: Concern remains on the potential adverse effects of metal ions. **Economic**: MoM hip resurfacing is more cost effective compared with watchful waiting followed by THR. MoM hip resurfacing is not more cost effective when compared directly with THR.
Agency for Healthcare Research and Quality (2006) *Horizon Scan on hip replacement surgery*	through 2006	N/A	• 3 RCTs ongoing (still in recruitment or data collection stage)	no	Review focused on THR	**Efficacy**: Data on effectiveness of hip resurfacing are limited, but this conclusion was based on older literature. **Safety**: Most common complication is periprosthetic fracture of the femoral head. **Economic**: Not addressed.
Center for Clinical Effectiveness (2002) *Hip resurfacing in patients with osteoarthritis*	through 11/2002	Birmingham hip resurfacing system (BHR)	• 1 SR (case series only included for review; f/u inclusion criteria > 5 years; N = NR; compared BHR with THR) • 1 HTA	yes	More recent literature available now	**Efficacy**: MoM resurfacing of the hip may be a viable and bone-conserving option for adults who are likely to outlive THR. **Safety**: Short-term revision rates were comparable between MoM and THR. **Economic**: THA was calculated to be more

Table 7. Overview of previous technology assessments of hip resurfacing.

Assessment (year)	Lit search dates	Prosthesis evaluated	Evidence base available[*†]	Critical Appraisal[‡]	Comments	Primary Conclusions
						cost-effective than MoM, but there was a lack of long-term data on health outcomes and revision rates.
The Canadian Coordinating Office for Health Technology Assessment (2003) *Metal-on-Metal hip resurfacing*	through 02/2003	BHR, CONSERVE Femoral Surface Replacement device	• 7 HTAs	no	More recent literature available now	**Efficacy**: MoM hip resurfacing was recommended as one option for active, younger patients with advanced hip disease. **Safety**: Patient selection is important for prosthesis viability. **Economic**: Need for cost-benefit analysis was stated.
Alberta Heritage Foundation for Medical Research (2002) *Metal-on-Metal hip resurfacing for young, active adults with degenerative hip disease*	through 2002	BHR, CONSERVE plus, Cormet 2000	• 5 HTAs • 0 RCTs • Several case series (not evaluated)	no	More recent literature available now	**Efficacy**: MoM resurfacing may be a viable and bone-conserving option for adults with degenerative hip disease who are likely to outlive THR. **Safety**: Concern of the toxicity of the metals over time if shed into the body. Patient selection is important; good bone stock required. **Economic**: Costs for BHR and THA were comparable in the UK.

BHR: Birmingham Hip Resurfacing System.
THR: Total hip arthroplasty.
HRS: Hip Resurfacing.
MoM: Metal-on-Metal.
NR: Not Reported.
N/A: Not Available.
[*]Percent follow-ups are weighted based on sample size, and were calculated using the N reported in the assessment. Percent follow-ups were not given for all RCTs or case series. Mean time to follow-up is reported here.
[†]N reflects numbers before loss to follow-up.
[‡]Critical appraisal refers to formal evaluation of individual study quality using criteria such as the Jadad or GRADE methods of scoring and the determination of overall strength of evidence.

2.12. Medicare and Representative Private Insurer Coverage Policies

Coverage policies are consistent for hip resurfacing for CMS and selected bell-weather payers. The payers will provide coverage for hip resurfacing as long as an FDA-approved device is used and certain patient conditions are met. Table 8 provides an overview of policy decisions.

- Medicare
  The Centers for Medicare and Medicaid Services (CMS) will consider total hip resurfacing medically necessary in select patients requiring primary hip resurfacing due to the following conditions:
    - Non-inflammatory arthritis (degenerative joint disease) such as osteoarthritis, traumatic arthritis, avascular necrosis, or dysplasia/developmental dislocation of the hip.
    - Inflammatory arthritis, such as rheumatoid arthritis.

  CMS listed hip resurfacing as a potential National Coverage Determination (NCD) topic in 2008 on their website, http://www.cms.hhs.gov/mcd/ncpc_view_document.asp?id=19

  CMS's assessments include information from three FDA premarket approval letters for the Birmingham Hip Resurfacing system (BHR). Safety and Effectiveness was assessed using the Smith and Nephew, Inc., Orthopaedics division (September 2005). Other information for assessment was derived from case-series reports.

- Aetna
  Aetna considers metal-on-metal hip resurfacing a medically necessary alternative to total hip arthroplasty for physically active members with osteoarthritis of the hip, or osteonecrosis of the femoral head.

- Blue Cross/Blue Shield
  Total hip resurfacing arthroplasty with an FDA-approved device may be considered medically necessary for patients with degenerative hip joint disease, or severe arthritis, or rheumatoid arthritis, or advanced avascular necrosis of the hip and meet all of the following criteria:
    - Skeletally mature and 55 years of age or less.
    - Patient with BMI of 39 or less.
    - Have failed conservative management, and would otherwise require total hip arthroplasty surgery.

- Cigna
  Cigna covers total hip resurfacing arthroplasty as medically necessary as an alternative to total hip arthroplasty when all of the following criteria are met:
    - Age less than 65 years.
    - Diagnosis of osteoarthritis or inflammatory arthritis.

- Individual has failed nonsurgical management and is a candidate for total hip arthroplasty.

- Harvard Pilgrim
  Total hip resurfacing arthroplasty is covered with FDA-approved devices for the treatment of hip disease in patients who are younger than age 55 and who meet the following criteria:
  - Have chronic, persistent pain and/or disability.
  - Are otherwise fit and active.
  - Have normal proximal femoral bone geometry and bone quality.
  - Would otherwise receive a conventional primary total hip arthroplasty, but are likely to live longer than a conventional THA is expected to last.

Table 8. Overview of payer technology assessments and policies for hip resurfacing.

Payer (year)	Lit search dates	Prosthesis evaluated	Evidence base available*†	Policy	Rationale/comments
Centers for Medicare and Medicaid Services (2007)	through 2007	BHR	(a) FDA premarket approval letters (2006-2007) (b) 6 case series (f/u, N NR)	The Centers for Medicare and Medicaid Services (CMS) will deem total hip resurfacing medically necessary in select patients requiring primary hip resurfacing due to the following conditions: o Non-inflammatory arthritis (degenerative joint disease) such as osteoarthritis, traumatic arthritis, avascular necrosis, or dysplasia/developmental dislocation of the hip. o Inflammatory arthritis, such as rheumatoid arthritis.	(a) Policy only valid for FDA-approved devices. (b) CPT codes if criteria are met: 00.85, 00.86, 00.87
Aetna Clinical Policy Bulletin (2008)	through 2007	BHR	• 3 HTAs • 3 case series (% f/u NR, 5 years) N = 506 • Listed 40 other references policy is based on.	Aetna considers metal-on-metal hip resurfacing a medically necessary alternative to total hip arthroplasty for physically active members with osteoarthritis of the hip, or osteonecrosis of the femoral head.	• No rationale for policy stated • CPT codes if selection criteria is met: 27125
BCBS Medical Policy (2006)	N/A	BHR	• 2 case series (f/u NR, N = NR)	Total hip resurfacing arthroplasty with an FDA-approved device may be considered medically necessary for patients with degenerative hip joint disease, or severe arthritis, or rheumatoid arthritis, or advanced avascular necrosis of the hip and meet all of the following criteria: o Skeletally mature and 55 years of age or less, and o Patient with BMI of 39 or less, and o Have failed conservative	▪ Hemi hip resurfacing of the femoral head is an established procedure for patients with osteonecrosis of the femoral head. ▪ While long-term studies are needed to address the use of total hip resurfacing in most patient populations, there is adequate evidence to supports the use of this procedure in patients at low risk for failure of the procedure.

Table 8. Overview of payer technology assessments and policies for hip resurfacing.

Payer (year)	Lit search dates	Prosthesis evaluated	Evidence base available*†	Policy	Rationale/comments
				management, and would otherwise require total hip arthroplasty surgery.	
Cigna Medical Coverage Policy (2009)	through 2008	BHR, Cormet (FDA-approved)	• 10 case series (% f/u NR); N = 403.	Cigna covers total hip resurfacing arthroplasty as medically necessary as an alternative to total hip arthroplasty when all of the following criteria are met: o Age less than 65 years o Diagnosis of osteoarthritis or inflammatory arthritis o Individual has failed nonsurgical management and is a candidate for total hip arthroplasty.	▪ Policy is designed for treatment of osteonecrosis of the femoral head when there is failure, contraindication or intolerance of nonsurgical management. ▪ CPT codes if selection criteria are met: 27125, 27130, 27299
Harvard Pilgrim (2008)	through 2007	BHR, Cormet	• 1 HTA • 4 case series (f/u NR, N = NR)	Total hip resurfacing arthroplasty is covered with FDA-approved devices for the treatment of hip disease in patients who are younger than age 55 and who meet the following criteria: o Have chronic, persistent pain and/or disability o Are otherwise fit and active Have normal proximal femoral bone geometry and bone quality, and would otherwise receive a conventional primary total hip arthroplasty, but are likely to live longer than a conventional THA is expected to last.	▪ Used BCBS, United, Cigna, and Aetna as benchmarks for policy decision. ▪ CPT codes if selection criteria are met: S2118, 27299, 27130
Washington State Payers					
Group Health Cooperative (2007)	through 2006	BHR, Cormet	• 2 SR (no information provided) • 2 RCTs used for critical appraisal (% f/u NR, 8.5 years – one study only); N = 128	Group Health members are covered when all of the following criteria are met: o The patient is 55 years of age or younger. o The device is FDA-approved. o The patient has been diagnosed with arthritis of the hip. o The patient would otherwise require a total hip arthroplasty.	• No rationale for policy stated

BHR: Birmingham Hip Resurfacing System.

NR: Not Reported.
*Formal critical appraisals were not reported in any of the payer HTAs except Group Health. Percent follow-ups were not given for RCTs or case series. Mean time to follow-up is reported here.
†N reflects numbers before loss to follow-up.

3. The Evidence

3.1. Methods of the Systematic Literature Review

3.1.1. Inclusion/exclusion

Inclusion and exclusion criteria are summarized in Table 9.

- *Population.* Studies of adults who underwent primary total HR for arthritis (non-inflammatory or inflammatory) developmental dysplasia, osteonecrosis were included.
- *Intervention.* Included studies that evaluated total HR using modern commercially available devices designed for hybrid (i.e. using cementless acetabular fixation) resurfacing: FDA-approved or unapproved devices in Phase III trials with ≥ 1 year of follow-up data in a peer-reviewed journal. Studies reporting on non-hybrid or hemi-resurfacing or minimally invasive surgery were excluded.
- *Comparator.* Included studies that compared hybrid total HR to primary THA: FDA-approved or unapproved devices in Phase III trials with ≥ 1 year of follow-up data in a peer-reviewed journal. Studies that reported on revision THA were excluded.
- *Outcomes.* Eligible studies reported on at least one of the following outcomes: physical function/disability (clinical success (e.g., Harris Hip Score, pain, activity, and motion), revision, or complications (including femoral neck fracture, femoral head collapse, avascular necrosis, dislocation, osteolysis, device migration or loosening, heterotopic ossification, impingement, infection or radiolucencies).
- *Study design.* Eligible studies compared total HR with THA utilizing a randomized or cohort study design. In order to provide additional context regarding key questions 2 and 3, registry studies as well as studies with historical/nonconcurrent controls and/or summaries of case series with > 5 years follow-up and >10 patients were included. Formal economic analyses published in peer-reviewed journals were eligible for inclusion to help answer key question 4 as were cost data reported in other systematic reviews or technology assessments.

Table 9. Summary of inclusion and exclusion criteria

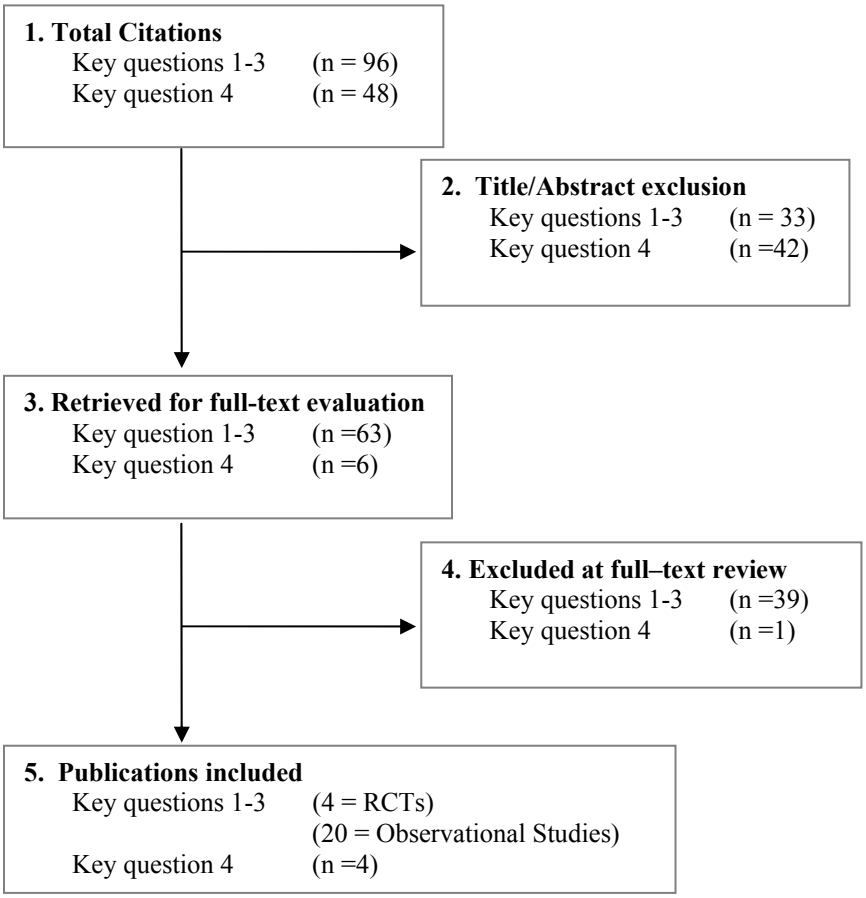
Study Component	Inclusion	Exclusion
Population	• Patients undergoing primary total HR	• Patients with contraindications to receive total HR
Intervention	• Total HR with a modern commercially available device designed for hybrid resurfacing: FDA-approved or un-approved devices in Phase III trials with ≥ 1 year of follow-up data in a peer-reviewed journal	• Non-hybrid or hemi resurfacing or use of minimally invasive surgery
Comparator	• Primary THA	• Revision THA
Outcomes	Studies must report on at least one of the following: • Physical function/disability (clinical success, pain, activity, or motion) • Revision • Complications (e.g., femoral neck fracture, infection, avascular necrosis, dislocation, osteolysis, device migration or loosening, heterotopic ossification, infection, and others) The following secondary outcomes are reported if presented with studies meeting the above criteria: • Quality of life (SF-36, SF-12, or EQ-5D)	• Non-clinical outcomes
Study Design	• Randomized controlled trials (RCTs) and comparative studies with concurrent controls were considered for question 1. • RCTs and comparative studies with concurrent controls were sought initially for questions 2 and 3. • In order to provide additional context regarding questions 2 and 3, registry studies as well as studies with historical/non-concurrent controls and/or summaries of case series with follow-up ≥ 5 years were obtained and very briefly summarized. • For question 4, formal economic analyses (e.g., cost-utility study) were sought. In the absence of formal economic analyses, cost data reported in other systematic reviews or technology assessments were briefly summarized.	• For question 1, studies other than RCTs or comparative studies with concurrent controls were excluded. • Case reports • Case series with fewer than 10 patients
Publication	• Studies published in English in peer reviewed journals	• Abstracts, editorials, letters • Duplicate publications of the same study which do not report on different outcomes • Single-site reports from multicenter trials • White papers or narrative reviews • Articles identified as preliminary reports when results are published in later versions

3.1.2. Data sources and search strategy

The clinical studies included in this report were identified using the algorithm shown in Appendix A. The search took place in four stages. The first stage of the study selection process consisted of a comprehensive literature search using electronic means and hand searching. We then screened all possible relevant articles using titles and abstracts in stage two. This was done by two individuals independently. Those articles that met a set of *a priori* retrieval criteria based on the criteria above were included. Any disagreement between screeners that were unresolved resulted in the article being included for the next stage. Stage three involved retrieval of the full text articles remaining. The final stage of the study selection algorithm consisted of the selection of those studies using a set of a priori inclusion criteria, again, by two independent investigators. Those articles selected form the evidence base for this report.

Electronic databases searched included PubMed, EMBASE, CINAHL, ClinicalTrials.gov, CRISP, HSTAT, *The Cochrane Library*, EconLIT, PsychINFO, AHRQ, and INAHTA for eligible studies, including health technology assessments (HTAs), systematic reviews, primary studies and FDA reports. Reference lists of all eligible studies were also searched. The search strategies used for PubMed and EMBASE, are shown in Appendix B. Figure 2 shows a flow chart of the results of all searches for included primary studies. Articles excluded at full-text review are listed in Appendix

Figure 2. Flow chart showing results of literature search



3.1.3. Data extraction

Reviewers extracted the following data from the included clinical studies: study population characteristics, study type, study period, patient demographics and preoperative diagnoses, study interventions, follow-up time, study outcomes (functional and clinical, motion, radiographic), adverse events (revision, femoral neck fracture, avascular necrosis, osteolysis, heterotopic ossification, device loosening or migration, elevated serum (etc.) ion concentrations, death), and other complications (intraoperative cracks or notching, radiographic lucency, infection, myocardial infarction, deep vein thrombosis, pulmonary embolism, etc.). An attempt was made to reconcile conflicting information among multiple reports presenting the same data. For economic studies, data related to sources used, economic parameters and perspectives, results, and sensitivity analyses were abstracted.

3.1.4. Study quality assessment: Level of evidence (LoE) evaluation

The method used by Spectrum Research, Inc. (SRI) for assessing the quality of evidence of individual studies as well as the overall quality of evidence incorporates aspects of the rating scheme developed by the Oxford Centre for Evidence-based Medicine,[61] precepts outlined by

the Grades of Recommendation Assessment, Development and Evaluation (GRADE) Working Group,[62] and recommendations made by the Agency for Healthcare Research and Quality (AHRQ).[63]

Details of the Level of Evidence (LoE) methodology are found in Appendix D. Each clinical/human study chosen for inclusion was given a LoE rating based on the quality criteria listed in Appendix D. Standardized abstraction guidelines were used to determine the LoE for each study included in this assessment.

3.2. Quality of Literature Available

3.2.1. Quality of studies retained

We initially found 144 citations using the search strategy in Appendix B. For Key Question 1 we identified 25 reports that compared total HR with THA (24 comparative studies and 1 meta-analysis). From among these, four RCTs (two studies reporting different outcomes for the same population) and nine cohort studies (one prospective and eight retrospective) were identified. The RCTs are all graded as LoE II, and the cohort studies all received LoE grades of III.

For Key Question 2 on safety, we included in addition to the studies cited in the preceding paragraph, three registry reports, two cohort studies and six case series. The registry reports received a LoE grade of II, the cohort studies a level III, while all the case series received the LoE grade of IV.

To address outcomes following total HR in special populations (Key Question 3), we included three registry reports and six retrospective cohort studies. The registry reports received a LoE grade of II while three of the cohort studies received the LoE grade of II and the remaining three received the LoE grade of III.

3.2.2. Critical Appraisal (APPENDIX E)

Randomized controlled trials

Garbuz (2009)
Garbuz et al.[64] published the results of a randomized controlled trial (RCT) in which 107 patients with an unreported number of hips were randomized to undergo either total HR with metal-on-metal (MoM) Durom components (48 patients underwent surgery) or MoM total hip arthroplasty (THA) with a large femoral head. Both groups received the same acetabular cup; all surgeries were performed using the posterior approach. The preoperative diagnosis was not reported; patients were included if they were considered suitable for total HR by the operating surgeon and were between 19 and 70 years of age. Mean patient age was 51.8 years, and 89.4% of patients were male. Although randomization was achieved using permuted blocks of two and four, patient treatment

groups were placed in sealed but not necessarily opaque envelopes that were opened the day prior to surgery. Patients were blinded to their treatment group, though the authors did not discuss whether the patients remained blinded throughout the follow-up period. There was no indication of intention-to-treat. The objectives of this study were to evaluate whether there were differences between the groups in quality of life outcomes as well as serum metal ion concentrations; interestingly, although the PAT-5D (Paper Adaptive Test in 5 Domains of Quality of Life in Arthritis Questionnaire) was the primary outcome, the outcome scores were not reported. A complete follow-up rate of only 68% was reported; most outcomes were reported at one or two years follow-up. The institution of one or more of the authors for this study received funding from Zimmer, Inc. (Warsaw, IN). This study received a level of evidence (LoE) grade of II.

Lavigne (2009)

Lavigne et al.[65] conducted a RCT in which 48 patients with 48 hips were randomized to undergo total HR with MoM Durom prostheses (24 patients) or THA with a MoM large femoral head (24 patients). The acetabular component was identical in both groups; all surgeries were performed using the posterior approach. The majority of patients had osteoarthritis (77.1%); other diagnoses included developmental dysplasia (6.3%), protrusion acetabuli (4.2%), posttraumatic osteoarthritis (2.1%), avascular necrosis of the femoral head (6.3%), postseptic arthritis (2.1%), and rheumatoid arthritis (2.1%). Mean patient age was 49.7 years, which ranged from 33 to 63 years, and 60.4% of patients were male. Patients were randomly assigned to a treatment group using a random number generator; corresponding sealed and numbered opaque envelopes contained the type of procedure and were opened by the surgeon the day of surgery. Unlike most surgical studies, most patients (85%) remained blinded to the treatment until one year postoperatively. Although one of the primary objectives of the study was to determine whether total HR patients had better walking speeds at one year, there was a significant difference in the preoperative normal walking speeds (and step length) between treatment groups that was not controlled for using stratification or multivariate analysis. Otherwise, baseline patient demographics (age, gender, body mass index, and diagnosis) were similar between treatment groups. The complete follow-up rate was 87.5%; the mean follow-up was 14 months, though most outcomes were reported at one year. With respect to conflict of interest for this study, one or more of the authors received funding from Zimmer, Inc. (Warsaw, IN). This study received a LoE grade of II (downgraded from a score of I due to inadequate sample size and lack of control adjusting for baseline walking speed differences).

Vendittoli (2006) and Rama (2009)

Vendittoli et al.[66] performed a RCT in which 191 patients with 210 hips were randomized to receive either total HR with MoM Durom components (107 hips) or MoM THA (103 hips); the posterior surgical approach was used for all patients. Most hips had a preoperative diagnosis of osteoarthritis (75.7%); other diagnoses included Perthes (2.9%), dysplasia (8.1%), osteonecrosis (2.4%), inflammatory arthritis (7.6%), postseptic arthritis (1.0%), and 2.4% of hips were posttrauma. Mean patient age was 49.8 years, which ranged from 23 to 65 years, and 65.2% of patients were male. Although randomization

was performed with statistical software that used a block randomization table, the authors did not disclose who had access to the program and when. Intention-to-treat was not used, as one patient was excluded following intraoperative conversion from total HR to THA. No mention of independent or blind assessment was made. Although most demographics were similar between groups, the THA group had a significantly higher body mass index (BMI) (29.6 versus 27.2 for the total HR group, (P = .01)) that was not controlled for. A complete follow-up rate of 97.6% was achieved, with outcomes reported at one year postoperation. Due to a number of limitations in study design, this RCT received a LoE grade of II.

Rama et al.;[67] reported heterotopic ossification rates and risk factors for the same population reported in Vendittoli.[66] Although radiographs were used to evaluate HO, the authors did not disclose whether they were assessed by independent or blinded evaluators. Rama had a complete follow-up rate of 95.2%; outcomes were reported at one year or later. This study also received a LoE grade of II.

Prospective cohort studies

Fowble (2009)
Fowble et al.[68]reported outcomes from a prospective cohort study in which 85 patients with 94 hips received either total HR using Conserve Plus prostheses under an FDA investigational device exemption (IDE) (50 patients with 50 hips) or THA with either a cross-linked polyethylene bearing (Marathon) (30 hips) or metal bearing (Ultamet) (14 hips); all surgeries were performed via the posterolateral approach. All patients who underwent total HR were self-referred, which is a potential source of bias. Preoperative diagnoses included osteoarthritis (93.6% of hips) and osteonecrosis (4.3% of hips). The mean patient age was 49.7 years, which ranged from 27 to 75 years, and 53.4% of patients were male. Patients were followed for a mean of 34.7 months, and a complete follow-up rate of 94.1% was obtained. One total HR patient had revision surgery and was excluded from all clinical outcomes, which may be a source of bias. Independent or blind assessment was not reported. Of note, there were several significant differences between the groups that were not controlled for, including (but not limited to): gender (62% male (total HR); 41% male (THA), (P = .03)), mean age (46 (total HR); 55 (THA), (P = .0001)), BMI (27.3 (total HR); 31.3 (THA), (P = .001)), Harris hip score (HHS) (P = .005), and UCLA activity score (P = .02). It is noted that financial support for this study was provided by Wright Medical Technology and the Los Angeles Orthopaedic Hospital Foundation. One investigator has a financial interest in the total hip replacement prostheses used in this research study (DePuy Pinnacle™, Summit™, and Ultamet™). The study received a LoE grade of III.

Retrospective cohort studies

Li (2009)
Li et al.[69] published the results of a retrospective cohort study in which 49 ankylosing spondylistis patients with 80 hips received either total HR (Durom) (24 consecutive

patients with 39 hips) or ceramic-on-ceramic cementless THA (Secur-Fit HA) (25 patients with 41 hips). All surgeries were performed via the posterolateral approach. The mean patient age was 30.9 years, and ranged from 20 to 47 years, and 81.2% of patients were male. A complete follow-up rate of 100% was achieved, but the length of follow-up was not reported. One total HR patient who underwent revision due to a femoral neck fracture was excluded from all clinical outcomes, which may skew the results for this group. This study received a LoE grade of III due to inadequate sample size, lack of independent or blind assessment, and for not controlling for possible confounding.

Li (2008)
Li et al.[70] evaluated outcomes of 42 patients (52 hips) with developmental dysplagia in a retrospective cohort study. MoM total HR (Durom) was performed in 21 consecutive patients with 26 hips, while the same number of matched patients (and hips) received ceramic-on-ceramic (Secur-Fit HTA) THA; all procedures were done via the posterolateral approach. Mean patient age was 47.4 years (range of 37 to 64 years), and 71.4% of patients were female. The complete follow-up rate at a mean of 26.5 months was 100%. A LoE grade of III was given for the same reasons as Li (2009).

Mont (2009)
Mont et al.[10]prospectively followed both total HR and THA patients before performing retrospective matching. Fifty-four patients with 54 hips in the total HR group received the Conserve Plus prosthesis (as part of an FDA IDE study) via the anterolateral approach, while the same number of matched patients and hips underwent THA (approach not disclosed). Patients in the total HR group came to the authors' institution specifically to request the procedure, creating a potential source of bias between the groups. The mean patient age was 55 years, and ranged from 35 to 79 years, and two-thirds of patients were male. Surgical indications included osteoarthritis, osteonecrosis, or hip dysplagia (percentages were not reported). Complete follow-up of 92.6% of patients was achieved with a mean follow-up of 39 months. Although patients were matched, total HR patients had a significantly higher mean preoperative activity score than their THA counterparts ($P = .01$); another limitation was that the two cohorts received different postoperative rehabilitative treatments. The primary author for this study acknowledges that he is a consultant for and has received funding from Stryker Orthopaedics (Mahwah, NJ) and Wright Medical Technology (Arlington, TN). This study was assigned a LoE grade of III.

Pattyn (2008)
Pattyn et al.[71] reported outcomes of 440 patients (number of hips not reported) who underwent either Birmingham MoM total HR via the posterolateral approach (250 consecutive patients) or ceramic-on-ceramic THA (Ancafit; uncemented) via the Harding lateral (73.7%) or posterolateral (26.3%) approach (190 patients). Patients had a mean age of 48.3 years, with a range of 14 to 78 years, and 63.0% were male. Preoperative diagnoses included osteoarthritis (70.1%), avascular necrosis (17.0%), rheumatoid arthritis (4.5%), and trauma (1.9%). The complete follow-up rate was 99.5%, and follow-up ranged from 36 to 72 months. One limitation of the study was notable

differences in patient demographics (mean age, gender, and surgical indications) between groups that were not controlled for by stratified or multivariate analysis. In addition, outcomes were not assessed in a blinded or independent manner. The study was given a LoE grade of III.

Pollard (2006)
Pollard et al.[72] retrospectively reviewed 113 patients with 117 hips who underwent Birmingham total HR or hybrid THA via the posterior approach. In the total HR group, the first 63 Birmingham total HR hips treated by the senior author were included; however, the authors excluded three hips that underwent revision within the first postoperative year as well as six hips not available for follow-up. One additional hip in the total HR group was lost to follow-up, thus clinical outcomes for the total HR group were reported for 53 hips (complications were reported for 56 hips). The THA group was comprised of 54 matched hips and was selected from a 64-month period, however three patients were lost to follow-up, leaving 51 hips available for review. Overall, the complete follow-up rate included 88.5% of all patients. The primary surgical indication was osteoarthritis (75.9%); other indications included avascular necrosis (10.2%) and dysplasia (5.6%). Mean patient age was 50.1 years (range of 18 to 67 years), and 76.9% of patients were male. Outcomes did not appear to be assessed by an independent or blinded observer. Although patient demographics were relatively similar between groups, the authors did not report whether there were any statistically significant differences. Interestingly, the authors excluded from clinical outcomes three patients from the total HR group who underwent revision in the first year following the procedure; however one total HR patient who underwent revision surgery at 62 months and four THA patients with planned revisions were all included in the clinical outcomes, making the reported clinical outcomes difficult to interpret. This study was the only one with mid-term (as opposed to short-term) follow-up, as patients were followed for an average of 5.9 years (range: 3.5–10 years) (THA: 6.7 (3.5–10) years; total HR: 5.1 (4.3–5.9) years). This study received a LoE grade of III.

Stulberg (2008)
Stulberg et al.[46] evaluated 603 patients with as many hips in a retrospective cohort study. A total of 337 patients were enrolled as part of a randomized FDA IDE study for the Cormet 2000 total HR System. The THA group was comprised of 266 matched patients who received the ceramic-on-ceramic Osteonics ABC System as part of a nonrandomized IDE study of this device. One potential source of bias is that the THA group served as an historical control, treated between 1996 and 1998, while the total HR patients were enrolled between 2001 and 2003. Preoperative diagnoses were osteoarthritis (84.9%), osteonecrosis (14.5%), and rheumatoid arthritis (0.7%). The mean patient age was 51.5 years (range was not reported), but was significantly higher in the THA group (53.3 years versus 50.1 years in the total HR group), and 65.2% of patients were male. The authors did perform propensity analysis to assess the comparability of patient demographics and baseline HHS between the cohorts and found no differences that would affect a

conclusion of non-inferiority of total HR. A complete follow-up rate of 90.8% was achieved, and only patients with a minimum of 24 months of follow-up were included in clinical outcomes. Of note, 16 patients in the total HR group and three patients in the THA group were excluded from all clinical outcomes because they received revision surgery with the first two postoperative years, while an additional eight patients in the total HR group and two in the THA group underwent revision after 24 months and were included in all clinical outcomes. The exclusion of some revision patients from clinical outcomes makes these data more difficult to interpret. The authors acknowledge that one or more of them received outside funding or grants from Stryker Orthopaedics. In addition, one or more of the authors or a member of his or her immediate family received payments or other benefits, or a commitment or agreement to provide such benefits from a commercial entity (Corin, Tampa, Florida). This study was given a LoE grade of III.

Vail (2006)
Vail et al.[19] published the results of a retrospective review of 231 patients with 261 hips who underwent MoM total HR (Conserve Plus) as part of an FDA IDE study or metal-on-polyethylene THA. Because the authors excluded patients with less than two years follow-up, only 55 patients (57 hips) and 84 patients (93 hips) were included in the outcomes for the total HR and THA groups, respectively. Total HR prostheses were implanted via the posterior approach, while the approach used for the THA procedures was not disclosed. The most common indication for surgery (prior to loss to follow-up) was osteoarthritis; others included osteonecrosis, dysplasia, posttraumatic arthritis, and rheumatoid arthritis. Prior to loss to follow-up, the mean patient age was 53.2 years (range: 17 to 92 years), and 52.9% of patients were female. There were several significant differences between the groups in terms of demographics and preoperative scores; however these differences were controlled for with the use of multivariate analysis. Outcomes were not assessed in an independent or blinded manner. The complete follow-up rate was only 59.6%, and patients were followed for a mean of 36 months. Each author in this study certifies that he has or may receive payments or benefits from a commercial entity related to this work (Wright Medical Technology, Inc). This study received a LoE grade of III.

Zywiel (2009)
Zywiel[73] conducted a retrospective cohort study of 66 patients with 66 hips. The total HR group consisted of 33 patients who received the Conserve Plus prosthesis as part of an FDA IDE study; the THA group consisted of 33 matched patients who underwent THA with either metal-on- polyethylene or ceramic-on-polyethylene prostheses. The anterolateral approach was used in all total HR procedures, however the approach used for THA was not reported. Although the total HR group originally consisted of 54 consecutive hips, the authors were unable to match and hence excluded 21 of these hips, providing an opportunity for bias. Surgical indications were not disclosed. The mean patient age was 53 years, and ranged from 37 to 79 years, and 69.7% of patients were male. Because patients were closely matched, preoperative demographics, HHS, and activity scores were similar between cohorts. The primary outcomes were patient-

reported and considered as reliable evidence in a retrospective study. Patients were followed for a mean of 43.5 months, and the complete follow-up rate was not reported. A senior author is a consultant for Stryker Orthopedics and Wright Medical Technology. This study was given a LoE grade of III.

Australian Joint Replacement Registry (2008)[74]
In 1999, the Australian Commonwealth Department of Health and Ageing established a Joint Replacement Registry with a staged implementation. The Registry became national in 2002 and receives information from some 292 public and private hospitals that perform joint replacement. Data for the Registry are collected on Registry forms completed at the participating hospitals at the time of surgery. Data validation is by multilevel comparison to data provided by state and territory health departments. For some territories, individual level patient/procedure validation is performed. For the 2006/07 Registry data, the initial validation resulted in 93% of Registry records verified against health department data. Follow-up on unreported and unmatched records yielded an almost complete set of data relating to hip replacement in Australia. The Eighth Annual Report details the findings from the Registry through December 2007, and includes analyses on 125,004 total hip arthroplasties, and 10,623 hip resurfacings. (Reference) The primary outcome is time to first revision described using Kaplan-Meier survival estimates.

The National Joint Registry for England and Wales (NJR) (2008)[75]
The NJR was established in October 2002 and began collecting and studying data on hip replacement surgeries in April 2003. Data are provided to the NJR by the National Health Service (NHS) and independent healthcare providers throughout England and Wales. The analyses are based on data on primary hip replacement undertaken between 1st April 2003 and 30 September 2007 that are linked to an episode in the Hospital Episode Statistics (HES) database. Linkage was not available for independently funded hip procedures in the private sector (estimated at 15% of the total number of procedures in the NJR during the above time frame). The main outcome of interest is survival to revision of implants in primary hip replacement surgeries. The 5th Annual Report was used for this Technology Assessment.

The Swedish Hip Arthroplasty Register (2007)[1]
The Swedish Hip Arthroplasty Register has been in existence for approximately 30 years. In the last 10 years, the registry has included such variables as patient-reported outcomes, short term complications and 10-year survival. As of the 2007 report, 1041 hip resurfacing implants have been registered in the Swedish Hip Arthroplasty Registry with a mean follow-up time of 2.2 years (SD, 1.7 years). Three implants have been used primarily: the BHR, Durom and ASR.

4. Results

For key question 1, we identified a total of 13 studies, three of which used FDA-approved devices, four used an FDA-510k investigational device and were conducted as part of an FDA investigational device exemption (IDE) trial, and five used devices not regulated by the FDA. For key question 2, we identified six additional case series: four used FDA-approved devices, one used an FDA-approved device in only 75% of patients, and one used an FDA-510k investigational device. Three registry studies reported results from both approved and non-approved FDA devices. For key question 3, we identified six retrospective cohort studies: two used FDA-approved devices, and four used an FDA-510k investigational device, Table 10.

Table 10. FDA-approval status for devices used to answer key questions

	design	Device used	FDA-approved	LoE
Key Question 1				
Garbuz (2009)	RCT	Durom	No	II
Lavigne (2009)	RCT	Durom	No	II
Vendittoli (2006) Rama (2009)	RCT	Durom	No	II
Fowble (2009)	Prospective cohort	Conserve Plus	No* (510-k investigational)	III
Li (2009)	Retrospective cohort	Durom	No	III
Li (2008)	Retrospective cohort	Durom	No	III
Mont (2009)	Retrospective cohort	Conserve Plus	No* (510-k investigational)	III
Pattyn (2008)	Retrospective cohort	Birmingham	Yes	III
Pollard (2006)	Retrospective cohort	Birmingham	Yes	III
Stulberg (2008)	Retrospective cohort	Cormet	Yes†	III
Vail (2006)	Retrospective cohort	Conserve Plus	No* (510-k investigational)	III
Zywiel (2009)	Retrospective cohort	Conserve Plus	No* (510-k investigational)	III
Key Question 2				
Australia (2008)	Registry report	Various		II
Swedish (2007)	Registry report	Various		II
UK (2008)	Registry report	Various		II
Amstutz	Case series	Conserve Plus	No (510-k	IV

Table 10. FDA-approval status for devices used to answer key questions

	design	Device used	FDA-approved	LoE
(2008)			investigational)	
McBryde/Revell) (2008)	Case series	Birmingham	Yes	IV
McMinn (2008)	Case series	Birmingham	Yes	IV
Ollivere /Duckett (2009)	Case series	Birmingham	Yes	IV
Revell (2006)	Case series	Birmingham (75%) Corin (25%)	Yes (75%) No (25%)	IV
Treacy (2005)	Case series	Birmingham	Yes	IV
Key Question 3				
Amstutz (2004)	Retrospective cohort	Conserve Plus	No (510-k investigational)	III
Beaule (2004)	Retrospective cohort	Conserve Plus	No (510-k investigational)	III
Le Duff (2007)	Retrospective cohort	Conserve Plus	No (510-k investigational)	III
McBryde (2007)	Retrospective cohort	Birmingham	Yes	III
Mont (2006)	Retrospective cohort	Conserve Plus	No* (510-k investigational)	III
Ollivere (2009)	Retrospective cohort	Birmingham	Yes	III

* performed under an FDA investigational device exemption (IDE) for the evaluation of the Conserve Plus hip resurfacing prosthesis.

4.1. Key question 1: What is the evidence of efficacy and effectiveness of total HR compared with THA?

4.1.1. Efficacy

WOMAC scores (Figure 3a)

No significant differences were identified in WOMAC scores between total HR and THA groups at one-year follow-up as reported in three RCTs.[64-66] A patient-reported outcome measure, WOMAC scores account for pain, stiffness, and physical functioning. All scores were normalized to a scale of 0 to 100; normalized postoperative WOMAC scores in the THA group ranged from 87.8 to 97.2, and from 90.4 to 96.9 in the total HR group. None of the studies reported a significant difference in the preoperative scores between groups.

SF-36 scores (Figure 3b)

No significant differences were identified in SF-36 quality of life scores between treatment groups at one-year follow-up as reported in two RCTs.[64, 65] Postoperative physical SF-36 scores ranged from 51.3 to 53.3 in the THA group and from 51.2 to 55.2 in the total HR group; postoperative mental SF-36 scores ranged from 52.1 to 55.1 in the THA group and from 51.9 to 53.9 in the total HR group. There were no statistically significant differences in preoperative scores between groups.

UCLA activity scores (Figure 3c)
Vendittoli[66] reported significantly higher one-year UCLA activity scores in the total HR group (7.1 versus 6.3 in the THA group) ($P = .037$); preoperative scores were not reported. The other two RCTs did not find a significant difference in one-year UCLA activity scores between groups.[64, 65] In the THA group, postoperative scores ranged from 6.3 to 8.3; in the total HR group, postoperative scores ranged from 6.8 to 8.0. Preoperative UCLA scores were only reported in one study, and were similar between groups. [64] The UCLA activity score is a patient-reported outcome measure.

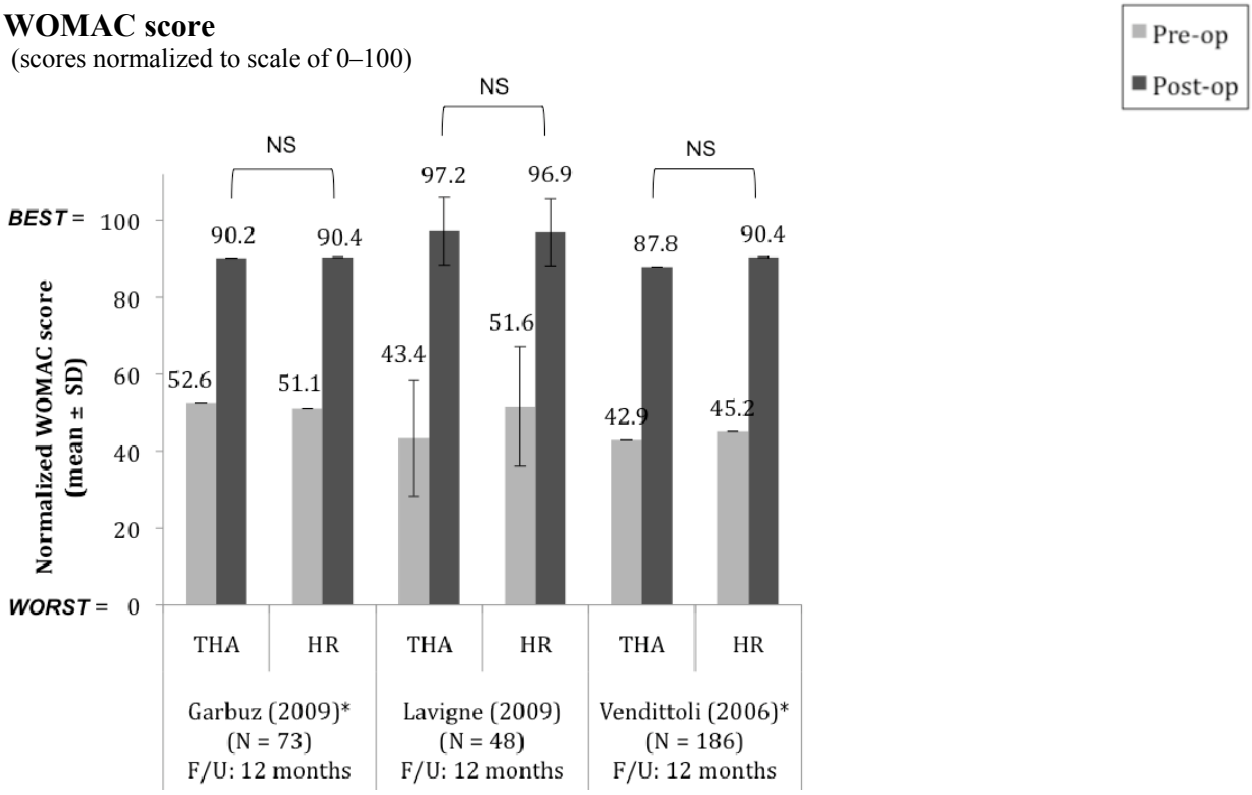
† Device had not yet received approval at the time of the , which was performed under an FDA investigational device exemption (IDE) for the evaluation of the Cormet hip resurfacing prosthesis.

Merle D'Aubigné (MA) scores (Figure 3d)
There were no significant differences in mean one-year MA scores between total HR and THA cohorts as reported by two RCTs.[65, 66] Postoperative scores ranged from 16.6 to 18.0 in the THA group, and from 16.7 to 17.9 in the total HR group. Mean scores increased from preoperative values by 6.4 to 7.5 points in the THA group and by 5.9 to 6.9 points in the total HR group. There were no significant differences in preoperative MA scores between groups. The MA score is a clinician-reported outcome measure that includes pain, ability, and walking ability components.

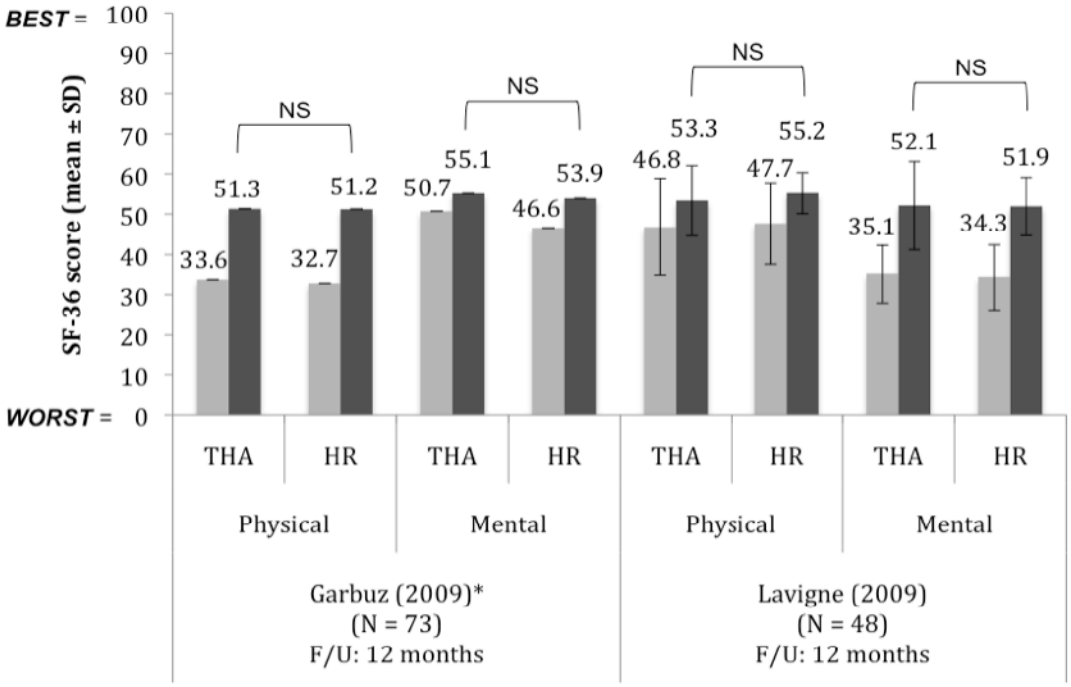
Pain scores (Figure 3e)
One RCT reported WOMAC pain scores.[64] There was no significant difference in pain scores at one-year follow-up between groups (THA: 90.0, total HR: 91.5). Preoperative pain scores were not significantly different between cohorts.

Figure 3. Short-term functional, quality of life, activity, and pain outcome measure scores from three randomized controlled trials.

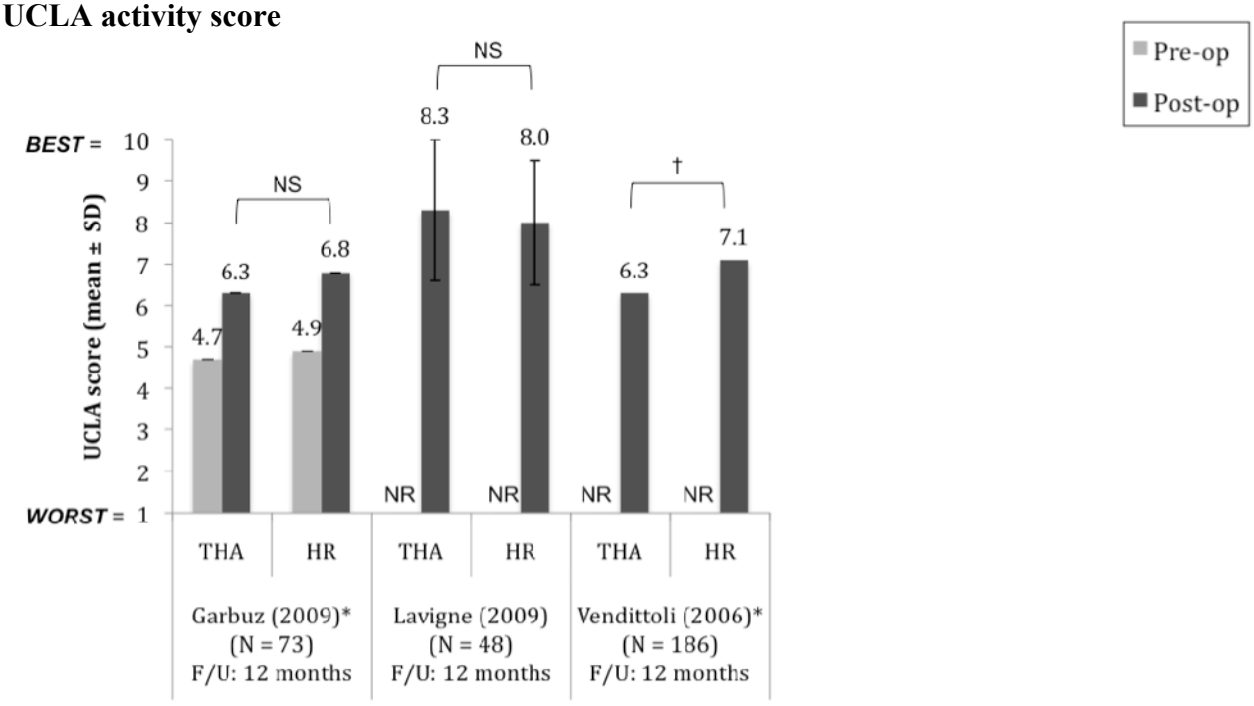
a. WOMAC score
(scores normalized to scale of 0–100)



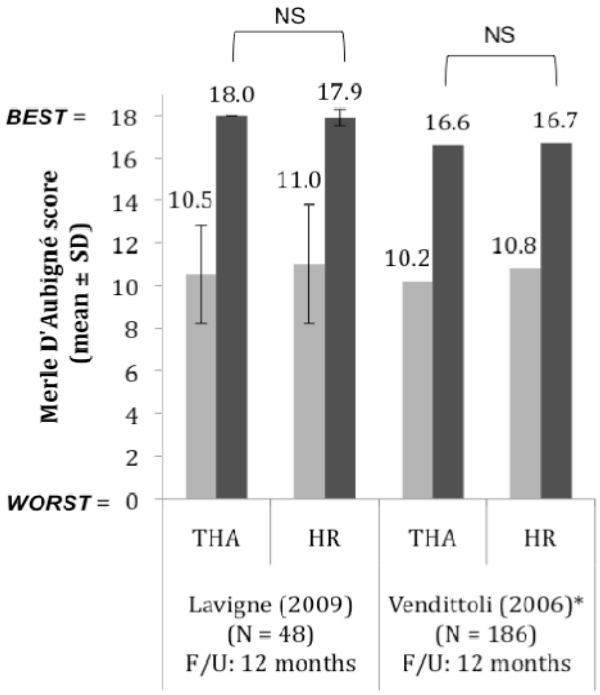
b. SF-36 scores



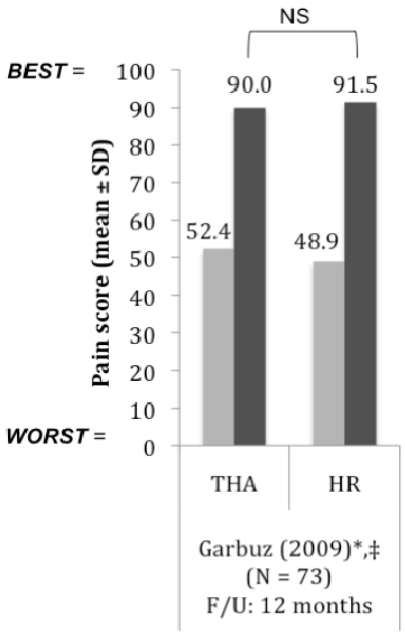
c. UCLA activity score



d. Merle D'Aubigné score



e. Pain score



HR: total hip resurfacing.
NR: *P*-value not report .
NS: not statistically significant ($P \geq .05$).
SD: standard deviation.
THA: total hip arthroplasty.

* Standard deviation not reported.
† $P < .05$.
‡ Garbuz (2009): Authors reported WOMAC pain score component normalized to a scale of 0–100.

4.1.2. Effectiveness (Figure 4)

Harris hip scores (HHS) (Figure 4a)
No significant differences were identified in postoperative HHS as reported by one prospective[68] and seven retrospective cohort studies.[10, 19, 46, 69-71, 73] Only short-term follow-up was available, as scores were obtained between a mean of 24 to 43.5 months in most studies. The mean postoperative HHS ranged from 89.7 to 96.2 in the THA group and from 90 to 98.1 in the total HR group. Postoperative HHS were slightly higher in the total HR group in six of the seven studies, although no differences between groups were statistically significant. Preoperative HHS were significantly higher in the THA group in one study by Fowble[68]; another study had significant differences in the preoperative HHS, however this difference was controlled for using multivariate statistical analysis.[19] Three studies excluded some or all patients who underwent revision surgery from this clinical outcome: Fowble excluded one hip in the total HR group, Li (2009) excluded one hip in the total HR group[69], and Stulberg[46] excluded 16 hips in the total HR group and three hips in the THA group but included an additional 10 hips that underwent revision

(eight in the total HR group and two in the THA group). Although all these hips are accounted for in the safety section, exclusion of patients that underwent revision surgery could bias results. The HHS is a clinician-reported outcome measure that accounts for pain, function, deformity, and range of motion.

Oxford score (Figure 4b)
No significant differences were found in postoperative Oxford scores as reported by Pollard et al (THA: 18.5, total HR: 15.9).[72] Mid-term follow-up was available for this study, with a mean follow-up of 70.7 months (5.9 years). However, Pollard excluded three hips in the total HR group yet included a total of five other hips that underwent revision (one in the total HR group and four planned revisions in the THA group), which makes the results somewhat difficult to interpret. Preoperative scores were not reported. The Oxford score is a patient-reported outcome measure that includes pain and function components.

Quality of life (Figure 4c)
EQ-5D
T Postoperative EQ-5D scores were significantly higher in the total HR group (0.9) compared to the THA group (0.78) as reported by Pollard et al.[72] Preoperative scores were not reported. Details of this study are provided in the preceding paragraph. The EQ-5D score is a patient-reported outcome measure that includes subscales to assess mobility, self-care, activity, pain, and mental health.

SF-12
Postoperative SF-12 physical scores were significantly higher in the total HR (53.6) group versus the THA (47.0) group as reported by one prospective cohort study (P = .002). [68] Although the preoperative scores were not significantly different between groups (P = .2), it is possible that at least some of the 6.6-point difference in the postoperative scores between groups may be accounted by a 7.8-point difference in the preoperative scores (THA: 25.8, total HR: 33.6). There was no significant difference in the SF-12 mental scores between groups either postoperatively (THA: 52.5, total HR: 54.6) or preoperatively (THA: 35.2, total HR: 44.2). Only short-term follow-up was available. The SF-12 is a patient-reported health survey.

Activity (Figure 4d)
UCLA
Postoperative UCLA activity scores were higher in the total HR group in all three studies that reported this outcome;[68, 69, 72] this difference reached statistical significance in two of the three studies,[68, 72] one of which had mid-term follow-up with a mean of 5.9 years. [68] Postoperative UCLA activity scores ranged from 3.6 to 6.8 in the THA group, and from 6.1 to 8.4 in the total HR group. Although the preoperative UCLA score was significantly higher in the total HR group (4.2) compared to the THA group (3.6) in one study[68] (P = .02), there was a 4-point increase between pre- and post-operative scores in the total HR group compared to a more modest 2.3-point increase in the THA group in this study. One study reported a *decrease* in postoperative UCLA activity scores

compared to their preoperative counterparts, but this is because the authors reported activity scores before patients were limited by pain and became an indication for arthroplasty.[72] The UCLA score is a patient-reported measure of activity.

Mont's scoring system

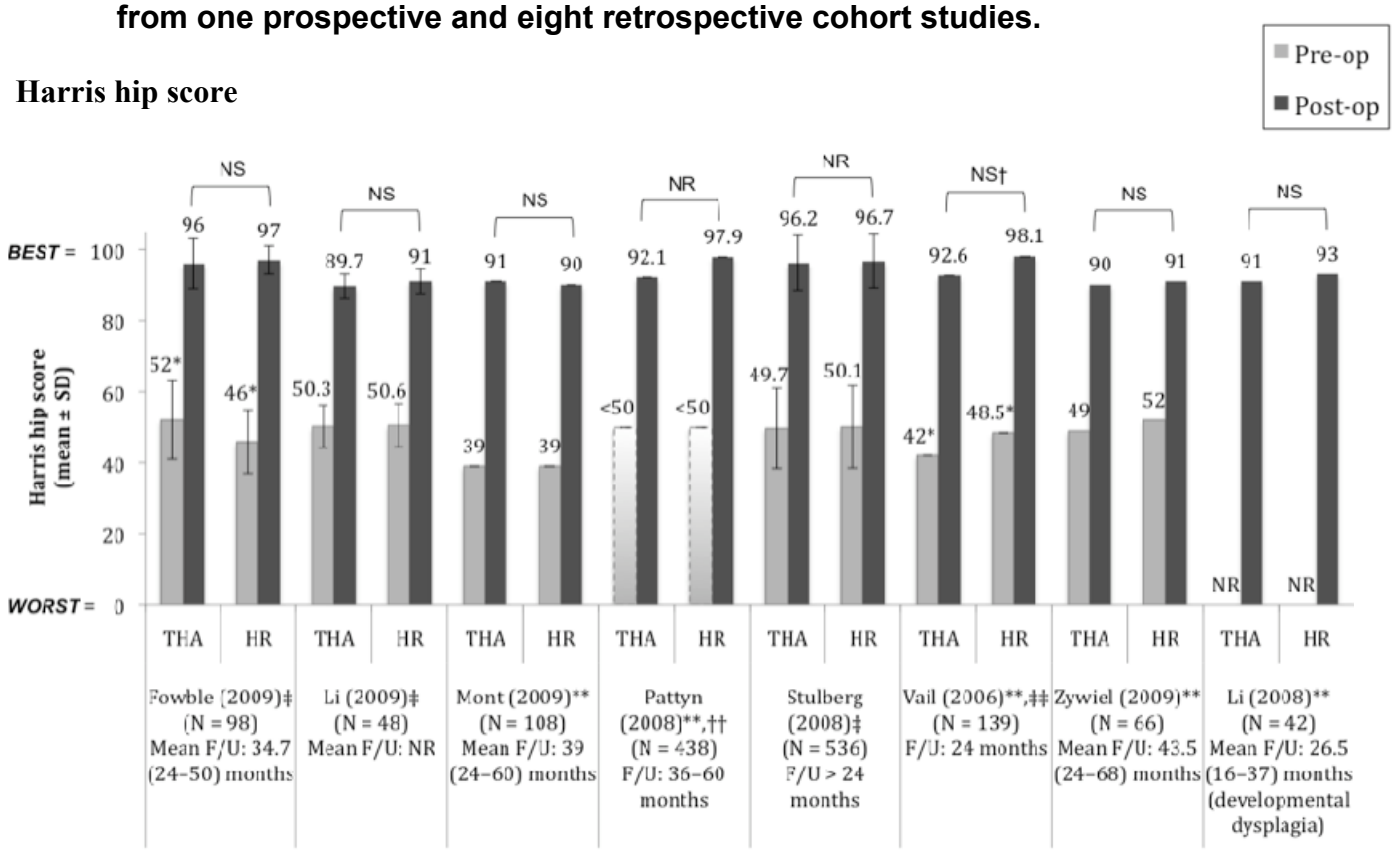
Using a scoring system devised by Mont (2009), two studies reported significantly higher activity scores for the total HR group compared to the THA group[10, 73] ($P = .0004$, $P < .001$). Postoperative scores were reported at a mean of 39 and 43.5 months for each of the two studies, and ranged from 5.3 to 7.0 in the THA group and from 10.0 to 11.5 in the total HR group. Although preoperative activity scores were significantly higher in the total HR group in one study (3 versus 2 in the THA group, ($P = .01$)), there was a pre- to postoperative 8.5-point increase in activity score in the total HR group compared to a 5-point increase in the THA group.[10] This activity scoring system is patient-reported, and takes into account the duration, frequency, and level of competitiveness of each activity the patient regularly participates in. From the description of the scoring system, there appears to be no maximum possible score.

Pain (Figure 4e)

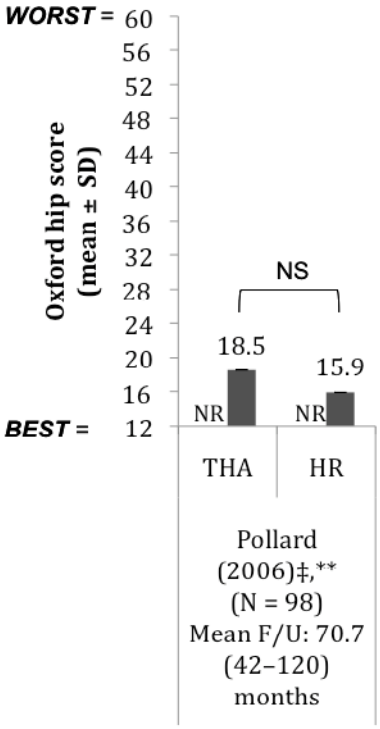
There appears to be no significant differences in the postoperative level of pain between total HR and THA treatment groups as reported by five retrospective cohort studies.[10, 19, 69, 70, 73] Postoperative pain scores ranged from 0.7 to less than 2 in the THA group, and from 0.9 to less than 2 in the total HR group, and were measured during short-term follow-up only (maximum mean of 43.5 months). All but one study reported pain scores from the VAS (visual analogue scale); Vail et al. reported Harris hip pain component scores, which were normalized here.[19] There were no significant differences in preoperative pain scores between groups as reported by two studies.[19, 69] VAS pain scores were patient-reported, with 0 indicating no pain and 10 indicating worst pain imaginable.

Figure 4. Functional, quality of life, activity, and pain outcome measure scores from one prospective and eight retrospective cohort studies.

a. Harris hip score



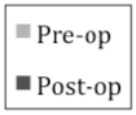
b. Oxford score

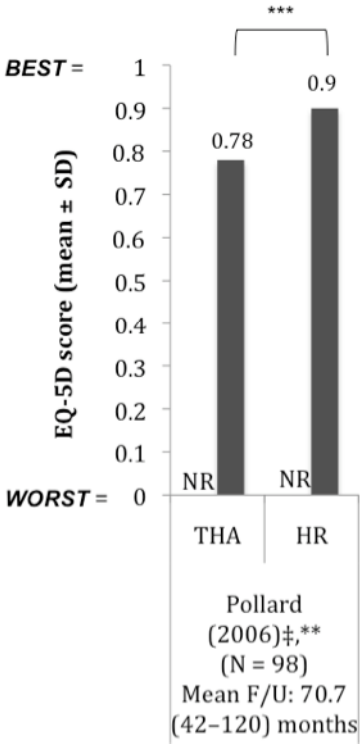


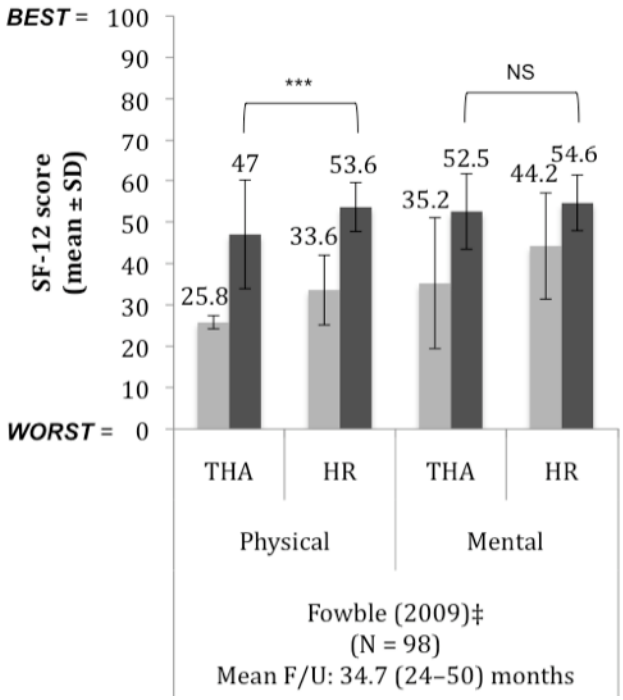
c. Quality of life scores

1. EQ-5D

2. SF-12 scores

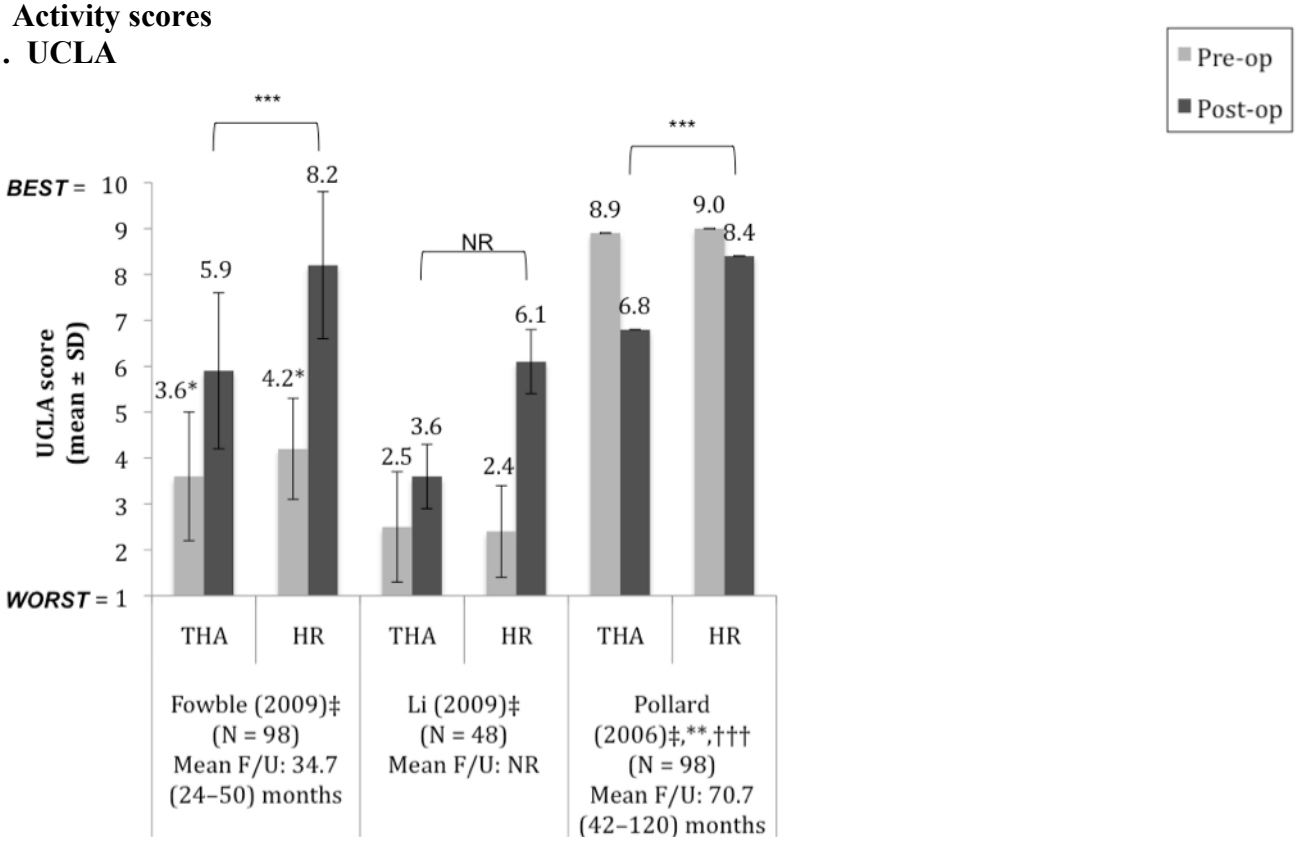




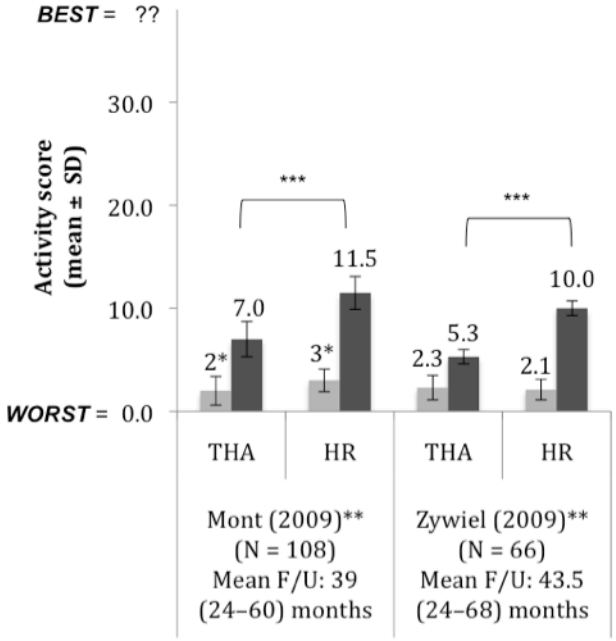

BEST =
WORST =
100
90
80
70
60
50
40
30
20
10
0
SF-12 score (mean ± SD)
***
NS
47
53.6
52.5
54.6
44.2
35.2
33.6
25.8
THA
HR
THA
HR
Physical
Mental
Fowble (2009)‡
(N = 98)
Mean F/U: 34.7 (24–50) months

d. Activity scores

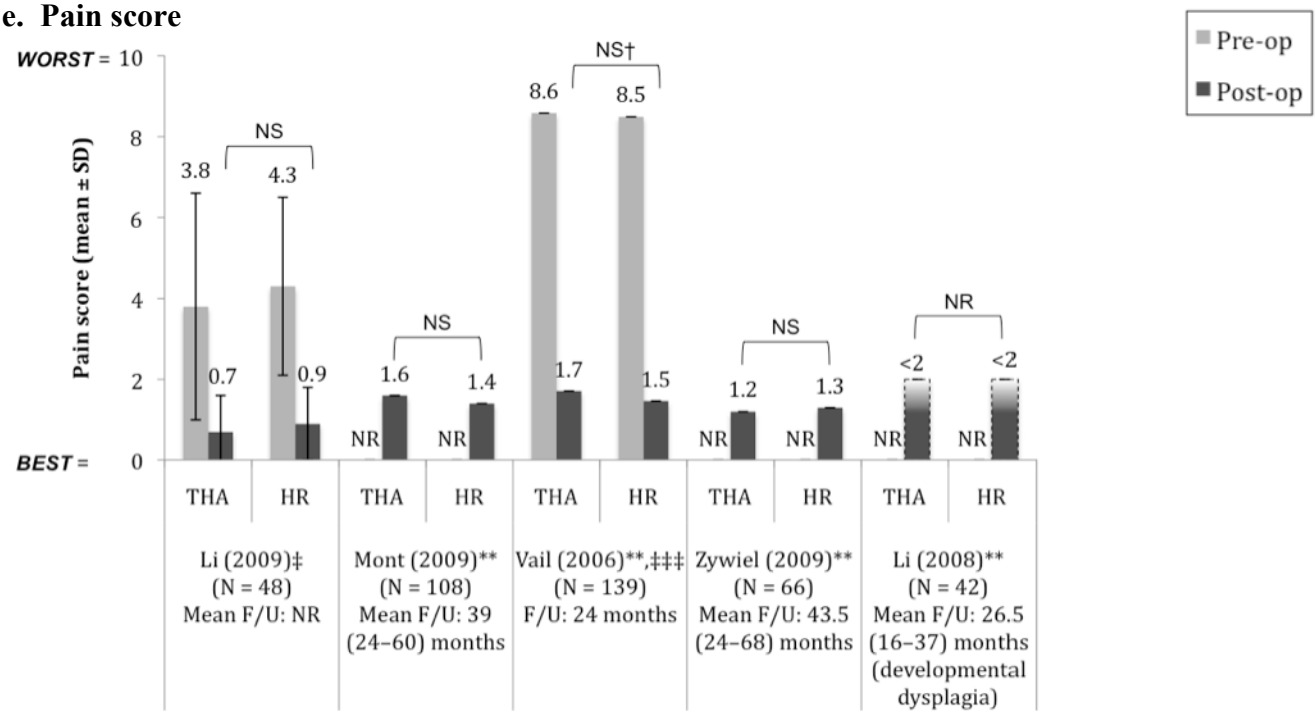
1. UCLA



2. Mont's scoring system



e. Pain score



HR: total hip resurfacing.
NR: *P*-value not reported.
NS: not statistically significant ($P \geq .05$).
THA: total hip arthroplasty.

* There was a statistically significant difference in the preoperative values between groups ($P < .05$).

† P-value reflects postoperative scores adjusted for age, gender, and preoperative values using multiple regression analysis because there were statistically significant differences in preoperative values and demographics.

‡ Author excluded patient(s) who underwent revisions from all clinical outcomes:
Fowble: one hip in the total HR group underwent revision (due to avascular necrosis).
Li (2009): one hip in the total HR group underwent revision (due to femoral neck fracture).
Pollard: three hips in the total HR group underwent early revision (due to avascular necrosis). Note that an additional five revisions were NOT excluded (one in the total HR and four planned revisions in the THA group).
Stulberg: 16 hips in the total HR group and 3 hips in the THA group underwent early revision (< 24 months) and were excluded from clinical outcomes.

** Standard deviation not reported.

†† Pattyn (2008): All patients had a preoperative HHS of less than 50, but the mean preoperative HHS was not reported.

‡‡ Vail (2006): Statistical significance for the postoperative scores was calculated *after* adjusting for age, gender, and preoperative values (the difference in the preoperative scores for THA versus total HR was statistically significant; there were also substantial differences in patient age and gender between the two groups (see patient demographics)).

*** $P < .05$.

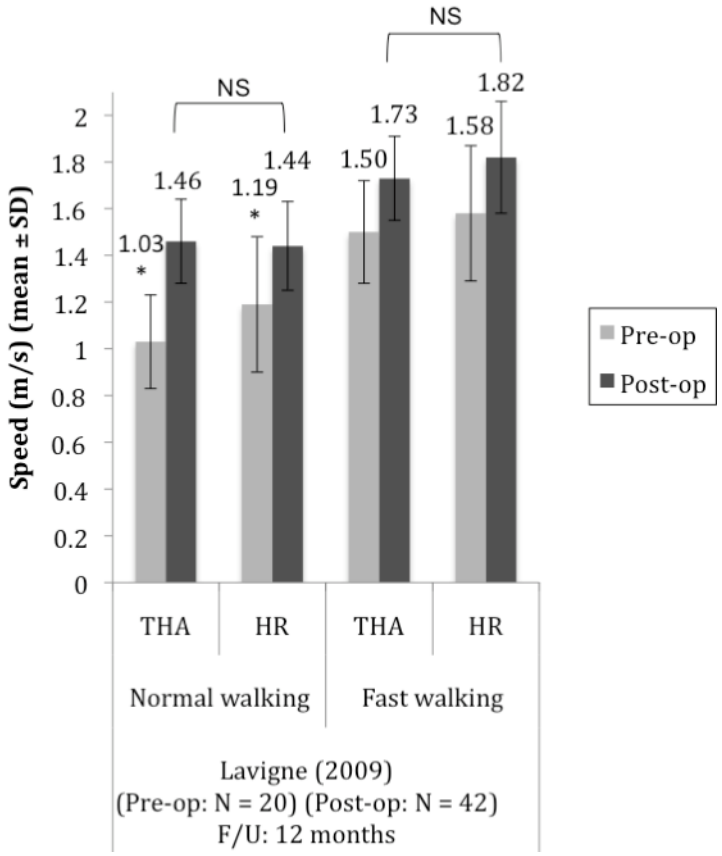
††† Preoperative scores were reported *before* patients were limited by pain; UCLA score was modified for the British population.

‡‡‡ Vail (2006): Authors reported the 45-point HHS pain score, which ranges from 0–44 (higher score indicates less pain). This score was normalized to an 11-point 0-10 scale with 10 indicating the worst pain: new pain score = 11 - (HHS pain score/4.5).

Walking speed (Figure 5)

One RCT reported normal and fast walking speeds at 12 months follow-up.[65] No significant differences were identified in postoperative walking speeds. The authors did note significantly higher preoperative normal walking speeds in the total HR group (1.19 m/s versus 1.03 m/s in the THA group), which could have biased the results.

Figure 5. Walking speeds from one randomized controlled trial.



HR: total hip resurfacing.
NS: not statistically significant ($P \geq .05$).
THA: total hip arthroplasty.
* There was a statistically significant difference in the preoperative values between groups ($P < .05$).

4.2. Key Question 2: What is the evidence of safety of hip resurfacing?

We present safety information in four sections: revisions, complications, surgical learning curve and metal ion safety. For revisions and complications, we stratified by follow-up period: short-term = 1 to 5 years, mid-term = 5 to 10 years, and long-term = >10 years. Short-term data are presented from available comparative studies and from three

international registries. Mid-term data are presented from available case-series and from one international registry. We found no long-term data.

4.2.1. Revision

Short-term follow-up

Most of the evidence from national registries and comparative studies suggest that short term follow-up revision rates are higher for total HR than THA.

Registry Studies (Figure 6)

Three national registry studies are consistent in reporting significantly higher revision rates in those receiving total HR compared with THA after three years of follow-up, Figure 6.

The National Joint Replacement Registry, Australia[75]

A comparison of time to revision revealed a significantly higher revision rate for total HR compared with conventional THA. The cumulative 3-year revision rate is 3.1% for total HR (95% CI = 2.8%, 3.5%) and 2.5% for conventional THA (95% CI = 2.5%, 2.6%).

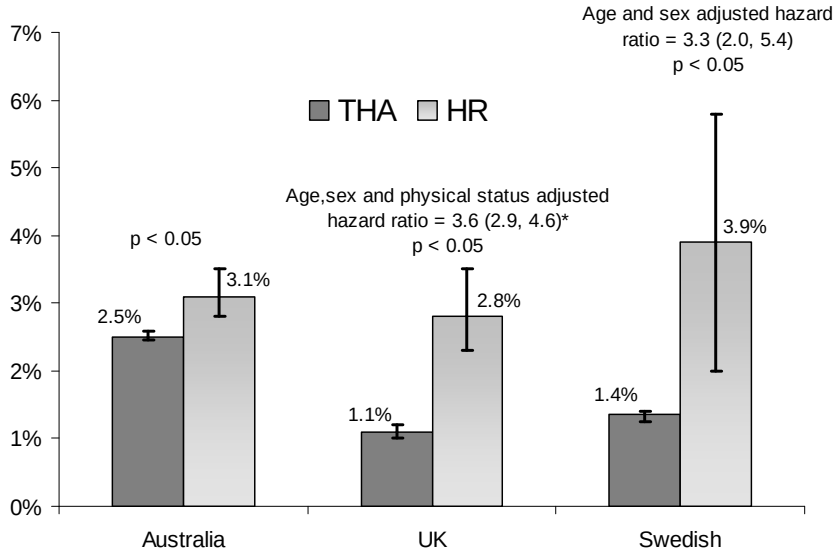
Swedish Registry[1]

The three year risk for revision is three times higher in those with hip resurfacing compared with THA, RR = 3.33 (95% CI = 2.04, 5.43). The risk among females is more than doubled compared with males, RR = 2.12 (95% CI = 1.03, 4.46).

National Joint Registry, UK[75]

The NJR overall 3-year revision rate varied according to type of prosthesis. The rates were lowest in patients who received cemented prostheses (0.7%, 95% CI: 0.6%, 0.8%) and highest after hip resurfacing (2.8%, 95% CI: 2.3%, 3.5%; hazard ratio = 3.6, 95% CI: 2.9%, 4.6%) in patients ≤65 years. The 3-year revision rate was 1.8% (95% CI: 1.6%, 2.1%) in patients who received an uncemented prosthesis and 1.3% (95% CI: 1.1%, 1.7%) in patients who received a hybrid.

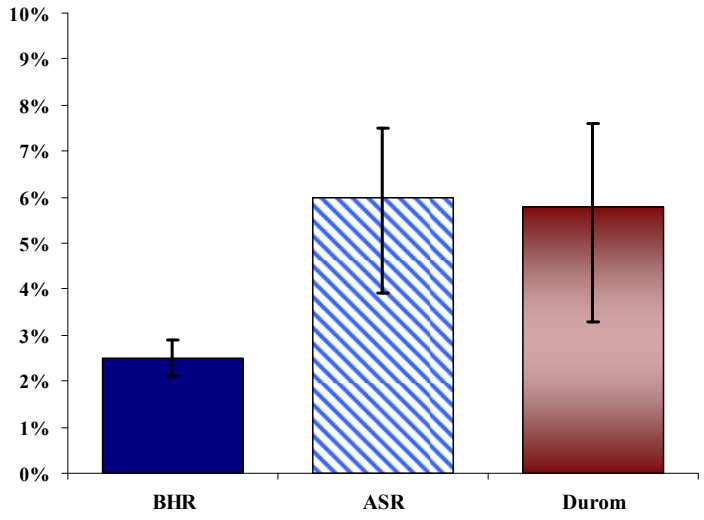
Figure 6. Three-year revision rates comparing hip resurfacing with total hip arthroplasty in three national registries.



*Hazard ratio given for comparison between cemented THA versus total HR.

Three total HR systems, the BHR, the ASR and the Durom, have over 1,000 observed component years available for comparison of revision rates in the Australian Registry. The 3-year cumulative percent revision is 2.5%, 6.0% and 5.8%, respectively, Figure 7.

Figure 7. Revisions of hip resurfacing per 100 observed component years for three different resurfacing systems.



Comparative Studies

Revision rates were reported in one RCT,[66] one prospective cohort study,[68] and six retrospective cohort studies,[10, 19, 46, 69, 71, 73] with short-term follow-up (< 5 years).

Overall, revision rates ranged from 0 to 4.3% of hips in the THA group, and from 0 to 7.8% of hips in the total HR group, Table 11. One RCT reported revision rates in 1.0% (THA) and 1.9% (total HR) of hips with one-year follow-up; a non-FDA-approved total HR prosthesis was used. Revision rates were slightly higher in the total HR group than in the THA group in four of eight studies reporting (risk difference ranged from 0.9–5.9%), the same in two studies, and lower in the total HR group than in the THA group in two studies (absolute risk difference ranged from 0.8–1.6). Causes of revision in the total HR group included aseptic loosening of the femoral head (n = 2),[66] avascular necrosis (n = 1),[68] femoral neck fracture (n = 11),[10, 19, 46, 69] acetabular cup migration (n = 1),[10] femoral component loosening (n = 11),[46] acetabular cup loosening (n = 5, one associated with infection),[19, 46] and dislocation (n = 1).[46] In the THA group, indications for revision included recurrent dislocations (n = 4),[66 71 19] dislocation (n = 1)[46], femoral fracture (n = 1),[46] acetabular cup migration (n = 1),[10] femoral component loosening (n = 2),[19, 46] aseptic loosening of the acetabular socket (n = 1),[19] infection (n = 3),[10 71 46] hip pain (n = 1),[46] and periprosthetic fracture (n = 1).[71]

Table 11. Revision and complication risks (%) from comparative studies

Author	Study design	Prosthesis FDA status	Mean F/U yrs (range)	No. of hips		**Revision (% hips)**		Complications
				THA	HR	**THA**	**HR**	
Lavigne[65]	RCT	Not approved	1.2 (12–18)	24	24	---	---	• Femoral component loosening THA: 0% HR = 0% • Heterotopic Ossification THA: 29.4% HR = 42.7%
Vendittoli [66] (Rama)	RCT	Not approved	1 (1)	102	103	**1**	**1.9**	• Femoral component loosening THA: 0% HR = 1.9% • Heterotopic ossification THA: 29.4% HR = 42.7%
Fowble*[68]	Prospective cohort	510-k investigational	2.9 (2–4.2)	44	50	**0**	**2**	• Femoral component loosening THA: 0% HR = 0% • Acetabular component loosening THA: 0% HR = 0% • Avascular necrosis THA: 0% HR = 2%
Li (2009)[69]	Retrospective cohort	Not approved	NR	41	39	**0**	**2.6**	• Femoral neck fracture THA: NA HR = 2.6% • Femoral component migration THA: 0% HR = 0% • Acetabular component migration THA: 0% HR = 0% • Heterotopic ossification THA: 14.6% HR = 5.1%
Li (2008)[70]	Retrospective cohort	Not approved	2.2 (1.3–3.1)	26	26	---	---	• Femoral component migration THA: 0% HR = 0% • Acetabular component migration THA: 0% HR = 0% • Heterotopic ossification THA: 0% HR = 0%
Mont (2009)[10]	Retrospective cohort	510-k investigational	3.3 (2–5)	54	54	**3.7**	**3.7**	• Femoral neck fracture THA: NA HR = 1.9% • Femoral component loosening THA: 0% HR = 0% • Acetabular component loosening THA: 0% HR = 0% • Femoral component migration THA: 0% HR = 0% • Acetabular component migration THA: 1.9% HR = 1.9%
Pattyn[71]	Retrospective cohort	Approved	NR (3–6)	250*	190*	**1.6**	**0**	• Femoral neck fracture THA: NA HR = 0.4% • Avascular necrosis THA: 0% HR = 0.4%
Stulberg[46]	Retrospective cohort	Approved	NR (2–NR)	253	283	**2**	**8.5**	• Femoral neck fracture THA: NA HR = 2.8% • Femoral component loosening THA: 0.4% HR = 3.9% • Acetabular component loosening THA: 0% HR = 1.4%
Vail[19]	Retrospective cohort	510-k investigational	3 (2–4)	93	57	**4.3**	**3.5**	• Femoral neck fracture THA: NA HR = 1.8% • Femoral component loosening THA: 1.1% HR = 0% • Acetabular component loosening THA: 1.1% HR = 1.8% • Heterotopic ossification THA: 0% HR = 10.5%

Zywiel[73]	Retrospective cohort	510-k investigational	3.6 (2–5.7)	33	33	**0**	**0**	NR

* Garbuz: Number of hips not reported; data based on the number of patients reported for each cohort.

Mid-term follow-up

Registry Studies

The Australian National Joint Replacement Registry has 7-year follow-up data for 10,623 HRs. A comparison of time to revision revealed a significantly higher revision rate for total HR compared with conventional THA, adjusted hazard ratio = 1.42 (1.24, 1.63), P < .001, Figure 8. The cumulative 7-year revision rate for total HR is 4.6% for total HR (95% CI: 3.9%, 5.4%) and 3.4% for conventional THA (95% CI: 3.2%, 3.7%).

Figure 8. Cumulative percent revision of conventional total hip arthroplasty and hip resurfacing requiring revision

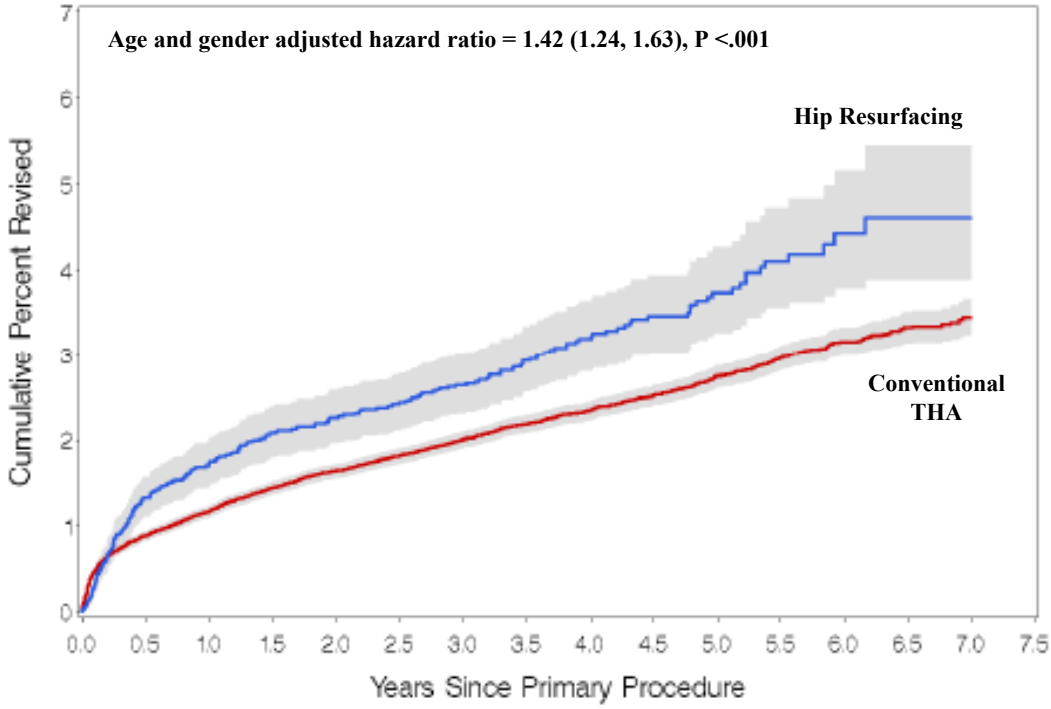


Figure used by permission from the AOA National Joint Replacement Registry, Australia.

Cohort studies and case-series

Revision rates were reported in one retrospective cohort study[72] and six case-series[76-81] with mid-term follow-up (5 to 10 years). The cohort study reported similar risks of revision, 7.8% in THA and 7.1% in total HR. Causes of revision in the total HR group included femoral neck fracture (n = 3, one was due to avascular necrosis) and femoral component failure (n = 1), and in the THA group included osteolysis (n = 3) and recurrent dislocations (n = 1).[72] In the case-series, revision rates in total HR-treated hips ranged from 0–7.7%, Table 12.

Table 12. Mid-term revision and complication risks (%) from case-series

Author	Study design	Prosthesis FDA status	Mean F/U (years) (range)	Preoperative diagnosis (N, %)	No. of hips	**Revision (% hips)**	Complications
Pollard[72]	Retro cohort	Approved	5.9 (3.5-10)	Osteoarthritis (74%) AVN (13%) Dysplasia (6%) Other (7%)	HR: 56 THA: 51	**HR: 7.1** **THA: 7.8***	◆ Femoral neck fracture = 5.4% (THA = NA) ◆ Avascular necrosis = 1.8% (THA= 0%) ◆ Osteolysis = 1.9% (THA = 17.6%) ◆ Femoral component migration = 9.4% (THA = 0%) ◆ Acetabular component migration = 0% ◆ (THA = 0%)
Amstutz[76] (2008)	Case series	510-k investiga-tional	6 (2–12)	Dysplasia (100%)	103	**7.7**	◆ Femoral neck fracture = 1.9% ◆ Osteolysis = 0% ◆ Femoral component loosening = 2.9% ◆ Acetabular component loosening = 0% ◆ Acetabular component migration = 0%
McBryde (Revell) (2008)†[77]	Case series	Approved	5.3 (2.0–9.4)	Osteoarthritis (100%)	909	**1.4**	◆ Femoral neck fracture = 0.1%† ◆ Avascular necrosis = 0.2%† ◆ Acetabular component loosening = 0.3%† ◆ Acetabular component migration = 0.2%†
McMinn[78]	Case series	Approved	7.8 (6–9.6)	Dysplasia (acetabular insufficiency) (100%)	110	**2.7**	◆ Femoral neck fracture = 0.9% ◆ Avascular necrosis = 0% ◆ Osteolysis = 0% ◆ Acetabular component loosening = 0%
Ollivere (Duckett) 2009[79]	Case series	Approved	5.1 (3.2–6.3)	NR	104	**0**	◆ Femoral neck fracture = 0% ◆ Avascular necrosis = 0% ◆ Osteolysis = 5.8% ◆ Femoral component loosening = 0% ◆ Acetabular component loosening = 0%
Revell[80]	Case series	Approved (75%)	6.1 (2–12)	Osteonecrosis (100%)	73	**6.8***	◆ Femoral neck fracture = 0% ◆ Avascular necrosis = 0% ◆ Femoral component loosening = 2.7%‡ ◆ Heterotopic ossification = 15.6%
Treacy[81]	Case series	Approved	≥5 (5–NR)	Osteoarthritis (87%) AVN (7%) Dysplasia (2%) Other (4%)	144	**2.1**	◆ Femoral neck fracture = 1.4% ◆ Avascular necrosis = 0.7% ◆ Osteolysis = 0% ◆ Femoral component loosening = 0.7% ◆ Acetabular component loosening = 0% ◆ Femoral component migration = 0% ◆ Heterotopic ossification = 28%

* Includes revisions planned: Pollard: all revisions in THA group planned; Revell: one revision planned.† McBryde: only reported causes of revision. There were two cases of femoral head collapse secondary to osteonecrosis, so these were not included as loosening or migration of the femoral component.
‡Revell: authors included one case of femoral head collapse as loosening of the femoral component.

4.2.2. Complications

Short- and mid-term complications are summarized below.

Table 13. Summary of short- and mid-term complications

	Short-term follow-up (Table 11)	Mid-term follow-up (Table 12)
Femoral neck fracture	Rates of femoral neck fractures ranged from 0.4–2.6% of resurfaced hips as reported by five retrospective cohort studies.[10, 19, 69, 71, 82] Femoral neck fracture was associated with female sex in one study,[44] but not in another.[83]	Femoral neck fracture rates were reported in one retrospective cohort study to be 5.4%[72] and in six case-series with mid-term follow-up at rates ranging from 0–1.9% of hips.[72, 76-81]
Avascular necrosis	The rate of AVN was higher in hips treated with total HR (0.4–2.0%) compared with those treated with THA (0%) as reported by one prospective and one retrospective cohort study.[68, 71]	Avascular necrosis was reported in one retrospective cohort study at a rate of 1.8% of total HR hips and in 0% of THA hips.[72] Rates of AVN ranged from 0–0.7% as reported by five case-series.[77-81]
Osteolysis	Osteolysis was not reported in any of the comparative studies with short-term follow-up.	Osteolysis occurred at a rate of 1.9% in total HR-treated and 17.6% of THA-treated hips, respectively, as reported by one retrospective cohort study.[72] The rate of osteolysis in hips that underwent total HR ranged from 0–5.8% as reported by four case-series.[76, 78, 79, 81]
Femoral component loosening	Femoral component loosening was detected in 0% and 0–1.9% of hips in the THA and total HR cohorts, respectively, as reported by two RCTs.[65, 66] One prospective cohort study noted that no hips in either group had evidence of femoral component loosening.[68] Finally, three retrospective cohort studies reported rates of femoral component loosening that ranged from 0–1.1% and 0–3.6% of hips (THA and total HR, respectively).[10, 19, 46]	Femoral component loosening was observed in 0–5.1% of resurfaced hips as reported by four retrospective studies.[76, 79-81]
Acetabular component loosening	One RCT and one prospective cohort study both reported no cases of acetabular component loosening in either group.[65, 68] Three retrospective cohort studies reported rates of acetabular cup loosening that ranged from 0–1.1% of THA-treated hips and 0–1.8% of total HR-treated	Acetabular component loosening was detected in a range of 0–0.3% of hips as reported by five retrospective studies.[76-79, 81]

Table 13. Summary of short- and mid-term complications

	Short-term follow-up (Table 11)	Mid-term follow-up (Table 12)
	hips.[10, 19, 46]	
Femoral component migration	Femoral component migration was not detected in any hips in either group as reported by three retrospective cohort studies.[10, 69, 70]	One retrospective study reported femoral component migration in 9.4% of total HR-treated hips and in 0% of THA-treated hips[72]; one case-series reported no cases of femoral component migration following total HR.[81]
Acetabular component migration	Acetabular cup migration rates were noted in a range of 0–1.9% of THA-treated and 0–1.9% of total HR-treated hips as reported by three retrospective cohort studies.[10, 69, 70]	There was no evidence of acetabular component migration in either group as reported by one retrospective cohort study[72]. Two case-series reported rates of acetabular component migration ranging from 0-0.2% of hips.[76, 77]
Heterotopic ossification	HO rates were as high as 29.4% in THA-treated and 42.7% in total HR-treated hips as reported in one RCT.[67] Resurfaced hips had a significantly higher rate of Brooker grade III and IV HO than hips treated with THA ($P = .02$). Of the 30 hips in the THA group with HO, 34.1% were Brooker grade I, 29.5% were Brooker grade II, 4.5% were Brooker grade III, and none were Brooker grade IV. Similarly, of the 44 hips in the total HR group with HO, 27.3% were Brooker grade I, 43.2% were Brooker grade II, 18.2% were Brooker grade III, and 11.4% were Brooker grade IV. Significantly more total HR-treated hips had HO in the central and lateral regions (27.3% versus 3.3% of THA-treated hips, ($P = .011$)), while there was no difference in the rates of hips with HO in the central region only. Male sex was a risk factor for HO in both groups (total HR+THA: males: 43.5%, females: 24.6%, ($P = .009$)), as well as in the total HR (males: 50.8%, females: 28.9%, ($P = .039$)) but not THA group (males: 36.4%, female: 19.4%, ($P \geq .05$)). Furthermore, males had a higher ossification grade ($P = .037$). Three retrospective cohort studies reported HO in 0–14.6% of hips in the THA group and in 0–10.5% of hips in the total HR group.[19, 69, 70] In the THA group, Brooker grade classification was I (100%), II (0%), and III (0%); in the total HR	Two case series with mid-term follow-up reported HO rates in 15.6–28.0% of resurfaced hips.[80, 81]

Table 13. Summary of short- and mid-term complications

	Short-term follow-up (Table 11)	Mid-term follow-up (Table 12)
	group, Brooker grades were I (83.3–100%), II (0%), and III (0–16.7%) as reported by two studies.[19, 69]	

4.2.3. Learning curve threshold

It has been suggested that there is a steep learning curve associated with hip resurfacing arthroplasty. A number of factors may affect the success of this procedure that are often improved with increased surgeon experience, including patient selection, [60, 84] optimal component positioning,[84-86] and component size[60, 84] The following studies were identified that address the association between patient outcomes and surgical experience.

Marker et al (2007)[60] performed 550 MoM total HR arthroplasties between 2000 and 2006. Although the overall risk for femoral neck fracture was 2.5% (14 fractures), 12 of these occurred in the first 69 hips treated (86% of fractures; incidence of 17.4%), while only two occurred in the remaining 418 hips (0.4%). Surgeons who treated patients that developed femoral neck fractures had significantly less experience performing total HR surgeries compared to those who treated patients who did not develop fractures (mean of 69 previous total HR procedures in the fracture group versus 279 in the non-fracture group ($P < .001$)). In addition, the authors found over the course of the study that they could reduce femoral neck notching by decreasing the thickness of the acetabular shell from 10 to 6 mm. This change took place after patient 78, and allowed for the use of a larger femoral component and improved femoral and acetabular bone conservation. Other changes were also made in surgical technique, including not cementing the femoral component on femoral neck or cysts and limiting the femoral component cement mantle to 2 mm. The authors also placed additional restrictions in total HR patient selection, and began to exclude patients who would need more than minimal notching of the femoral neck.

Mont et al (2007)[84] compared the outcomes of the first 292 patients treated with total HR (Conserve Plus) to the subsequent 724 patients. The authors evaluated their techniques and patient outcomes after treating the first group of patients and made a number of changes to their surgical indications and techniques in the 724 subsequent total HR patients. Three types of risk factors for failure were identified, which included preoperative (femoral head cysts, abnormalities in the head-neck junction, inadequate bone density), operative (not covering reamed femoral bone, using smaller femoral component to conserve acetabular bone, malpositioning of the acetabular component, and leaving the femoral component proud), and postoperative risk factors (patient does not heed postoperative restrictions, traumatic events). The improved patient selection and operative technique led to a decrease in the rate of femoral neck fractures from 7.2% to 0.8% ($P = .0001$) and in the revision rate from 13.4% to 2% ($P < .001$).

Nunley et al (2009)[85] evaluated the first 100 hip resurfacing procedures performed by each of five surgeons. None of the surgeons had prior training in this technique yet all performed a high volume of other joint reconstruction surgeries (mean of 220 hip arthroplasties per year) and had many years in clinical practice (mean of 21.4 years). The rate of major complications (femoral neck fracture, nerve injury, dislocation, infection, and acetabular bone in-growth) was stratified by the first 25 cases per surgeon, the second set of 25 cases, etc. The rate of major complications for all surgeons was significantly higher in the first 25 cases (5.6%) compared to the second 25 cases (1.6%) ($P < .002$); the third and fourth subsequent sets of 25 cases each had a major complication rate of 1.6%.

O'Neill et al (2009)[87] evaluated the first 50 cases performed by each of five surgeons with no prior training in hip resurfacing but who performed at least 100 THA per year, and found a postoperative revision rate of 3.2%. The authors suggested that this relatively low revision rate may be due to the surgeons' high-volume practices.

Sielbel et al (2006)[88] noted a decrease in the revision rate with increased surgical experience, from 5% in the first 100 cases to 2% in the next 100 to 1% in the last 100 resurfacing procedures. However, this trend was not statistically significant ($P = .308$). Although longer follow-up time was available for the first set of patients, the mean follow-up in this study was quite short at 202 days.

Witjes et al (2009)[86] evaluated the learning curve for optimal component positioning in the first 40 cases performed by a single surgeon. Implant positioning was determined radiographically and compared against a set of predetermined "optimal" measurements for cup abduction angle, stem shaft angle, and cup head angle. Although the number of cases was too small to achieve statistical significance, a trend towards more optimal component positioning was found.

4.2.4. Total Hip Resurfacing Revision Surgery

We identified two small studies that attempted to address whether converting a total HR to a THA was safe and successful. Both studies compared outcomes of total HR patients revised to THA with outcomes of patients who underwent primary THA. However, neither evaluated morbidity following revision of HR compared to revision of THA, which is important given that younger patients who undergo hip arthroplasty are likely to need revision surgery at some point in the future.

Ball et al[28] retrospectively evaluated outcomes of 20 patients (21 hips) following conversion from total HR (Conserve Plus; Wright Medical Technology) to conventional THA due to femoral component failure and compared the results with those of 58 patients (64 hips) who had undergone primary total hip arthroplasty with conventional or cross-linked polyethylene bearings and MoM bearings during the same time period. All surgeries were performed via a posterior approach. Mean patient age was 50.2 years (range, 23–72) in the total HR group, 50.8 years (range, 27–64) in the THA group, and

55% and 65% of patients were male, respectively. The primary but not exclusive diagnosis in both groups was osteoarthritis. Indications for conversion from total HR to THA were femoral neck fracture (5 hips) and femoral component loosening (16 hips). Patients were followed for a mean 3.8 (1–9.4) years and 4.8 (2–8.8) years, respectively. All outcomes were found to be comparable between revision of the total HR and the THA groups, with no significant differences reported for surgery time (178 versus 169 minutes), blood loss (509 versus 579 mL), length of hospital stay (4.0 versus 4.2 days), SF-12 physical (48.6 versus 47.1) and mental (54.2 versus 50.3) scores, UCLA pain (9.3 versus 9.6), walking (9.4 versus 9.2), function (9.3 versus 8.8) and activity (6.8 versus 6.4) scores, and Harris Hip scores (92.2 versus 90.3), respectively. Furthermore, no differences between groups were seen on radiographic evaluation in relation to stem or acetabular fixation scores, limb length discrepancy, femoral offset, and the position of the center of rotation of the hip. One further two-stage revision was required in the THA group for a deep infection. Overall complication rates, to include femoral nerve palsy, proximal femoral fracture, infection, and perioperative myocardial infarction, were also comparable between groups, total HR revision, 15% (n = 3); THA revision, 10% (n = 6).

Grammatopoulos et al[89] conducted a retrospective analysis of 53 Birmingham total HRs revised to conventional THA and compared the outcomes of those revised for the formation of a pseudotumor with those revised for other reasons and with that of a matched cohort of 103 primary THA patients. Mean age of both groups was 54 years (range, 20–71) and 36% and 38% of the patients were male in the total HR and THA groups, respectively. For the total HR patients, the main diagnosis was primary osteoarthritis in 40 patients (76%). Eighty percent of the revision surgeries were performed via a posterior approach, 15% using an anterolateral approach 5% using the Smith-Petersen approach. Indications for revision from total HR to THA were inflammatory pseudotumor in 16 patients (30%), femoral neck fracture in 21 (40%), and other to include loosening, infection, avascular necrosis/collapse, and recurrent dislocations in 16 (30%). In the matched THA group, 32 patients (31%) had pseudotumors, 41 (40%) had femoral neck fractures, and 30 (29%) had other indications. Patients were followed for a mean 3 years (0.8–7.2). All outcomes following revision for pseudotumor were found to be significantly worse than for any other indication and when compared with primary THA. No differences were found between the two groups for all other indications. The length of surgery differed within the total HR revision group with a mean 161.6 minutes for pseudotumors, 129.4 minutes for other, and 99.6 minutes for fracture ($P < .002$). When the total HR group was compared with the THA group, only the pseudotumor groups differed in surgery duration, with the total HR group taking significantly longer (161.6 minutes versus 113.1 minutes; $P < .001$). Blood transfusion after surgery was required in 12 cases (75%) of total HR revision for pseudotumor versus seven (33%) for fracture and seven (44%) for other, and compared with only four (10%) cases in the pseudotumor THA group. The pseudotumor group had significantly worse function as judged by the mean Oxford Hip Score (OHS) of 20.9 versus 40.2 and 37.8 for the fracture and other groups, respectively; $P < .001$. Furthermore, OHS scores between the total HR revision and THA groups differed only for the pseudotumor groups, with those following total HR revision showing worse outcomes (20.9 vs. 39.1; $P < .001$).

Likewise, significantly lower activity levels as indicated by the mean UCLA score were seen for pseudotumors, 3.8 versus 7.0 for fracture and 6.7 for other, P < .001. UCLA scores were not available for the THA group. Further revisions were required in five (38%) of the patients in the pseudotumor group compared with three (14%) in the fracture group and 2 (13%) in the other group. No further revisions were reported in the THA group. The rate of major complications was also greater for the revised pseudotumor group with eight cases (50%) versus three (14%) and two (13%) cases for the fracture and other groups, respectively, and when compared with that of the THA pseudotumor group (50% vs. 6.2%, *P* < .01). The authors concluded that the outcome of revision for pseudotumour is poor and consideration should be given to early revision to limit the extent of the soft-tissue destruction. The outcome of resurfacing revision for other causes is good.

4.2.5. Metal Ion Safety

Short-term follow-up

Serum concentrations of metal ions

One RCT reported significantly higher concentrations of cobalt and chromium ions in the serum of THA-treated patients compared to that detected in total HR-treated patients at both one- and two-years follow-up.[64] Cobalt was detected in the serum of THA patients at concentrations of 5.09 and 5.38 μg/L at one- and two-years follow-up, respectively, whereas cobalt levels were only 0.51 and 0.54 μg/L at the same time points in patients who underwent total HR (*P* = .000 for one-year follow-up, *P*–value not reported for two-year follow-up). Serum chromium levels were 2.14 and 2.88 μg/L in THA-treated patients, and only 0.84 and 0.81 in total HR-treated patients at one- and two-years follow-up, respectively (*P* = .023 for on-year follow-up, *P*–value not reported for two-year follow-up).

Background and significance

Patients with metal-on-metal (MoM) joint replacements are likely to experience elevated metal levels throughout the life of the prosthesis. Metal release is a potential issue not just for MoM total HR, but for MoM THA as well. Concerns have been raised regarding the safety of and risks associated with prolonged exposure to metal ions, and whether such exposure may increase the risk of cancers or metabolic disorders.[90]

The primary metals used in MoM prostheses are cobalt (Co) and chromium (Cr). Although these metals are essential trace elements that are important for many biological processes in the human body (and are found in food and water supplies), they are considered toxic and hazardous by inhalation. Cobalt and chromium exposure is regulated by the Occupational Safety and Health Administration (OSHA). According to OSHA's occupational safety and health standards limits for air contaminants (expressed as milligrams of substance per cubic meter of air [mg/m]), cobalt metal, dust, and fume (as Co) has a limit of 0.1 mg/m and chromium metal and insoluble salts (as Cr) has a limit of 1 mg/m.[91] To date, no OSHA regulations exist regarding metal levels following orthopaedics procedures.

MoM articulations generate a much larger number of particles every year than do conventional metal-on-polyethylene components. However, the particles produced through MoM wear are smaller (generally < 50 nm)[92] than those generated by polyethylene wear (generally > 100 nm)[93], which results in a lower actual volumetric wear. The processes by which these particles are taken up by cells in the human body differ depending on their size. While smaller metal particles (< 150 nm) are taken up by cells through endocytotic processes (non-specific receptor-mediated endocytosis and pinocytosis),[94] larger particles (> 150 nm) stimulate phagocytosis in specialized cells called macrophages.[95] The response of macrophages to wear debris is thought to be responsible for implant loosening in patients with metal-on-polyethylene bearings. In contrast, the smaller particles created by MoM bearings have limited ability to activate macrophages.[96, 97] Once internalized into a cell, metal particles can induce cytotoxicity,[98] chromosomal abnormalities,[99] and oxidative stress.[100] Metal ions released from orthopaedic implants have also been shown to induce apoptosis and/or necrosis in a range of cells, with Co(II) and V(III) among the most cytotoxic.[98, 101] The major theoretical concerns regarding MoM hip resurfacing include hypersensitivity-related failures, allergic reaction, aseptic lymphocytic vasculitis-associated lesions (ALVAL), local tissue toxicity, impaired renal function, chromosomal damage, and possible malignant cellular transformation/cancer.

The variability of methods used to assess metal levels in orthopaedic studies, such as analytical technique, time of collection, units of measurement, and specimen can make reliable comparisons difficult between studies.[90] Historically, metal ion concentrations in total HR and THA patients have been measured using serum plasma levels; other methods include whole blood, red blood cells, and urine specimens. However, whole blood measurement has been shown to be more accurate than serum plasma levels to indirectly measure metal wear and systemic metal exposure,[50, 102, 103] as has daily output of metal ions in urine.[104] One study found a significant difference between serum and whole blood cobalt and chromium concentrations such that there was an over-estimation of cobalt and chromium ion concentrations in serum levels compared with whole blood levels.[105]

Metal Ion Concentrations in MOM Articulations

Patients with MoM articulations are likely to experience elevated metal levels throughout the life of the prosthesis. How or if these levels change following prosthesis implantation and the degree to which these levels differ between hip resurfacing and total hip replacement is analyzed.

Metal-on-Metal Hip Resurfacing
In general, studies tend to report a pattern of early increase in metal ion levels followed by a steady decrease in levels thereafter. This decline may be suggestive of a "running-in" period indicative of MoM hip resurfacing articulations. Though the trend is for a reduction in metal ions over time, the concentrations still remain high after long-term follow-up, the consequences of which remain unclear.

Allan et al. monitored serum cobalt (Co) and chromium (Cr) levels at multiple time points over a three year period in 35 patients who underwent unilateral hip resurfacing with the Cormet 2000 MoM implant.[106] Co and Cr levels were significantly higher at all time points when compared to preoperative values, with the highest levels seen at one year. The mean preoperative serum Co levels were 0.28 μg/L, which increased to 3.70 μg/L at six months and 4.31 μg/L at one year, and then began to show signs of decline at two and three years (2.75 μg/L and 2.50 μg/L, respectively). The mean preoperative serum Cr level was 0.32 μg/L, which increased to 4.52 μg/L at six months and 5.12 μg/L at one year, and again declined at two and three years (3.75 μg/L and 4.26 μg/L, respectively).

Back et al. also noted a pattern toward an early increase in ion levels followed by a decline thereafter in 16 patients following hip resurfacing with the Birmingham MoM prosthesis.[50] Serum cobalt levels increased from a preoperative level of 5.69 μmol/L to 50 μmol/L at three months and 56.56 μmol/L at six months, and then gradually declined thereafter: 9 months 51.5 μmol/L (nine months), 40.2 μmol/L (one year), and 31.83 μmol/L (two years). Although the two-year serum Co concentrations were 5.6 times greater than preoperative levels, they had decreased significantly ($P = .007$) from their peak at six months. Serum chromium levels showed a similar pattern: 6.00 μmol/L (preoperatively), 75.7 μmol/L (three months), 85 μmol/L (six months), peaked at 95μmol/L (nine months), and then declined to 76.9 μmol/L (one year), and 67.92 μmol/L (two years). At two years, mean serum chromium ion concentrations were 11.3 times greater than preoperative levels; however, there was a significant decrease ($P = .02$) in mean ion concentrations from the nine-month peak.

Similarly, in a study looking a outcomes in 64 patients following hip surfacing with the Durom MoM high carbon content device,[105] mean whole blood concentrations of chromium and cobalt increased steeply early in follow-up, reaching their highest levels by three months ($P = .0006$ and $P < .001$, respectively), and then declined steadily thereafter. Preoperative levels of chromium and cobalt were 0.92 μg/L and 0.15 μg/L, respectively, followed by 2.01 μg/L and 0.90 μg/L at three months, 1.89 μg/L and 0.80 μg/L at six months, 1.61 μg/L and 0.67 μg/L at one year, and 1.37 μg/L and 0.59 μg/L at two years. The only significant decline in chromium concentrations was seen when comparing the one- and two-year follow-up points ($P = .0416$). In contrast, cobalt concentrations declined significantly between three and six months ($P = .0003$) and six months and one year ($P = .0114$), but did not differ significantly at the one- and two-year marks.

Daniel et al. (2007) conducted a prospective, longitudinal study that investigated blood and urine metal ion levels over a four-year period in 26 young and active male patients after unilateral Birmingham hip resurfacing arthroplasty.[104] A second study in the same population was published a few years later which extended the results to six years of follow-up.[107] The daily urinary output of cobalt (μg/day) increased for six months and then declined thereafter (0.3 μg/day (preoperative median) increased to 3.6 (five days), 7.1 (two months), and 12.1 (six months), then decline to11.9 (one year), 10.6 (two years), and 8.1 (four years)). Compared to the preoperative values, the Co levels were significantly higher

all time points, even as early as five days. However, the differences between the levels at two months and four years were not statistically significant, which suggests that they may have reached a plateau. The daily urinary output of chromium (μg/day) followed a similar, yet less pronounced increasing trend up to one and two years followed by a decrease at four years (preoperative median: 0.5 μg/day (preoperative median) increased to 1.8 (five days), 2.7 (two months), 4.5 (six months), and then 4.8 (one year and two years), and then declined to 3.6 (four years)). Again, there was a significant increase in the urine chromium concentrations at all time points compared to the preoperative values, but the levels between two months and four years were not statistically significant. The mean whole blood cobalt concentrations at one and four years were 1.3 μg/L and 1.2 μg/L, respectively, and mean chromium concentrations were 2.4 μg/L and 1.1 μg/L, respectively. The differences between the preoperative and one-year levels for both cobalt and chromium were statistically significant (0.2 μg/L vs. 1.3 μg/L and 0.3 μg/L vs. 2.4 μg/L). Chromium levels at four years showed a significant reduction compared to 1-year levels (21.3 mg/L). For cobalt, the decrease in ion levels from one to four years was only marginal and not statistically significant.

At six years follow-up,[107] the declining trend seen in the urine output of cobalt continued (7.8 μg/day vs. 8.1 μg/day at four years), with a significantly lower output at six years than at six months ($P < .01$). In contrast, the median chromium output at six years was slightly, though not significantly, greater (3.9 μg/day vs. 3.6 μg/day at four years). Whole blood serum concentrations of cobalt and chromium continued a downward trend at six years, with levels of chromium still significantly lower than those after one year (1.11 μg/L vs. 2.41 μg/L). Both the daily output of metal ions and whole blood levels showed an early increase followed by a progressive decrease, which continued until six years. At six years, all patients were active and had well-functioning arthroplasties.

Metal-on-Metal Hip Resurfacing versus Metal-on-Metal Total Hip Arthroplasty
Direct comparisons of metal ion levels between MoM total HR and THA provide conflicting evidence as to which method, if either, results in greater metal ion concentrations, Table 14. Two of the studies reported higher Co and Cr concentrations in total HR patients compared with THA,[52, 108] and one study reported greater levels in the total HR group at early, but not late, follow-up.[109] Significantly higher concentrations of Co and Cr metal ions were reported for THA in one study,[64] while another found no differences in Co or Cr ion levels between the two groups.[110] Molybdenum concentrations were insignificant in all MoM groups in the studies in which this metal was considered. Little is known about the circulation levels of molybdenum in patients with MoM bearings and little is known about its toxicity.[111]

Clarke et al. compared metal ion levels at a median 16 months following large diameter total HR in 22 patients with a matched group of 22 patients following small diameter (28 mm) THA.[52] The median serum levels of Co and Cr were significantly greater following total HR versus THA (Co: 38 nmol/L versus 22 nmol/L; Cr: 53 nmol/L versus 19 nmol/L, respectively ($P < .001$)). Furthermore, the maximum serum levels of cobalt and chromium

recorded after resurfacing arthroplasty (165 nmol/L and 144 nmol/L, respectively) were twice those found following THA (87 nmol/L and 58 nmol/L, respectively). The upper limit of normal, physiological values for patients without implants is typically 5 nmol/L, and these results indicate significant increases above normal.

Witzleb et al. investigated the serum concentrations of cobalt, chromium, and molybdenum over two years in 111 patients following Birmingham total HR and in 74 patients following THA.[108] Implant-free patients (n = 130) were used as control subjects and had median chromium and cobalt levels of < 0.25 μg/L and 0.25 μg/L, respectively. In the unilaterally implanted THA patients, median chromium concentrations at three months, one year, and two years were 0.83 μg/L, 1.62 μg/L, and 1.22 μg/L, respectively. In the bilaterally implanted THA patients, concentrations were a median of 4.42 μg/L, 3.62 μg/L, and 2.50 μg/L, respectively. In total HR patients, levels rose from a median 1.96 μg/L to 4.20 μg/L and 5.12 μg/L, respectively. There was a significant difference between each patient's serum chromium concentration and the concentration of the control individuals. Compared to the concentrations in the patients who received unilateral implants, serum chromium levels in bilateral total HR patients were significantly higher at each assessment time (P = .003, P = .005, and P < 0.001, respectively). Similar differences were found in the bilateral THA patients at three and 24 months after surgery (P = .04 and P = .05, respectively). Cobalt concentrations in THA patients were highest at two years, with a median of 1.70 μg/L for the unilaterally implanted patients and 3.18 μg/L for the bilaterally implanted patients. In total HR patients, median levels rose from 2.17 μg/L at three months to 4.28 μg/L at two years after surgery. Serum cobalt levels differed significantly from the control levels except for bilaterally implanted THA patients at the two-year assessment. Levels in total HR patients exceeded the levels in unilaterally implanted THA patients significantly at each assessment point (P = .04, P = .002, P < .002). No statistically significant differences were found when comparing bilaterally implanted THA patients with unilaterally implanted patients and total HR patients. The median molybdenum serum level of the controls was 2.11 μg/L, and the levels of the patient groups showed no statistically significant difference versus the controls at any assessment point.
Antoniou et al. compared metal ion levels in the blood as well as oxidative stress markers (plasma levels of total antioxidants, peroxides, and nitrated proteins) following MoM total HR in 70 patients, MoM THA with either a 28 mm (n = 28) or 36 mm (n = 58) femoral head, and metal-on-polyethylene (MoP) THA in 18 patients. Forty pre-resurfacing patients were also used as a control group.[109] At six months postoperatively, median cobalt and chromium levels were significantly lower (P < 0.001 and P < .01, respectively) in the 36 mm MoM THA group (1.8 μg/L and 0.25 μg/L) compared with the 28 mm MoM THA group (2.5 μg/L and 0.35 μg/L) and the total HR group (2.3 μg/L and 0.50 μg/L). At one year, the levels of cobalt and chromium in all of the three groups with MoM bearings were significantly higher than those observed in the control and metal-on-polyethylene groups (P ≤ 0.001). However, neither metal level differed significantly between the three MoM groups at one year (2.3 μg/L vs. 2.6 μg/L vs. 2.4 μg/l, respectively). Median levels of molybdenum were not significantly different between groups, including the control group, at any time point. No significant difference in the plasma concentration of oxidative stress

markers was seen in the MoM groups as compared to controls both at six months and at one year. No correlation was found between these markers and the concentrations of either cobalt or chromium ions ($r^2 \leq 0.01$).

As reported in our data above, Garbuz et al. compared 48 total HR patients with 56 large diameter head THA patients and found that overall, patients receiving THA had elevated serum ions levels compared to the resurfacing arm.[64] In the THA group, compared to median preoperative levels (cobalt: 0.11 µg/L, chromium: 0.20 µg/L), cobalt levels at one year had increased 46-fold to a median of 5.09 µg/L, and chromium levels had increased 10.7-fold to a median of 2.14 µg/L. In the total HR group, median preoperative levels of serum cobalt and chromium were 0.13µg/L and 0.15 µg/L, respectively, and by one year median cobalt levels had increased 3.9-fold to a median of 0.51µg/L, and chromium levels had increased 5.4-fold to a median 0.81 µg/L. Intergroup comparisons at one year revealed that serum cobalt was 10-fold higher and serum chromium 2.6-fold higher than the resurfacing arm.

Moroni et al. compared serum concentrations of chromium, cobalt, and molybdenum between 20 patients who underwent large diameter (48 mm) total HR and 26 patients who received small diameter (28 mm) THA at a median follow-up of 25 months.[110] Forty-eight subjects with no prostheses were included as a control group. No differences were found in the levels of any of the three metals between the two MoM groups. As would be expected, the MoM hip resurfacing group had higher serum concentrations of chromium, cobalt, and molybdenum than the control group (2.30 ng/mL vs. 0.25 ng/mL, 1.40 ng/mL vs. 0.29 ng/mL, and 0.90 ng/mL vs. "less than detectable limit", respectively, ($P < .0001$)).

Table 14. Metal concentration in patients with MoM total HR or THA

Author	Specimen	Timing	Metal	HR	THA (28 mm)
Clarke (2003)	Serum	16 months	Co	38 nmol/L	22 nmol/L
			Cr	53 nmol/L	19 nmol/L
Witzleb (2006)	Serum	3 months	Co	2.17 µg/L	NR
			Cr	1.96 µg/L	0.83 µg/L
		24 months	Co	4.28 µg/L	1.70 µg/L
			Cr	5.12 µg/L	1.22 µg/L
Antoniou (2008)	Whole blood	6 months	Co	2.3 µg/L	2.5 µg/L
			Cr	0.50 µg/L	0.35 µg/L
		12 months	Co	2.4 µg/L	2.6 µg/L
			Cr	0.5 µg/L	0.6 µg/L
Garbuz (2009)	Serum	12 months	Co	0.51 µg/L	5.09 µg/L
			Cr	0.81 µg/L	2.14 µg/L
Moroni (2008)	Serum	25 months	Co	1.40 ng/mL	0.29 ng/mL
			Cr	2.3 ng/mL	0.25 ng/mL
			Mo	0.90 ng/mL	undetectable

NR = not reported.

Metal Ion Levels During Pregnancy
Data derived from a study that evaluated metal levels during pregnany confirmed that the placenta plays a modulatory role in the rate of metal transfer and that the transfer rate is different with different metals.

Ziaee et al. investigated whether elevated concentrations of cobalt and chromium ions in a pregnant woman's blood lead to raised levels of these metals in the umbilical cord blood of the baby by analyzing whole blood specimens from women who had undergone MoM total HR and a control group without metal implants.[103] In both groups, cobalt and chromium ions crossed the placenta. The mean concentrations of cobalt and chromium in the umbilical cords of the study group were comparatively lower than those of maternal levels, 0.839 μg/L versus 1.39 μg/L and 0.378 μg/L versus 1.28 μg/L, respectively, and the difference was significant for chromium ($P < 0.05$). The differences between the maternal and umbilical cord blood levels in the control group were marginal and not statistically significant. The mean cord blood level of cobalt in the study patients was significantly greater than that in the control group ($P < 0.01$) and although the mean umbilical cord blood chromium level was almost twice as high in the study patients (0.378 μg/L vs. 0.193 μg/L), this difference was not statistically significant ($P > 0.05$). In the controls, the rate of ion transfer across the placenta was in excess of 95% for both metals but only 29% for chromium and 60% for cobalt in the study patients, indicating that at low maternal metal ion levels there is a higher transfer rate than at higher maternal levels.

Effects of Metal Debris on Cellular and Biological Processes

Results from studies conducted in patients with MoM bearings have raised the possibility of serious adverse consequences from exposure to metal ion debris, including increased DNA and chromosomal changes, and immunological responses. A small number of studies have also raised the possibility that early osteolysis in patients with MoM THA is associated with abnormalities consistent with delayed hypersensitivity to metal wear debris.

DNA and chromosomal effects
In a study conducted by Davies et al. in THA patients, samples of synovial fluid retrieved at revision arthroplasty and cultured for human fibroblast cells revealed that all six samples from MoM implants and four of six samples from metal-on-polyethylene implants had statistically significant higher levels of DNA damage compared with control levels in human fibroblasts *in vitro*.[112]

We identified two studies that investigated ion levels in patients with cobalt-chromium-alloy THA implants. These studies reported a 2.5-fold increase and a 2- to 4-fold increase in aneuploidy (an increased number of chromosomes), as well as a 3.5-fold and 1.5-fold increase in chromosomal translocations in peripheral blood lymphocytes that could not be explained by confounding factors.[113, 114]

Hypersensitivity and immunological responses

Park et al. reported a higher rate of hypersensitivity reaction to cobalt in patients who underwent conventional MoM THA compared with controls (P = .031).[115] Histopathological and immunohistochemical analyses in these patients revealed perivascular accumulations of CD3-postive T-cells and CD68-positive macrophages, as well as bone-resorbing cytokines produced mainly by lymphocytes. Willert et al. similarly reported diffuse and perivascular infiltrates of T and B lymphocytes, macrophages with droplike inclusions, and eosinophilic granulocytes and necrosis from MoM articulations that had been retrieved at revision.[97] The prevalence of wear debris osteolysis and allergic reactions appears to be < 1% but longer-term, more inclusive data is needed to delineate the true prevalence of these complications.[47]

Hart et al. evaluated the relationship between MoM bearings, the levels of cobalt and chromium ions in whole blood, and the numbers of circulating lymphocytes by comparing 106 patients who underwent unilateral or bilateral MoM total HR with 33 and 25 patients who underwent metal-on-polyethylene (MoP) and ceramic-on-ceramic (CoC) THA, respectively.[116] Whole blood levels (parts per billion) of cobalt and chromium ions were significantly higher in the MoM groups compared to the non-metal-on-metal groups. The bilateral MoM hips had the highest rates at a median 2.45 and 2.35, respectively, followed by the unilateral MoM hip group (1.71 and 2.33), the MoP group (0.44 and 0.65), and the CoC group (0.21 and 0.32), (P < 0.001). Peripheral blood counts of all T-lymphocyte subtypes were considerably reduced in MoM hips as compared to the non-MoM groups. However, T-cytotoxic cells ($CD8^+$) were only subtype which was significantly reduced in both the unilateral MoM group (0.34) and the bilateral group (0.31) when each was compared to the MoP group (0.49) and the CoC group (0.50) (P = .024 vs. MoP and CoC, P = .024 vs MoP and 0.046 vs CoC, respectively). When all patients were analyzed together, the level of cobalt or chromium in the blood inversely correlated with the absolute counts of T- and B-lymphocytes. Long-term studies are needed to determine whether the moderate lymphopenia associated with "high" levels of cobalt or chromium after the use of MoM bearing is detrimental or even beneficial to longevity.

Ollivere et al. examined the rate and mechanism of early failure in 463 patients who underwent Birmingham total HR which is the most widely used of the current MoM resurfacing devices.[117] Overall, 12 patients (2.8%) were revised. Thigh and groin pain was the presenting symptom in seven patients, fracture in three, dislocation in two, and infection in one. Histological changes associated with metallosis and a response to wear debris was seen in nine of the 12 cases of failure. All seven of the patients revised for pain and two for fracture (one had no histological samples available) were diagnosed with aseptic lymphocytic-vasculitis-associated lesions (ALVAL), an abnormal tissue reaction associated with the release of metal ions. Soft-tissue and bone necrosis were also typical operative findings in these patients. MRIs were completed in three patients who presented with pain and fluid-filled soft-tissue tumors were identified in all three patients. The rate of metallosis related failure in this series was 1.9% at a mean 3.5 years and survival analysis indicated a rate of over 3% at five years.

Diseases

Cancer
The International Agency for Research on Cancer (IARC), has classified soluble cobalt as possibly carcinogenic and metallic chromium and chromium III compounds and implanted orthopaedic alloys as unclassifiable.[118] A review conducted by the IARC in 2000, which summarized the findings of 14 epidemological cohort studies after total knee or total hip replacement with various devices including MoM articulations, revealed inconsistent findings regarding the risk of malignancy following joint replacement surgery. Four studies suggested that there was an increased risk for specific types of cancers, including Hodgkin's disease, non-Hodgkin's lymphoma, leukemia, and renal cancer while the other studies were inconclusive. The overall incidence of cancer was reported to have decreased in all but one study which reported a slight increase. Currently, there is no consensus regarding safe exposure limits for these metals in hip arthroplasty.[119]

Occupational metal exposure such as at that of chromium has been linked, though weakly, to an increased risk of lung cancer[120] and concerns that metal-induced DNA damage, such as that from MoM wear particles, may lead to carcinogenesis are gaining consensus.[90] However, available data do not support a causal link between THA and the development of cancer and some studies have reported no increased cancer risk in patients with conventional MoM total hip devices.[121, 122] It is important to note that, in order to detect a rise in such adverse events, large numbers of patients would be required to be followed for several decades.[123] The effects of accumulating concentrations over time remain to be determined and continual monitoring of patients with MoM bearings is encouraged until a better understanding of the possible risks associated with metal ions in circulating blood is achieved.

Pseudotumors
Pandit et al. reported the clinical details of a subset of 17 patients (20 hips) with a mean age of 53 years who, following MoM total HR with various devices, all developed a soft-tissue mass or "pseudotumor" associated with the implant.[124] The mass was neither malignant nor infective. Histological examination was completed in 13 of the 20 hips. Common features noted included extensive necrosis of dense connective tissue within the pseudotumor that was sometimes associated with obvious cystic degeneration; metal wear particles contained within the tumors, though gross clinical metallosis was not seen in most cases; a scattered, focally heavy macrophage and lymphocytic infiltrate including lymphoid aggregates; and plasma cells and eosinophil or neutrophil polymorphs. It is important to note that these patients represented a subset of patients from a series of 1300 hip resurfacing procedures completed at the same institution, and the authors estimate the incidence of psuedotumors to be approximately 1% at five years. Further studies are necessary to determine the true incidence of pseudotumors.

There is some evidence to suggest that patients who have a revision from a total HR to a THA as a result of a pseudotumor have poorer outcomes than those who have revision for other reasons. Grammatopolous et al. reported worse Oxford hip scores following revision

for pseudotumor than for fracture or for other causes.[89] The clinical outcome of revision for pseudotumor was also significantly worse than the outcome of matched primary total hip replacements. By contrast, the outcome for fracture and other causes was not significantly different from that of matched primary total hip replacements. After revision for pseudotumor there were three cases of recurrent dislocation, three of femoral nerve palsy, one of femoral artery stenosis and two of component loosening. The authors concluded that outcome of revision for pseudotumor is poor, while outcome of resurfacing revision for other causes is good.

Possible factors associated with increased levels of metal debris

The size of the femoral head and the position of the acetabular component of MoM total HR and THA prostheses have been investigated as possible factors affecting the level of metal ions following these procedures.

Femoral head size

The relationship between the size of the femoral head and metal ion concentration is not entirely clear, but there does appear to be a trend toward higher levels of metal ions in patients with smaller diameter femoral components, Table 15. Two studies reported an inverse relationship between femoral implant size and chromium and cobalt ion concentrations,[102, 110] one study found an inverse correlation between the component size and the levels of chromium, but not the levels of cobalt ions,[105] and one larger study reported no association between the head size and the chromium, cobalt, or molybdenum serum concentrations.[108]

Table 15. Correlation between head size and metal ion concentrations.

Author	Chromium	Cobalt	Molybdenum
Langton (2008)	r = -0.328; *P* = .004	r = -0.315; *P* = .006	NR
Moroni (2008)	r = -0.64; *P* = .002	r = -0.51; *P* = .01	NR
Vendittoli (2007)	r = -.298; *P* = .018	no correlation	NR
Witzleb (2006)	no correlation	no correlation	no correlation

NR = not reported

Acetabular positioning

Acetabular components with high inclination angles have been shown to demonstrate increased wear secondary to rim loading.[42] We identified four studies that looked at acetabular positioning, and similar to femoral head size, the results are mixed. One study found a positive correlation between the inclination of the acetabular component and the whole blood serum concentrations of chromium (r = 0.372 (*P* = .01) and cobalt (r = 0.439 (*P* < .001)) in a group with the smaller (≤ 51 mm) femoral component (as compared to the larger component group, ≥ 53 mm) in a series of MoM total HR arthroplasties.[102] Likewise, the anteversion of the acetabular component correlated significantly with the levels of metal ions in whole blood (r = 0.338 (*P* = .008) and r = 0.330 (*P* = .01), respectively). These correlations were not reflected in the group with the larger femoral component, however.

De Haan et al. also found a positive correlation between the inclination angle of the acetabular component and serum levels of chromium and cobalt in patients treated with a MoM resurfacing device.[125] Significantly higher levels of metal ions were found in patients with steeply inclined components (*P* = .002 for chromium, *P* = .003 for cobalt) and an arc of cover of less than 10 mm was highly correlated with a greater risk of high concentrations of serum metal ions (*P* < 0.001).

Allan et al. report persistent outlier elevations of chromium (13.88 µg/L to 69.95 µg/L) and cobalt (37.89 µg/L to 124.94 µg/L) in two subjects throughout all follow-up time points up to three years after hip resurfacing with the Cormet 2000 device.[106] Both subjects demonstrated radiographic evidence of excessive cup tilt (83° and 60°), but showed no clinically significant problems and maintained good functional mobility.

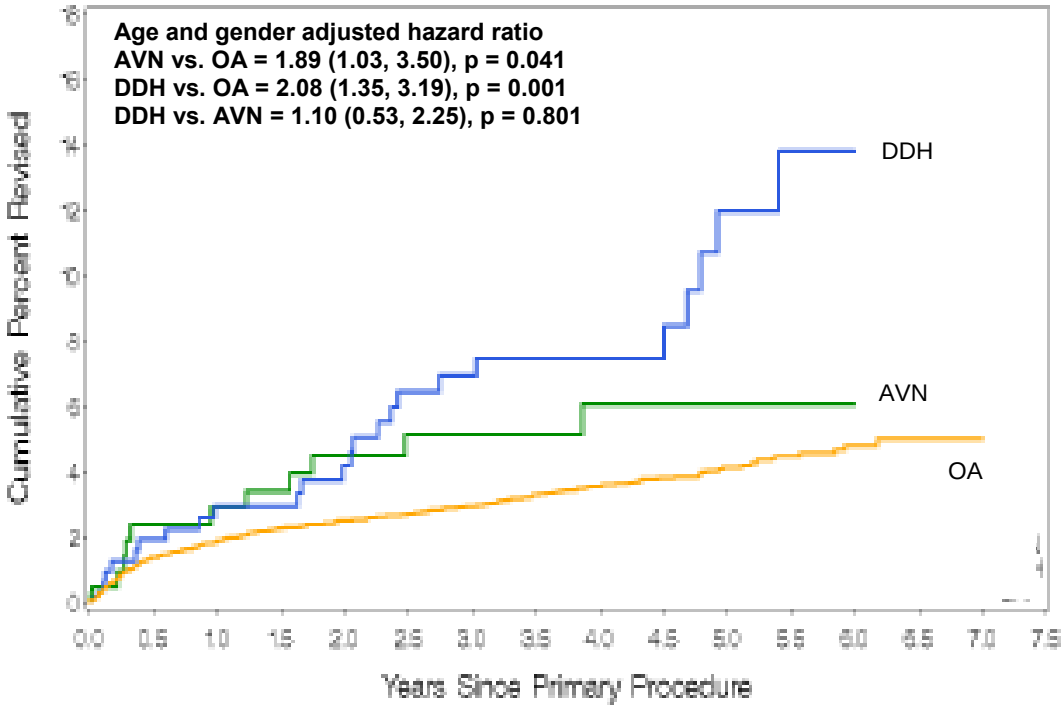
Finally, in a large study, Witzleb et al. found no correlations between the position of the acetabular cup and the levels of chromium, cobalt, and molybdenum.[108]

4.3. Key Question 3: Is there evidence of differential efficacy or safety issues with use of hip resurfacing?

Dysplasia versus other arthritic etiologies

Australian National Joint Replacement Registry
Three diagnoses were compared for total HR in the Australian National Joint Replacement Registry. The Registry reports a significantly higher risk of revision with resurfacing procedures undertaken for dysplasia (hazard ratio = 2.08; 95% CI 1.35, 3.19) compared with osteoarthritis, Figure 9. The five-year cumulative percent revision for dysplasia is four times greater in those receiving total HR compared with THA, 12% versus 3%. [74]

Figure 9. Cumulative percent revision of primary hip resurfacing by primary diagnosis.



AVN = avascular necrosis.
DDH = developmental dysplasia of the hip.
OA = osteoarthritis.
Figure used by permission from the AOA National Joint Replacement Registry, Australia.

Published Cohort Studies

McBryde (2008)

McBryde et al.[126] published the results of a retrospective cohort study in which the outcomes of 85 patients (96 hips) with osteoarthritis secondary to developmental dysplasia (Crowe I (41.7%), II (22.9%), III (12.5%), and IV (5.2%)) were compared with those of 93 matched patients (96 hips) with primary osteoarthritis (OA) following Birmingham hip resurfacing. The mean age of patients was 45 years, and 78% of patients were female. Although patients were matched, the mean age of patients in the dysplasia group was 43 (range: 14–65) years compared to a mean age of 47 (range: 22–76) years in the OA group. Additionally, previous acetabular and/or femoral surgery had been performed in 10 and 6 hips, respectively, in the dysplasia group, while no hips in the OA group had undergone prior surgery. A dysplasia cup was used in 34 of the hips in the dysplasia group. Patients were followed for a mean of 4.5 years.

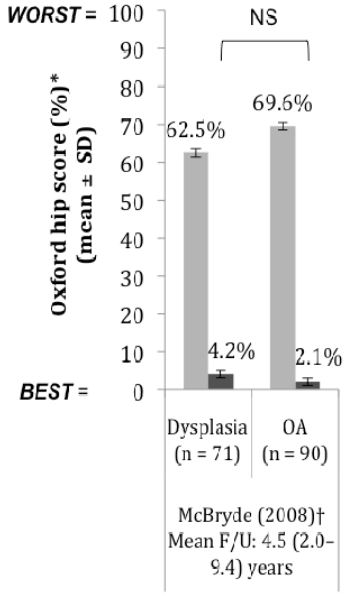
No significant difference was found in the postoperative Oxford hip scores between groups, which the authors expressed as a percentage (lower percentage indicating a better outcome): 4.2% in the dysplasia group (interquartile range: 0–6.3%) (reported by 71 patients) versus 2.1% in the OA group (interquartile range: 0–16.7%) (reported by 90 patients) (Figure 10a).

Five-year survival rates were based on a subset of patients, and were significantly higher in the OA group (100%; 95% CI: 100, 100%; N = 36 hips) compared with the dysplasia group (96.7% ; 95% CI: 90, 100%; N = 26 hips) (Figure 10c). No hips in the OA group required revision at any time during follow-up (mean 4.5 (2.0–9.4 years)), while five hips (5.2%) in the dysplasia group were revised. Causes of revision included failure of the acetabular component in four hips (4.2%) and femoral neck fracture in one hip (1.0%). Other reported complications included thinning of the femoral neck (dysplasia: 7.7% of the 78 unrevised hips, OA: 0%) and loosening of the acetabular (dysplasia: 1.3%, OA: 0%) and femoral (0% in both groups) components (Figure 10d).

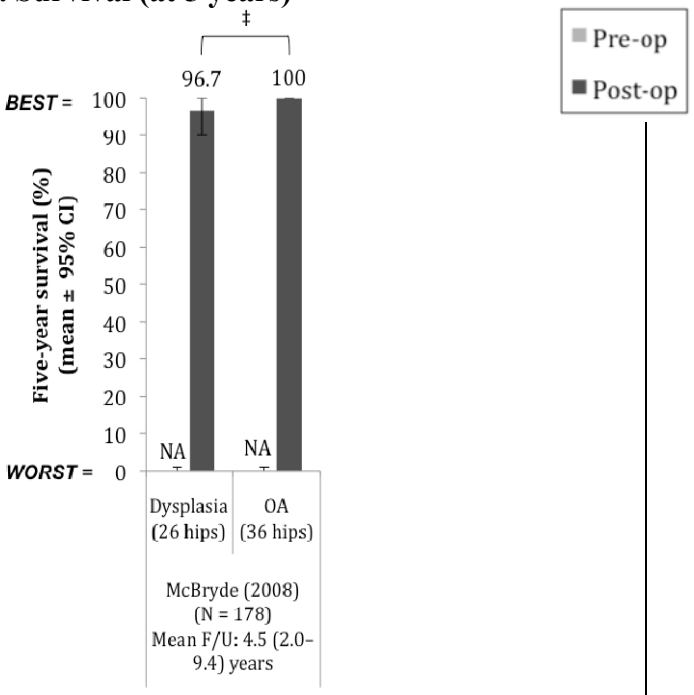
Amstutz (2008)
Amstutz et al.[76] reported on a series of 90 patients with 103 hips who underwent hip resurfacing for osteoarthritis secondary to developmental dysplasia (Crowe I (94%), II (4%), or III (2%)). Nine patients (10 hips) had undergone prior surgery, including pelvic osteotomies (6 hips), femoral osteotomies (7 hips), and a shelf procedure (1 hips). All patients received the Conserve Plus prosthesis. The mean UCLA activity score in these patients was compared to that of similarly treated patients (897 hips) with other indications for surgery, including idiopathic osteoarthritis (77.6%), osteonecrosis (9.3%), posttraumatic osteoarthritis (4.8%), and others. No other details on these patients were reported, and this series should not be interpreted as a cohort study. Patients with developmental dysplasia had a significantly lower mean postoperative UCLA activity score (7.0, range of 2 to 10) than patients with other preoperative diagnoses (7.5, range not reported) ($P = .003$) (Figure 10b). While these results suggest that patients with developmental dysplasia have lower activity outcome scores than other patients, the paucity of data on the control group of patients here makes the results difficult to interpret. No other reported outcomes of interest for the dysplasia patients were compared to outcomes for patients with other indications for surgery, so they are not included here.

Figure 10. Dysplasia versus other arthritic etiologies: functional/activity measures and survival rates following hip resurfacing from two retrospective studies.

a. Oxford hip score (%)*



b. Survival (at 5 years)



c. Complications

	Dysplasia (% hips)	OA (% hips)
Revision	5.2	0
Femoral neck fracture	1.0	0
Thinning of the femoral neck	7.7	0
Loosening of the femoral component	0	0
Loosening of the acetabular component	0	0
Failure of the acetabular component	4.2	0

NS: not statistically significant ($P \geq .05$).
SD: standard deviation.
* McBryde reported Oxford hip scores as percentages, presumably representing the percentage of the total score which normally ranges from 12–60, with lower scores indicating better outcome.
† Standard deviation not reported.
‡ $P < .05$.

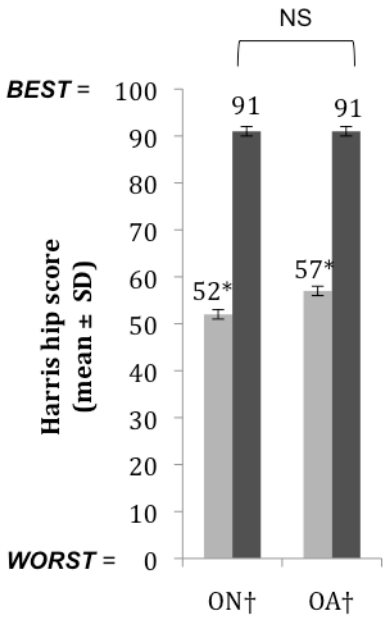
Osteonecrosis versus osteoarthritis

Australian National Joint Replacement Registry
As shown above in Figure 9, preoperative diagnoses of avascular necrosis (AVN) and osteoarthritis were compared for total HR in the Australian National Joint Replacement Registry. The Registry reports a significantly higher risk of revision with resurfacing procedures undertaken for AVN (hazard ratio = 1.89; 95% CI 1.03, 3.5) compared with osteoarthritis, Figure 9. The five-year cumulative percent revision for AVN is approximately 2-times greater in those receiving total HR compared with THA, 6.1% versus 3%.[74]

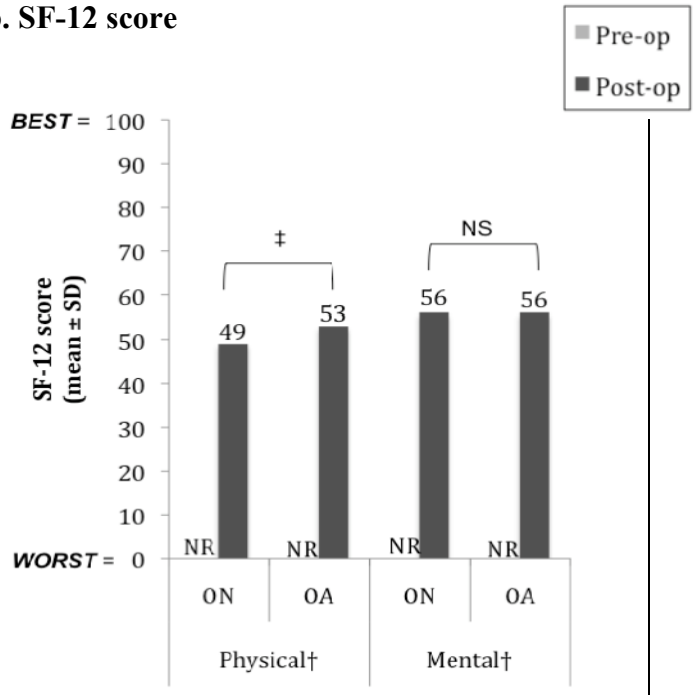
Mont (2006)
Mont et al.[127] retrospectively reviewed outcomes of 78 patients with a preoperative diagnosis of either osteonecrosis (ON) (36 patients with 42 hips) or osteoarthritis (OA) (41 matched patients with 42 hips) at a mean of 3.4 years following hip resurfacing with the Conserve Plus prosthesis as part of an FDA IDE study. Mean patient age was 42 years, and 68.8% were male. Although patients with osteoarthritis had a significantly higher mean preoperative HHS (57 versus 52 (ON), (P = .008)), at final follow-up, both groups had the same mean HHS (91) (P = .941) (Figure 11a). Patients with osteoarthritis had a significantly higher mean postoperative SF-12 physical component score than those with osteonecrosis (53 versus 49, respectively, (P = .008)), while there was no significant difference in the mean postoperative SF-12 mental component scores between the groups (56 in each group) (Figure 11b). At final follow-up, there was no significant difference in device survival between groups (ON: 94.5% versus OA: 95.2%, (P = .977)) (Figure 11c). Similarly, revision rates were the same in both groups (4.8%); reasons for revision in the ON group were femoral neck fracture in one hip and aseptic loosening of the femoral component in the other, while two hips in the OA group were revised following femoral neck fracture (Figure 11d). Complications in the OA and ON groups included: femoral neck fracture (4.8%, 2.4%, respectively), acetabular cup migration (2.4%, 0%), heterotopic ossification (9.5%, 14.3%), and nonprogressive femoral (0%, 2.4%) and acetabular (4.8%, 7.1%) radiolucencies. Overall, the short-term outcomes for patients with a preoperative diagnosis of osteonecrosis were similar to those that occur in osteoarthritic patients.

Figure 11. Osteonecrosis (ON) versus osteoarthritis (AO): Functional and quality of life outcomes measures scores and survival rates following hip resurfacing from one retrospective cohort study (Mont 2006).

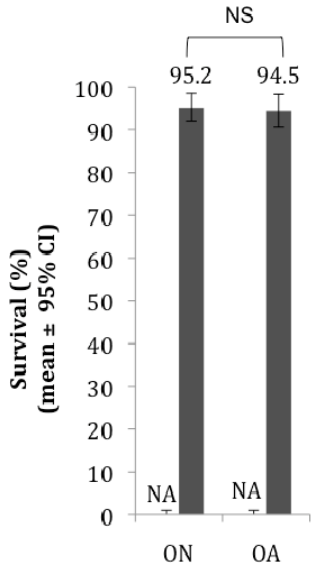
a. Harris hip score



b. SF-12 score



c. Survival (at final follow-up, mean of 3.4 (2–5.1) years)



d. Complications

	OA (% hips)	ON (% hips)
Revision	4.8	4.8
Femoral neck fracture	4.8	2.4
Acetabular cup migration	2.4	0
Heterotopic ossification	9.5	14.3

NS: not statistically significant ($P \geq .05$)
SD: standard deviation
* There was a statistically significant difference in preoperative values between groups ($P < .05$).

† Standard deviation not reported.
‡ $P < .05$

Female versus male

Registry Data
The Australian National Joint Replacement Registry[74] reports that the 5-year cumulative percent revision risk for females is 2.5 times higher than males, 6.5% versus 2.6%, Figure 12. The Registry recently reported an inverse relationship between femoral component head size and the risk of revision. As the head size increases, the five year cumulative percent revision decreases. After adjusting for femoral component head size, the Registry found no significant difference in the risk of revision between males and females, Figure 13. The Swedish Registry reported a two-fold increased risk of revision for females.[1] The National Joint Registry from the UK reported a 57% increased risk of revision after three years, 3.6% (95% CI 2.7%, 4.8%) for females and 2.3% (95% CI 1.9%, 3.3%) for males.[75]

Figure 12. Cumulative percent revision of hip resurfacing by gender (primary diagnosis OA excluding infection).

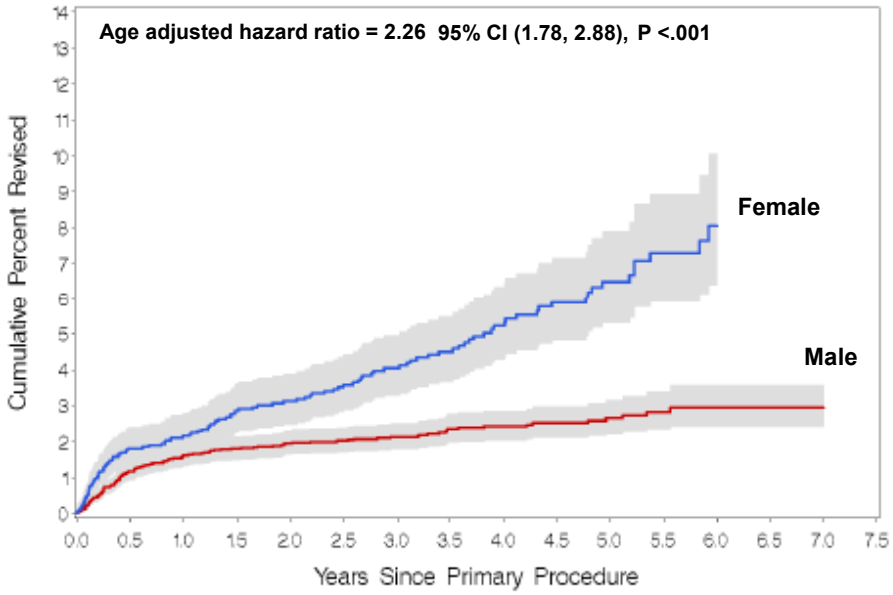


Figure used by permission from the AOA National Joint Replacement Registry, Australia.

Figure 13. Cumulative percent revision of hip resurfacing by gender and femoral component head size (primary diagnosis OA excluding infection).

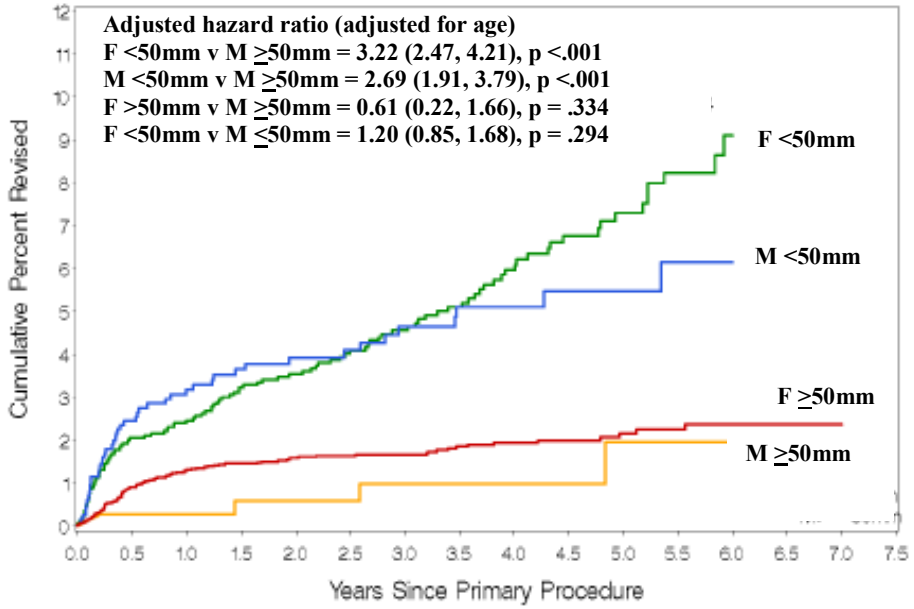


Figure used by permission from the AOA National Joint Replacement Registry, Australia.

Published cohort studies

Amstutz (2004): risk of metaphyseal stem radiolucencies

Amstutz et al.[21] reported outcomes of 355 patients with 400 hips following hip resurfacing with the FDA-510k investigational Conserve Plus prosthesis. Mean patient age was 48.2 years, and 73% of patients were male; patients were followed for an average of 3.5 years. Using the Cox multivariate proportional-hazards model, the authors found that females were at significantly higher risk of developing metaphyseal stem radiolucencies than males, with a hazard ratio of 3.1 (95% confidence interval was 1.4 to 6.3) ($P = .005$). In addition, males with a smaller (≤ 2 mm) component size had an increased risk of metaphyseal stem radiolucencies, hazard ratio of 1.31 (95% CI: 1.09, 1.59, $P = .005$). However, patients with this type of radiolucency had similar pain and activity scores as the remaining patients. Radiographs of 372 hips were evaluated by an independent observer, and a radiolucency was considered to be present if there were lucencies in all regions of the metaphyseal stem or migration of the metaphyseal stem. Notably, no radiolucencies were detected in any of the 59 hips in which the metaphyseal stem was cemented. Revised hips were excluded from this analysis (12 hips).

Ollivere (2009): risk of metallosis

Ollivere et al.[117] reviewed a series of 463 consecutive patients who underwent Birmingham hip resurfacing. The mean age of patients was 56 years, and 66.3% were male. Patients were followed for an average of 3.6 years following surgery, and this study focused on the

twelve patients (with 13 hips) who underwent revision. Seven of these patients, all of whom underwent revision for unexplained pain and two of whom had late femoral neck fractures, had evidence of metallosis (aseptic lymphocytic-vasculitis-associated lesions (ALVAL)) upon histopathological examination by one of the authors. Females had a significantly higher risk of metallosis than their male counterparts, with a relative risk of 4.94 (95% CI, 1.33 to 18.31, $P = .018$). Additionally, the authors reported that those that required revision had a significantly lower mean femoral component size (44 mm) than the patients in the entire series (48 mm) ($P = .002$). However, more patients with metallosis should be evaluated to determine whether females are at increased risk for developing this potentially adverse immunological reaction.

Surface Arthroplasty Risk Index (SARI)

Beaule (2004)

Beaule et al.[128] developed the Surface Arthroplasty Risk Index (SARI) in 2004 as a way to preoperatively evaluate a patient's risk for radiologic changes or revision surgery following modern hybrid hip resurfacing. The SARI was based on multivariate logistic regression analysis of potential risk factors, four of which were found to be associated with significantly increased risk of total HR failure and were assigned one or two points based on the value of the corresponding odds ratio: UCLA activity score > 6 (1 point); previous arthroplasty (1 point); weight less than 180 pounds (2 points); and femoral cysts with a diameter of more than 1 cm (2 points). SARI was validated in this series of 83 MoM total HR patients, all of whom were under the age of 40, with a minimum of two years follow-up. Patients with problematic hips (13 hips, including three conversions and 10 significant radiological changes on femoral side) had a significantly higher mean SARI score (4.7, all ≥ 3) than those without hip problems (79 hips) (mean SARI score: 2.6) ($P = .001$). A SARI score of four or higher was associated with a twelve-fold increase in early complications compared to patients with a score of three or less.

Amstutz (2004)

In the case series of 355 patients with 400 hips described above, Amstutz et al.[21] noted that a SARI score greater than 3 was associated with earlier time to revision ($P = .004$). Similarly, patients with a SARI score > 3 had a lower four-year survival (88.8% (95% CI, 80.0 to 97.6%)) than those with a SARI score ≤ 3 (97% (95% CI, 94 to 100%)). The risk of femoral radiolucency was also 4.2-fold higher in patients with a SARI score higher than 3 than in patients with a lower SARI score.

Obesity

We identified two studies that assessed whether high body mass index (BMI) was correlated with poorer outcomes following hip resurfacing.

Le Duff (2007) (Figure 14)

Le Duff et al.[129] evaluated the relationship between BMI and patient outcomes following hip resurfacing with the Conserve Plus prosthetic system. A total of 125 patients (144

hips) with a BMI ≥ 30 (obese) were included from a series of 1000 procedures; an additional 531 patients (626 hips) with a BMI < 30 with similar preoperative diagnoses were selected from this series to serve as the control group. Mean patient age was 49.4 years, and 74.7% of patients were male. Patients were followed for an average of 5.9 years. Patient demographics were similar between the groups, although not surprisingly, patients with a higher BMI had a higher mean body weight (P = .0001) as well as lower mean UCLA activity scores (P = .0075) and SF-12 physical component scores (P = .0003) preoperatively. These confounding factors were not controlled for and make the data somewhat more difficult to interpret.

The obese group had a significantly lower mean postoperative HHS compared to that of the control group (90.6 versus 93.8, respectively, (P = .0003)). Although patients in the obese group had a significantly lower postoperative SF-12 physical component score (49.3) than did those in the control group (51.4) (P = .0129), it is possible that this 2.1-point difference may be attributed to the similar 2.8-point difference in the corresponding preoperative scores (30.1 versus 32.9, respectively, (P = .0003)). No significant differences between groups were identified in the preoperative or postoperative SF-12 mental component scores. The mean UCLA activity score was significantly lower in patients with a BMI ≥ 30 than in those with a lower BMI both preoperatively (4.3 versus 4.7, respectively, (P = .0075)) and postoperatively (7.1 versus 7.6, respectively (P = .0021)), and the point difference between groups was similar at both time points (0.4 preoperatively and 0.5 points postoperatively), Figure 14a-c.

Interestingly, patients in the obese group had a significantly higher five-year prosthesis survival rate (98.6% (95% CI: 94.5%, 99.7%)) than those in the control group (93.6% (95% CI: 90.4, 95.8%); P = .0401), Figure 14d. When patients were further subdivided based on their BMI, this trend continued: BMI < 25 (number of patients not reported) (90.6%; 95% CI; 84.0%, 94.5%), BMI 25–29 (number of patients not reported) (95.3%; 95% CI: 91.2, 97.5%), and BMI ≥ 30 (98.6%; 95% CI: 94.5%, 99.7%). Five-year survival rates in the obese group were significantly higher than in the BMI < 25 group (P = .013). Correspondingly, mean SARI scores were lower (indicating less risk) in obese versus control patients (1.5 versus 2.6 (BMI < 30), P < .0001). This result is likely attributable to differences in patient weight and possibly activity level between the groups. The SARI will be two points higher in patients that weigh less than 180 lbs, and patients in the obese group had a mean weight of 231 lbs (range: 163–362 lbs) while those in the control group had a mean weight of 173 lbs (range: 93–249 lbs). The SARI also assigns one point to patients with a preoperative UCLA activity score greater than 6 points. Although obese patients had a significantly lower preoperative UCLA activity score (4.3; range: 1–9) than their control counterparts (4.7 (range: 1–10; P = .0021), it is not clear that this difference would have a significant impact on SARI scores.

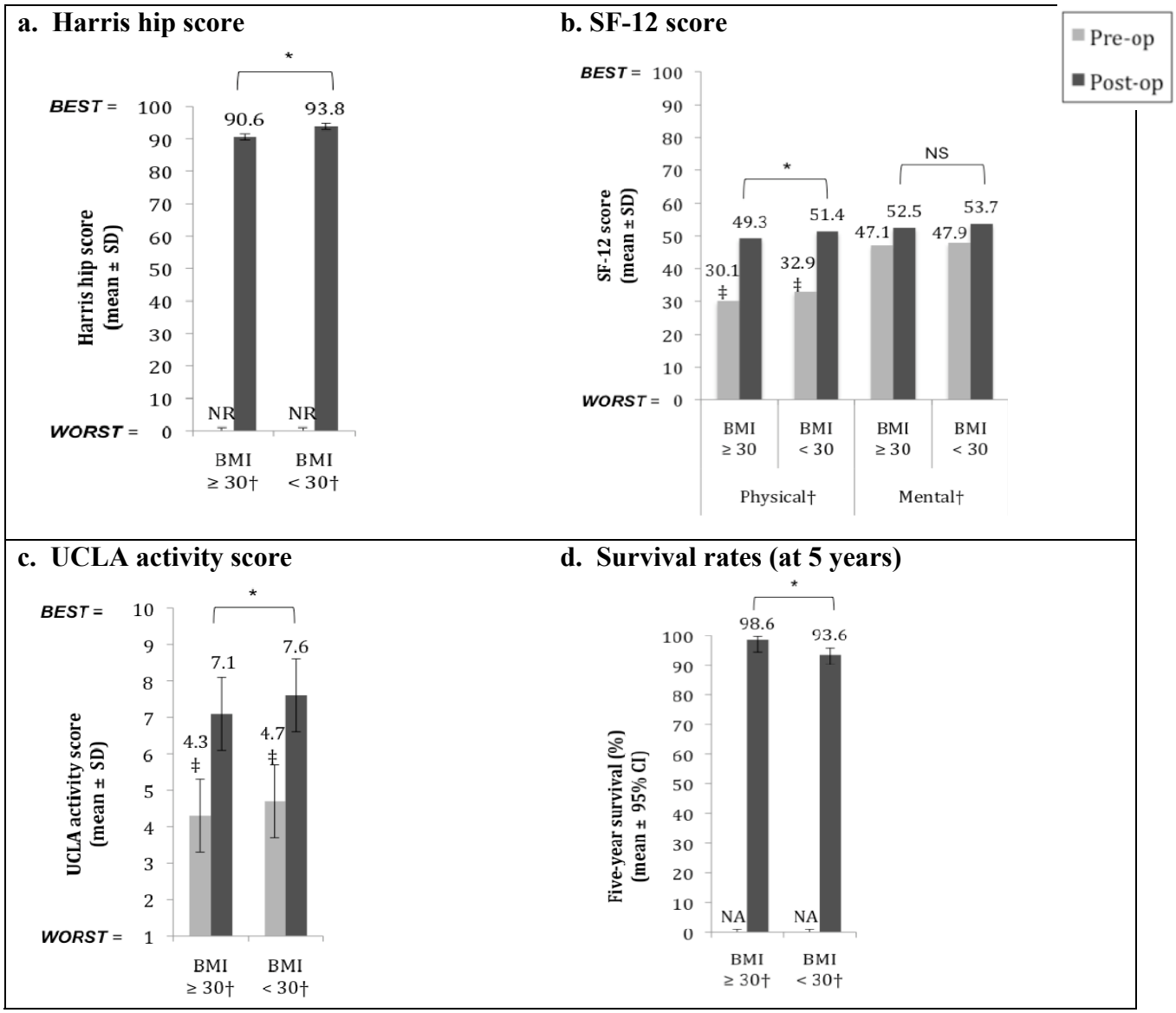
The obese group had a revision rate of 8.2% (two hips) due to femoral neck fracture in one hip and protrustion of the acetabular component into the medial acetabular wall in the other; the control group had a lower revision rate of 5.0% (31 hips) due to loosening of the femoral component (20 hips), femoral neck fracture (7 hips), infection (2 hips), recurrent

subluxations (1 hip), and an undetermined cause (1 hip revised at different location) (Figure 14e). The rate of femoral neck fractures was slightly higher in the control group (1.1% versus 0.7% of hips in the obese group). Similarly, loosening of the femoral component occurred at a higher rate in the control group (3.2% versus 0% of hips in the obese group), while there were no cases of loosening of the acetabular component in either group (Figure 14e). There was no significant difference in the rate of other complications (infections, dislocations, femoral nerve palsy, hematoma or bleeding, thromboembolic events, and radiolucencies between the groups (4.2% (BMI ≥ 30) versus 6.1% (BMI < 30, $P = .389$).

Ollivere (2009)

In the case series of 463 consecutive Birmingham hip resurfacing patients discussed above (female gender section) Ollivere et al.[117] found that the twelve patients who required revision had a significantly higher mean BMI (30.4; 95% CI, 29.3, 31.6) than all patients in the series (27.7; 95% CI, 27.5, 27.9); $P = .034$). However, a larger group of patients with revised hips needs to be evaluated in order to determine whether obese patients are at higher risk of failure.

Figure 14. Functional, quality of life, and activity measure scores, and survival and complication rates following hip resurfacing in patients with a BMI ≥ 30 compared to patients with a BMI < 30 from one retrospective cohort study (Le Duff (2007)).



e. Complications

	BMI ≥ 30 (% hips)	BMI < 30 (% hips)
Revision	8.2	5.0
Femoral neck fracture	0.7	1.1
Loosening of femoral component	0	3.2
Loosening of acetabular component	0	0

NS: not statistically significant ($P \geq .05$)
SD: standard deviation
* $P < .05$
† Standard deviation not reported.
‡ There was a statistically significant difference in preoperative values between groups ($P < .05$).

4.4. What is the evidence of cost implications and cost effectiveness of hip resurfacing?

Overall, there is limited evidence on the cost-effectiveness of hip resurfacing. We found two previous HTAs and two published articles that address the economic implications of hip resurfacing, Tables 11 and 12.

Evidence from other HTAs.
The Ontario HTA reports only on the McKenzie study, which we describe here in more detail (see below). They also conducted a very brief budget impact estimate for a Canadian setting, which includes only the cost per patient of hip resurfacing surgery. We did not include their impact estimate since it is not a complete economic evaluation and does not include any estimates for a US market.

The NICE HTA (Vale 2002) notes the lack of economic evaluations on hip resurfacing, but does give a fairly extensive assessment of the one "relatively complete economic evaluation" industry submission it received (Midland Medical Technologies, MMT) as part of the technology assessment process. The MMT study does not appear to have been subsequently published in the peer-reviewed literature. Details of the study, as reviewed by the HTA, are in Tables 11 and 12.

The MMT submission included a cost-utility analysis submitted via spreadsheet. It compared Birmingham hip resurfacing (BHR) to either total hip arthroplasty or watchful waiting. As requested by NICE, the study took a health system perspective. The analysis was estimated for a hypothetical cohort of 1000 candidates for hip replacement at 5 10, 15, and 20 years post-procedure, with focus on people age 45-65 based on the premise that THA was the superior option for people over 65. BHR effectiveness data were taken from internal industry data on 1693 BHRs conducted by four surgeons with limited follow-up for four years (complete follow-up data available for one percent, or 21 patients at four years); cost data were taken from NHS estimates and from the published literature; utility estimates were also from the published literature. Sensitivity analyses were conducted on varying levels of revision rate, cost, and QOL. Although Vale et al found this to be a reasonably well-conducted economic evaluation, they described several concerns that give reason for caution in the interpretation of the results—mainly the lack of long-term follow up data on BHR, and the model's assumption that patients do not exit watchful waiting for hip replacement but only for death.

Economic studies on hip resurfacing
We found two published, peer-reviewed articles on the economic impact of hip resurfacing. One (McKenzie 2003[130]) is a well-conducted economic evaluation whose main limitation is the paucity of clinical data available at that time on hip resurfacing, especially on revision rates. Weighted QHES score was 100 [possible score 0 (worst) to 100 (best)] for this study.

The McKenzie study was conducted in the UK on behalf of NICE. They conducted a cost utility analysis using a Markov model to integrate cost and outcomes of MoM hip resurfacing compared to either immediate THA or watchful waiting followed by THA. Taking a UK health service perspective focusing only on direct medical costs, they created two separate models based on age of entry for younger and older "typical" patients with advanced hip disease. Costs were taken from literature and interviews with manufacturers, clinical data was from published literature and expert opinion, and utilities were from published literature. The main strengths of the study are the use of a cost utility model addressing several alternative clinical pathways and the 20-year time horizon. The main limitations are the use of expert opinion for some cost and clinical pathway inputs and the general lack of data on the effectiveness of BHR, especially revision rates, but overall it is a higher-quality economic study, and the authors' conclusions are commensurate with the quality of the data available.

The other study (Buckland 2008) was brought to our attention by [Smith & Nephew, Inc.] from a journal which does not appear to be indexed by Medline or EMBASE. We have included it as it provides some additional context and more recent data on the cost-effectiveness of hip resurfacing. Weighted QHES score was 67 [possible score 0 (worst) to 100 (best)] for this study.

Buckland and colleagues (2008) conducted a cost consequences study, which provides costs and QALYs separately. It takes a US health insurance payer perspective, comparing early hip resurfacing to five years of conservative management followed by THA. They provide estimates for a hypothetical population of people with moderate to severe hip disease at several age groups: 45-49, 50-54, and 55-59. Costs were estimated from Medicare fee schedules, average wholesale price for medications, and expert opinion; clinical pathways were determined by expert opinion; utility scores were based on published pain-related health states; and revision rates were from the Swedish National Hip Arthroplasty Register and the Oswestry Outcome Centre registry, which has 8-year follow up data on almost 5000 people receiving hip resurfacing. The sensitivity analysis is not clearly described but appears to be a threshold-type analysis of individual model inputs, including revision rate, cost, and discount rate. Overall, the quality of this study is undermined since many of the methods were not clearly described and the scope was relatively limited.

Results

The MMT submission described by Vale (2002) found that the cost utility of BHR increased dramatically over time. The cost per QALY at five years compared to THA was estimated to be £13,125, while BHR was dominant at 20 years. Compared to watchful waiting, Cost per QALY was £1101 at 5 years and BHR dominant (less costly, more effective) at 20 years. The improvement in QALYs for BHR was small at each time point and based on the assumption of a higher revision rate for THA. The submission indicated a "worst-case scenario" where THA revision rates equaled BHR revision rate; in this scenario THA would be dominant (less costly with same QALY improvements). Sensitivity analysis found that the break-even point of equal cost at 20 years for BHR compared with THA was if BHR revision rates were 85% of THA rates for a 55 year-old patient.

McKenzie found that that MoM hip resurfacing dominated the watchful waiting option at all time points in the younger population. Compared with THA, THA was found to be the dominant option in both the younger and older populations modeled. Hip resurfacing revision rate was influential in the model results, suggesting that with increasing THA revision rate, hip resurfacing ceases to be dominated when its revision rate is 80% of the THA rate. The model was not sensitive to prostheses cost up to 300% of the base case or variations in utility estimates up to 0.97, but was to watchful waiting costs (up to £620). The authors conclude that MoM hip resurfacing warrants further study given the lack of long-term data on hip resurfacing effectiveness, especially given how influential it was in model results.

Buckland (2008) found that for all age groups, immediate hip resurfacing was dominant over conservative management followed by THA. This study used implant survival rather than revision rate; e.g., a 81.6% implant survival at 13 years follow-up for a 50-59 year old, compared with 97.0% implant **survival** at eight years follow-up for a 50-54 year old. To illustrate further, consider the 50-54 age group, modeled to age 65, for conservative treatment. The cost per patient was $22,160, and QALYs per patient were 10.03, compared to immediate total HR, which was both less costly ($17,144 per patient) and saved more QALYs (11.51). Sensitivity analysis suggested that revision rates do not change the overall results, even at equal revision rates for hip resurfacing and THA. The cost of drug treatment associated with conservative management would need to be between 26% and 42% of the estimated modeled, and the discount rate would need to be 10% or higher to change the study conclusions.

Discussion

Overall, there is limited evidence to inform a strong recommendation about the economic value of hip resurfacing. It appears that the most convincing evidence of cost-effectiveness of hip resurfacing is in patients under age 65 and that there is little evidence of cost-effectiveness for extended watchful waiting/conservative management. Further study of the value of hip resurfacing is warranted, especially in the context of emerging body evidence about its clinical value that was unavailable at the time of two of the three studies we examined.

Although each study provides some very interesting context and data, they differed from one another enough to preclude strong recommendations. The following are take-away points from synthesis of the three studies and directly inform recommendations for future economic evaluations that would provide more definitive evidence:

- Revision rate of hip resurfacing appears to be an influential factor in all studies we reviewed. As such, the most long-term, comprehensive, and highest-quality follow-up data on hip resurfacing revision rates is crucial to understanding the economic value of hip resurfacing. Any potential subgroups (besides age) likely to benefit the most from this intervention would also be useful, including pre-surgical health and activity levels.

- No studies used a societal perspective. Considering the likely dramatic impact of hip disease on productivity, out of pocket costs, and quality of life of degenerative hip disease especially in younger people, a thorough economic evaluation would take a societal perspective and include factors such as pain relief, adverse effects from therapy, productivity, functional status, health-related quality of life and such as out-of-pocket costs for subsequent diagnostic or interventional costs, rehabilitation, lost productivity.

- All studies used different comparators and forms of hip resurfacing, describing different clinical pathways—including one study that assumed that people choosing watchful waiting never proceed on to receive a hip replacement. As such it is difficult to compare the results of the studies as all made different assumptions about the costs and outcomes associated with conservative management or the exact clinical pathway through which patients receive THA or hip resurfacing.

Conclusion

Although further study is necessary to include more current data, there is currently insufficient evidence to warrant a conclusion about the economic value of hip resurfacing in a US setting. In particular, the estimates for revision used in these studies are not current and don't appear to match the contemporary data.

Table 16. Summary demographic information for economic studies.

Study	**Design**	**Data sources and population**	Model inputs	**Primary strengths and limitations**
McKenzie 2003	Cost utility analysis Markov model: costs and outcomes for 20 cycles at 1 cycle/year; five year bands Intervention: Metal on metal hip resurfacing (MoM) Comparators: Total hip replacement (THA); watchful waiting (WW) plus THA UK health service perspective focused on direct medical costs 20 year time horizon; 6% discount rate Sensitivity analysis: Altered key parameter values (revision rate, cost, QOL); time horizon assessed at 5-, 10-, and 15-year cycles	Separate models for persons aged 45-50 on entry and aged 65-70 on entry "Typical" patient with advanced hip disease (no other information provided) Costs: published literature, contact with manufacturers (Fitzpatrick 1998 updated to 2000 UK£) Probabilities: THA and MoM: published literature (MoM inputs largely from McKinn 1996); WW from contact with local medical staff QOL: published literature	Hip resurfacing: 1.52% THA: 1.36%	<u>Strengths:</u> Cost utility analysis with extended time horizon Use of several alternative management strategies <u>Limitations:</u> Survival rates used to model outcome unlikely Lack of robust long term data about MoM

Table 16. Summary demographic information for economic studies.

Study	Design	Data sources and population	Model inputs	Primary strengths and limitations
Vale 2002	NICE's evaluation of industry submission received for NICE HTA (Midland Medical Technologies) Cost utility analysis provided via spreadsheet Time horizon: 20 years (analysis at 5, 10, 15, 20 years) Intervention: BHR Comparators: THA, watchful waiting Health services perspective Sensitivity analysis: Probabilistic analysis allowed variation of BHR revision rate, cost of surgery, utility.	Population: hypothetical cohort of 1000 patients. Not clear what ages were modeled, but submission focused recommendations for people ages 45-65, saying older patients are well managed with THA. Costs: UK Department of Health (interventions) , Personal Social Services Research Unit data (hospital stay) reference costs for devices; published literature BHR effectiveness: Industry data on 1693 BHRs conducted by four surgeons (82% by one surgeon) with limited four-year follow-up (66% fu at 1 year, 1% at 4 years). Ages 15-86. THA revision: Swedish national hip register Watchful waiting: resource use measures for meds, GP visits, hospitalization Utilities: Published literature for THA	Hip resurfacing: 0.5% (year 2) 2.5% (year 11+) THA: 1.0% (year 2) 5% (year 11+)	Strengths: Use of person-level data "Reasonably complete" economic evaluation of BHR Limitations: Survival rates used to model outcome unlikely Lack of data on long-term revision rates of BHR Questionable model assumption that people only exit WW for death.
Buckland 2008	Cost consequences analysis U.S. private insurance payer perspective Intervention: early hip resurfacing Comparator: Five years of conservative treatment (analgesics and anti-inflammatory rx) followed by THA Two time horizons: to age 65 and to death Net present value of direct costs and patient utilities Costs and utilities discounted at 4% Sensitivity analysis: not clearly described, but appeared to do a threshold-type analysis at varying levels of revision rate, cost, and discount rate	Hypothetical patients with moderate to severe symptoms of degenerative hip disease (age groups: 45-49, 50-54, 55-59) Costs: 2006 Medicare fee schedule, CPT codes, average wholesale price for meds, interviews with managed care directors Clinical pathways and resource use: interviews with orthopedic surgeons and gastroenterologists Utility scores: published pain-related health states for people with degenerative hip disease Life expectancy: US Life Tables Revision rates: Swedish National Hip Arthroplasty Register (THA), Oswestry Outcome Center (HR) [registry of BHR outcomes since 1997—8 year followup on 4691 pts]	Hip resurfacing: 94.3% (age <40) – 94.8% (age 70+) THA: 72.1% (age <50) – 95.2% (age >75)	Strengths: Use of recently available data on revision rates Clear description of clinical pathways Limitations: Survival rates used to model outcome unlikely Methods of sensitivity analysis unclear Patient characteristics not described beyond age Use of expert opinion to determine clinical pathways and some costs

Table 17. Summary of results for economic studies.

	Relevant results	Results of sensitivity analysis	Author conclusions
McKenzie 2003	MoM versus WW: MoM dominates WW in younger population (Additional cost for MoM= £-179, QALYs gained 3.73) MoM versus THA: THA dominates in both younger and older populations (additional cost for MoM=£1357 and £1362, QALYs gained -0.02 for both younger and older)	MoM becomes cost-effective as THA revision rate increases or MoM revision rate decreases; with increasing THA revision rate, MoM ceases to be dominated when MoM revision rate is 80% of THA rate. Decreasing MoM revision rate, MoM ceases to be dominated when the revision rate is 88% of THA rate. THA continued to dominate MoM at THA prostheses costs up to 300% of base case and at all time horizons MoM dominant over WW up to WW cost £620 MoM dominant over WW even at QOL values 0.97.	MoM warrants further study, especially since long-term data is not yet available and since revision rates were influential in the model and may substantially affect cost-effectiveness.
Vale 2002	BHR versus THA: QALYs gained at 5 and 20 years: 29, 112 Cost gained at 5 and 20 years: £378,125, -£321,333 ICER (cost per QALY) at 5 and 20 years: £13,125; BHR dominates BHR versus WW: QALYs gained at 5 and 20 years: 2499, 8963 Cost gained at 5 and 20 years: £2.752,517, -£298,997 ICER (cost per QALY) at 5 and 20 years: £1101; BHR dominates	BHR vs THA: by year 20, BHR dominated 57% of the time; THA dominated 15% of the time, BHR less effective and less costly 28% of the time, BHR more costly and more effective 0% of the time. Improvement in QALYs for BHR was small and based on the assumption of a higher revision rate for THA. In a"worst-case scenario" where THA revision rates equaled BHR revision rate, THA would be dominant (less costly with same QALY improvements). Break-even point of equal cost at 20-years: BHR revision rates 85% of THA rates for a 55-year old patient. BHR vs WW: BHR continued to dominate at 20 year follow-up	Industry conclusions: equity issues with denying younger patients BHR if WW is the alternative. Vale et al conclusions: Evidence of utility gains at 5 years at reasonable cost, however some concern with model assumptions about revision rates, assumption that people only exited watchful waiting for death, not THA. Paucity of data, especially on revision rates, are reason for caution in recommending BHR beyond four years
Buckland 2008	For each age group, immediate total HR is the dominant option E.g.: For age group 50-54, to age 65: Conservative tx: Cost per patient: $22,160; QALY per patient: 10.03 Immediate total HR: Cost per patient: $20,476; QALY per patient: 11.51	Revision rates influential, but does not change conclusions even at total HR rates=THA rates. Cost of drug treatment need to be between 26% and 42% of current estimate to change conclusions Discount rate would need to be 10% or higher to change conclusions	total HR is superior to non-surgical management of degenerative hip disease. For younger, more active patients, hip resurfacing is superior to THA.

5. Summary by Key Question

Key Question 1: What is the evidence of efficacy and effectiveness of hip resurfacing?					
HR vs. THA	**Strength of evidence**	**Conclusions/Comments**	**Quality**	**Quantity**	**Consistency**
1. Efficacy (≤1 year)	**Moderate**	• There is moderate evidence from three small randomized controlled trials that total HR is similar to THA with respect to short-term (1 year) functional, quality of life and activity outcomes.	+	–	+
> 1 year	**No evidence**	• There are no data available to assess efficacy beyond one-year follow-up.	none	none	none
2. Effectiveness Short-term (<5 years)	**Low evidence**	• There is low evidence from studies directly comparing total HR with THA to suggest that short-term (≤5 years) patient-reported outcomes, clinician-based outcomes, and pain are similar comparing total HR and THA. Activity scores tend to be slightly higher (better) in total HR patients.	–	+	+
Mid-term (5–10 years)	**Very low evidence**	• There is very low evidence from one cohort study to suggest that at an average of 5.9 years follow-up, patients treated with total HR may have better quality of life and activity outcome scores, but similar functional scores, compared with those treated with THA.	–	–	–

Key Question 2: What is the evidence about the safety profile for hip resurfacing?					
	Strength of evidence	**Conclusions/Comments**	**Quality**	**Quantity**	**Consistency**
1. Revision Short-term (<5 years)	**Moderate evidence**	• There is moderate evidence that short term revision rates are slightly higher in patients treated with total HR compared with those treated with THA. The difference in 3-year revision rates between total HR and THA in 3 registry studies range from 0.6% to 2.5% in favor of THA. The difference in 1-year revision rates in one RCT is 0.9% in favor of THA. The difference in short-term revision rates between total HR and THA in eight cohort studies varied: 4 favored THA, 2 favored total HR and 2 reported equal rates.	+	–	+
Mid-term (5-10 years)	**Low evidence**	• There is low evidence from one large registry study that 7-year revision rates are higher in patients receiving total HR versus THA (hazard ratio = 1.42, rate difference = 1.3%). Data from one small cohort study with a mean follow-up of 5.9 years reports revision rates that are similar between total HR and THA.	+	–	–
Long-term (10+ years)	**No evidence**	• There is no evidence comparing long-term revision rates between total HR and THA.	none	none	none
2. Other complications	**Low evidence**	• Reported risks of other complications in the short-term for total HR are generally low except for heterotopic ossification; the risk of femoral neck fractures range from 0.4–2.6%, avascular necrosis from 0.4-2%, femoral component loosening from 0-3.6%, acetabular component loosening from 0-1.8%, acetabular	–	+	+

Key Question 2: What is the evidence about the safety profile for hip resurfacing?					
	Strength of evidence	**Conclusions/Comments**	**Quality**	**Quantity**	**Consistency**
		component migration from 0–1.9%, and femoral component migration was not detected in any hips. Heterotopic ossification rates ranged from 0-42.7%.			
3. Learning curve threshold	**Very low evidence**	• A number of studies identified that the rate of major complications (including femoral neck fracture and revisions) decrease as surgeons gain experience performing total HR. The studies suggested that experience is associated with improved surgical technique and patient selection. However, with respect to identifying the number of procedures necessary for improved outcome, no consistent threshold was identified.	–	+	–
4. Metal ion safety	**Very low evidence**	• Patients with metal-on-metal total HR are likely to experience elevated metal serum levels (Co and Cr). Concerns have been raised regarding the safety of and risks associated with prolonged exposure to metal ions, and whether such exposure may increase the risk of cancers or metabolic disorders. However, an association between total HR and cancer or metabolic disorders has not been reported with the current length of follow-up. The results from long-term monitoring will be needed to assess the risk of metal ion exposure.	–	–	–

Key Question 3: Is there evidence of differential efficacy or safety issues with use of hip resurfacing?					
	Strength of evidence	**Conclusions/Comments**	**Quality**	**Quantity**	**Consistency**
1. Dysplasia vs. other arthritic conditions	**Low evidence**	• There is low evidence to suggest that short-term revision rates are twice as high in patients who receive total HR for a primary diagnosis of dysplasia compared with patients of primary osteoarthritis. The 5-year cumulative revision percent for dysplasia is four times greater in those receiving total HR compared with THA (12% vs. 3%) in one registry study. One small prognostic study supported this data, with 5.2% revision rates in dysplasia patients compared with 0% revision rates in osteoarthritic patients.	+	–	–
2.Osteonecrosis (AVN) vs. other arthritic conditions	**Low evidence**	• There is low evidence to suggest that short-term revision rates are slightly higher in patients who receive total HR for a primary diagnosis of ostenecrosis (AVN) compared with patients of primary osteoarthritis. The 5-year cumulative revision percent for dysplasia is two times greater in those receiving total HR compared with THA (6% vs. 3%) in one registry study and rates are the same in one small prognostic study.	+	–	–
3. Gender	**Moderate evidence**	• There is moderate evidence from three registries that 3- and 5-year revision rates are higher in females than in males (hazard ratios range from 1.57 to 2.5). Much of the difference in rates between sexes disappeared in one	+	–	+

		study when controlling for femoral component head size; the smaller the head, the higher the failure rate.			
4. Obesity	**Very low evidence**	• Two low quality studies evaluated the effect of obesity on total HR with conflicting results. One reported lower revision risk with increasing obesity, and one reported higher.	–	–	–

Key Question 4: What is the evidence of cost implications and cost effectiveness of hip resurfacing?					
	Strength of evidence	**Conclusions/Comments**	**Quality**	**Quantity**	**Consistency**
	Very low evidence	• There is limited evidence on the economic implications of hip resurfacing from two published articles and one HTA. Revision rates are important input factors in the prediction models, and no study estimated the revision rates current data.	–	–	–

References

1. Swedish Hip Arthroplasty Register. Annual Report 2007. In: Department of Orthopaedics, ed. Sahlgrenska University Hospital.

2. Kurtz SM, Lau E, Ong K, Zhao K, Kelly M, Bozic KJ. Future Young Patient Demand for Primary and Revision Joint Replacement: National Projections from 2010 to 2030. *Clin Orthop Relat Res.* Apr 10 2009.

3. Hurst NP, Jobanputra P, Hunter M, Lambert M, Lochhead A, Brown H. Validity of Euroqol--a generic health status instrument--in patients with rheumatoid arthritis. Economic and Health Outcomes Research Group. *Br J Rheumatol.* Jul 1994;33(7):655-662.

4. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care.* Jun 1992;30(6):473-483.

5. Ware J, Jr., Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care.* Mar 1996;34(3):220-233.

6. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol.* Dec 1988;15(12):1833-1840.

7. Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. *J Bone Joint Surg Br.* Mar 1996;78(2):185-190.

8. Harris WH. Traumatic arthritis of the hip after dislocation and acetabular fractures: treatment by mold arthroplasty. An end-result study using a new method of result evaluation. *J Bone Joint Surg Am.* Jun 1969;51(4):737-755.

9. D'Aubigne RM, Postel M. Function al results of hip arthroplasty with acrylic prosthesis. *J Bone Joint Surg Am.* Jun 1954;36-A(3):451-475.

10. Mont MA, Marker DR, Smith JM, Ulrich SD, McGrath MS. Resurfacing is comparable to total hip arthroplasty at short-term follow-up. *Clin Orthop Relat Res.* Jan 2009;467(1):66-71.

11. Amstutz HC, Thomas BJ, Jinnah R, Kim W, Grogan T, Yale C. Treatment of primary osteoarthritis of the hip. A comparison of total joint and surface replacement arthroplasty. *J Bone Joint Surg Am.* Feb 1984;66(2):228-241.

12. Million R, Hall W, Nilsen KH, Baker RD, Jayson MI. Assessment of the progress of the back-pain patient 1981 Volvo Award in Clinical Science. *Spine (Phila Pa 1976).* May-Jun 1982;7(3):204-212.

13. Mont MA, Ragland PS, Etienne G, Seyler TM, Schmalzried TP. Hip resurfacing arthroplasty. *J Am Acad Orthop Surg.* Aug 2006;14(8):454-463.

14. Charnley J. Arthroplasty of the hip. A new operation. *Lancet.* May 27 1961;1(7187):1129-1132.

15. Schachter AK, Lamont JG. Surface replacement arthroplasty of the hip. *Bull NYU Hosp Jt Dis.* 2009;67(1):75-82.

16. Amstutz HC, Le Duff MJ. Background of metal-on-metal resurfacing. *Proc Inst Mech Eng H.* Feb 2006;220(2):85-94.

17. Nasser S, Campbell PA, Kilgus D, Kossovsky N, Amstutz HC. Cementless total joint arthroplasty prostheses with titanium-alloy articular surfaces. A human retrieval analysis. *Clin Orthop Relat Res.* Dec 1990(261):171-185.

18. Beaule PE, Antoniades J. Patient selection and surgical technique for surface arthroplasty of the hip. *Orthop Clin North Am.* Apr 2005;36(2):177-185, viii-ix.

19. Vail TP, Mina CA, Yergler JD, Pietrobon R. Metal-on-metal hip resurfacing compares favorably with THA at 2 years followup. *Clin Orthop Relat Res.* Dec 2006;453:123-131.

20. Daniel J, Pynsent PB, McMinn DJ. Metal-on-metal resurfacing of the hip in patients under the age of 55 years with osteoarthritis. *J Bone Joint Surg Br.* Mar 2004;86(2):177-184.

21. Amstutz HC, Beaule PE, Dorey FJ, Le Duff MJ, Campbell PA, Gruen TA. Metal-on-metal hybrid surface arthroplasty: two to six-year follow-up study. *J Bone Joint Surg Am.* Jan 2004;86-A(1):28-39.

22. Hing CB, Back DL, Bailey M, Young DA, Dalziel RE, Shimmin AJ. The results of primary Birmingham hip resurfacings at a mean of five years. *Journal of Bone and Joint Surgery - Series B.* Nov 2007;89(11):1431-1438.

23. Crawford JR, Palmer SJ, Wimhurst JA, RN V. Bone loss at hip resurfacing: A comparison with total hip arthroplasty. *Hip Int.* 2005;15(4):195-198.

24. Vendittoli PA, Lavigne M, Girard J, Roy AG. A randomised study comparing resection of acetabular bone at resurfacing and total hip replacement. *J Bone Joint Surg Br.* Aug 2006;88(8):997-1002.

25. Kishida Y, Sugano N, Nishii T, Miki H, Yamaguchi K, Yoshikawa H. Preservation of the bone mineral density of the femur after surface replacement of the hip. *J Bone Joint Surg Br.* Mar 2004;86(2):185-189.

26. Girard J, Lavigne M, Vendittoli PA, Roy AG. Biomechanical reconstruction of the hip: A randomised study comparing total hip resurfacing and total hip arthroplasty. *Journal of Bone and Joint Surgery - Series B.* Jun 2006;88(6):721-726.

27. Gore DR, Murray MP, Gardner GM, Sepic SB. Hip function after total vs. surface replacement. *Acta Orthop Scand.* Oct 1985;56(5):386-390.

28. Ball ST, Le Duff MJ, Amstutz HC. Early results of conversion of a failed femoral component in hip resurfacing arthroplasty. *J Bone Joint Surg Am.* Apr 2007;89(4):735-741.

29. Capello WN, Trancik TM, Misamore G, Eaton R. Analysis of revision surgery of resurfacing hip arthroplasty. *Clin Orthop Relat Res.* Oct 1982(170):50-55.

30. Amstutz HC, Grigoris P, Dorey FJ. Evolution and future of surface replacement of the hip. *J Orthop Sci.* 1998;3(3):169-186.

31. McMinn D, Treacy R, Lin K, Pynsent P. Metal on metal surface replacement of the hip. Experience of the McMinn prothesis. *Clin Orthop Relat Res.* Aug 1996(329 Suppl):S89-98.

32. Chan FW, Bobyn JD, Medley JB, Krygier JJ, Yue S, Tanzer M. Engineering issues and wear performance of metal on metal hip implants. *Clin Orthop Relat Res.* Dec 1996(333):96-107.

33. Dowson D, Hardaker C, Flett M, Isaac GH. A hip joint simulator study of the performance of metal-on-metal joints: Part II: design. *J Arthroplasty.* Dec 2004;19(8 Suppl 3):124-130.

34. Isaac GH, Siebel T, Schmalzried TP, et al. Development rationale for an articular surface replacement: a science-based evolution. *Proc Inst Mech Eng H.* Feb 2006;220(2):253-268.

35. Rieker CB, Schon R, Konrad R, et al. Influence of the clearance on in-vitro tribology of large diameter metal-on-metal articulations pertaining to resurfacing hip implants. *Orthop Clin North Am.* Apr 2005;36(2):135-142, vii.

36. Brandt KD, Dieppe P, Radin E. Etiopathogenesis of osteoarthritis. *Med Clin North Am.* Jan 2009;93(1):1-24, xv.

37. National Collaborating Centre for Chronic Conditions. *Osteoarthritis: national clinical guidelines for care and management in adults.* London: Royal College of Physicians; 2008.

38. Majithia V, Geraci SA. Rheumatoid arthritis: diagnosis and management. *Am J Med.* Nov 2007;120(11):936-939.

39. U.S. Food and Drug Administration. Summary of Safety and Effectiveness Data: Birmingham Hip Resurfacing (BHR) System - P040033. In: Medical Device Approvals and Clearances, ed; 2006.

40. U.S. Food and Drug Administration. Summary of Safety and Effectiveness Data: Cormet Hip Resurfacing System - P050016. In: Medical Device Approvals and Clearances, ed; 2007.

41. Ong KL, Manley MT, Kurtz SM. Have contemporary hip resurfacing designs reached maturity? A review. *J Bone Joint Surg Am.* Aug 2008;90 Suppl 3:81-88.

42. Campbell P, Beaule PE, Ebramzadeh E, et al. The John Charnley Award: a study of implant failure in metal-on-metal surface arthroplasties. *Clin Orthop Relat Res.* Dec 2006;453:35-46.

43. Freeman MA. Some anatomical and mechanical considerations relevant to the surface replacement of the femoral head. *Clin Orthop Relat Res.* Jul-Aug 1978(134):19-24.

44. Shimmin AJ, Back D. Femoral neck fractures following Birmingham hip resurfacing. *Journal of Bone and Joint Surgery - Series B.* Apr 2005;87(4):463-464.

45. Little CP, Ruiz AL, Harding IJ, et al. Osteonecrosis in retrieved femoral heads after failed resurfacing arthroplasty of the hip. *J Bone Joint Surg Br.* Mar 2005;87(3):320-323.

46. Stulberg BN, Trier KK, Naughton M, Zadzilka JD. Results and lessons learned from a United States hip resurfacing investigational device exemption trial. *J Bone Joint Surg Am.* Aug 2008;90 Suppl 3:21-26.

47. Shimmin A, Beaule PE, Campbell P. Metal-on-metal hip resurfacing arthroplasty. *J Bone Joint Surg Am.* Mar 2008;90(3):637-654.

48. Back DL, Smith JD, Dalziel RE, Young DA, Shimmin A. Incidence of heterotopic ossification after hip resurfacing. *ANZ Journal of Surgery.* August 2007;77(8):642-647.

49. Mabilleau G, Kwon YM, Pandit H, Murray DW, Sabokbar A. Metal-on-metal hip resurfacing arthroplasty: a review of periprosthetic biological reactions. *Acta Orthop.* Dec 2008;79(6):734-747.

50. Back DL, Young DA, Shimmin AJ. How do serum cobalt and chromium levels change after metal-on-metal hip resurfacing? *Clin Orthop Relat Res.* Sep 2005;438:177-181.

51. Brodner W, Bitzan P, Meisinger V, Kaider A, Gottsauner-Wolf F, Kotz R. Serum cobalt levels after metal-on-metal total hip arthroplasty. *J Bone Joint Surg Am.* Nov 2003;85-A(11):2168-2173.

52. Clarke MT, Lee PT, Arora A, Villar RN. Levels of metal ions after small- and large-diameter metal-on-metal hip arthroplasty. *J Bone Joint Surg Br.* Aug 2003;85(6):913-917.

53. MacDonald SJ, McCalden RW, Chess DG, et al. Metal-on-metal versus polyethylene in hip arthroplasty: a randomized clinical trial. *Clin Orthop Relat Res.* Jan 2003(406):282-296.

54. Savarino L, Granchi D, Ciapetti G, et al. Ion release in patients with metal-on-metal hip bearings in total joint replacement: a comparison with metal-on-polyethylene bearings. *J Biomed Mater Res.* 2002;63(5):467-474.

55. Dumbleton JH, Manley MT. Metal-on-Metal total hip replacement: what does the literature say? *J Arthroplasty.* Feb 2005;20(2):174-188.

56. Sieber HP, Rieker CB, Kottig P. Analysis of 118 second-generation metal-on-metal retrieved hip implants. *J Bone Joint Surg Br.* Jan 1999;81(1):46-50.

57. Grigoris P, Roberts P, Panousis K, Jin Z. Hip resurfacing arthroplasty: the evolution of contemporary designs. *Proc Inst Mech Eng H.* Feb 2006;220(2):95-105.

58. Radcliffe IA, Taylor M. Investigation into the affect of cementing techniques on load transfer in the resurfaced femoral head: a multi-femur finite element analysis. *Clin Biomech (Bristol, Avon).* May 2007;22(4):422-430.

59. Lilikakis AK, Vowler SL, Villar RN. Hydroxyapatite-coated femoral implant in metal-on-metal resurfacing hip arthroplasty: minimum of two years follow-up. *Orthop Clin North Am.* Apr 2005;36(2):215-222, ix.

60. Marker DR, Seyler TM, Jinnah RH, Delanois RE, Ulrich SD, Mont MA. Femoral neck fractures after metal-on-metal total hip resurfacing: a prospective cohort study. *J Arthroplasty.* Oct 2007;22(7 Suppl 3):66-71.

61. Phillips B, Ball C, Sackett D, et al. Levels of evidence and grades of recommendation. Available at: http://www.cebm.net/levels_of_evidence.asp. Accessed December 2, 2006, 2006.

62. Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. *BMJ.* Jun 19 2004;328(7454):1490.

63. West S, King V, Carey TS, et.al. Systems to Rate the Strength of Scientific Evidence. Evidence Report/Technology Assessment No. 47 (Prepared by the Research Triangle Institute-University of North Carolina Evidence-based Practice Center, Contract No. 290-97-0011): Agency for Healthcare Research and Quality, Rockville, MD; 2002.

64. Garbuz DS, Tanzer M, Greidanus NV, Masri BA, Duncan CP. The John Charnley Award: Metal-on-Metal Hip Resurfacing versus Large-diameter Head Metal-on-Metal Total Hip Arthroplasty: A Randomized Clinical Trial. *Clin Orthop Relat Res.* Aug 21 2009.

65. Lavigne M, Therrien M, Nantel J, Roy A, Prince F, Vendittoli PA. The John Charnley Award: The Functional Outcome of Hip Resurfacing and Large-head THA Is the Same: A Randomized, Double-blind Study. *Clin Orthop Relat Res.* Jun 20 2009.

66. Vendittoli PA, Lavigne M, Roy AG, Lusignan D. A prospective randomized clinical trial comparing metal-on-metal total hip arthroplasty and metal-on-metal total hip resurfacing in patients less than 65 years old. *Hip Int.* 2006;16 Suppl 4:73-81.

67. Rama KR, Vendittoli PA, Ganapathi M, Borgmann R, Roy A, Lavigne M. Heterotopic ossification after surface replacement arthroplasty and total hip arthroplasty: a randomized study. *J Arthroplasty.* Feb 2009;24(2):256-262.

68. Fowble VA, dela Rosa MA, Schmalzried TP. A comparison of total hip resurfacing and total hip arthroplasty - patients and outcomes. *Bull NYU Hosp Jt Dis.* 2009;67(2):108-112.

69. Li J, Xu W, Xu L, Liang Z. Hip Resurfacing Arthroplasty for Ankylosing Spondylitis. *J Arthroplasty.* Aug 12 2009.

70. Li J, Xu W, Xu L, Liang Z. Hip resurfacing for the treatment of developmental dysplasia of the hip. *Orthopedics.* Dec 2008;31(12).

71. Pattyn C, De Smet KA. Primary ceramic-on-ceramic total hip replacement versus metal-on-metal hip resurfacing in young active patients. *Orthopedics.* Nov 2008;31(11):1078.

72. Pollard TC, Baker RP, Eastaugh-Waring SJ, Bannister GC. Treatment of the young active patient with osteoarthritis of the hip. A five- to seven-year comparison of hybrid total hip arthroplasty and metal-on-metal resurfacing. *J Bone Joint Surg Br.* May 2006;88(5):592-600.

73. Zywiel MG, Marker DR, McGrath MS, Delanois RE, Mont MA. Resurfacing matched to standard total hip arthroplasty by preoperative activity levels - a comparison of postoperative outcomes. *Bull NYU Hosp Jt Dis.* 2009;67(2):116-119.

74. Australian Orthopaedic Association National Joint Replacement Registry. Annual Report. In: Adelaide:AOA; 2008, ed.

75. National Joint Registry for England and Wales. 5th Annual Report. In: The NJR Centre, ed. Hernel Hempstead, Hertfordshire.

76. Amstutz HC, Le Duff MJ, Harvey N, Hoberg M. Improved survivorship of hybrid metal-on-metal hip resurfacing with second-generation techniques for Crowe-I and II developmental dysplasia of the hip. *J Bone Joint Surg Am.* Aug 2008;90 Suppl 3:12-20.

77. McBryde CW, Revell MP, Thomas AM, Treacy RB, Pynsent PB. The influence of surgical approach on outcome in Birmingham hip resurfacing. *Clin Orthop Relat Res.* Apr 2008;466(4):920-926.

78. McMinn DJ, Daniel J, Ziaee H, Pradhan C. Results of the Birmingham Hip Resurfacing dysplasia component in severe acetabular insufficiency: a six- to 9.6-year follow-up. *J Bone Joint Surg Br.* Jun 2008;90(6):715-723.

79. Ollivere B, Duckett S, August A, Porteous M. The Birmingham Hip Resurfacing: 5-year clinical and radiographic results from a District General Hospital. *Int Orthop.* Jun 9 2009.

80. Revell MP, McBryde CW, Bhatnagar S, Pynsent PB, Treacy RB. Metal-on-metal hip resurfacing in osteonecrosis of the femoral head. *J Bone Joint Surg Am.* Nov 2006;88 Suppl 3:98-103.

81. Treacy RB, McBryde CW, Pynsent PB. Birmingham hip resurfacing arthroplasty. A minimum follow-up of five years. *J Bone Joint Surg Br.* Feb 2005;87(2):167-170.

82. Stulberg BN, Fitts SM, Bowen AR, Zadzilka JD. Early Return to Function After Hip Resurfacing Is It Better Than Contemporary Total Hip Arthroplasty? *J Arthroplasty.* Jul 28 2009.

83. Steffen RT, Foguet PR, Krikler SJ, Gundle R, Beard DJ, Murray DW. Femoral neck fractures after hip resurfacing. *J Arthroplasty.* Jun 2009;24(4):614-619.

84. Mont MA, Seyler TM, Ulrich SD, et al. Effect of changing indications and techniques on total hip resurfacing. *Clin Orthop Relat Res.* Dec 2007;465:63-70.

85. Nunley RM, Zhu J, Brooks PJ, et al. The Learning Curve for Adopting Hip Resurfacing Among Hip Specialists. *Clin Orthop Relat Res.* Sep 25 2009.

86. Witjes S, Smolders JM, Beaule PE, Pasker P, Van Susante JL. Learning from the learning curve in total hip resurfacing: a radiographic analysis. *Arch Orthop Trauma Surg.* Oct 2009;129(10):1293-1299.

87. O'Neill M, Beaule PE, Bin Nasser A, et al. Canadian academic experience with metal-on-metal hip resurfacing. *Bull NYU Hosp Jt Dis.* 2009;67(2):128-131.

88. Siebel T, Maubach S, Morlock MM. Lessons learned from early clinical experience and results of 300 ASR hip resurfacing implantations. *Proc Inst Mech Eng H.* Feb 2006;220(2):345-353.

89. Grammatopolous G, Pandit H, Kwon YM, et al. Hip resurfacings revised for inflammatory pseudotumour have a poor outcome. *J Bone Joint Surg Br.* Aug 2009;91(8):1019-1024.

90. Keegan GM, Learmonth ID, Case CP. Orthopaedic metals and their potential toxicity in the arthroplasty patient: A review of current knowledge and future strategies. *J Bone Joint Surg Br.* May 2007;89(5):567-573.

91. United States Department of Labor: Occupational Safety and Health Administration (OSHA). Occupational Safety and Health Standards, Toxic and Hazardous Substances; Table Z-1 Limits for Air Contaminants. http://www.osha.gov/pls/oshaweb/owadisp.show_document?p_table=STANDARDS&p_id=9992.

92. Doorn PF, Campbell PA, Worrall J, Benya PD, McKellop HA, Amstutz HC. Metal wear particle characterization from metal on metal total hip replacements: transmission electron microscopy study of periprosthetic tissues and isolated particles. *J Biomed Mater Res.* Oct 1998;42(1):103-111.

93. Elfick AP, Green SM, Krikler S, Unsworth A. The nature and dissemination of UHMWPE wear debris retrieved from periprosthetic tissue of THR. *J Biomed Mater Res A.* Apr 1 2003;65(1):95-108.

94. Shukla R, Bansal V, Chaudhary M, Basu A, Bhonde RR, Sastry M. Biocompatibility of gold nanoparticles and their endocytotic fate inside the cellular compartment: a microscopic overview. *Langmuir.* Nov 8 2005;21(23):10644-10654.

95. Trindade MC, Lind M, Sun D, Schurman DJ, Goodman SB, Smith RL. In vitro reaction to orthopaedic biomaterials by macrophages and lymphocytes isolated from patients undergoing revision surgery. *Biomaterials.* Feb 2001;22(3):253-259.

96. Milosev I, Trebse R, Kovac S, Cor A, Pisot V. Survivorship and retrieval analysis of Sikomet metal-on-metal total hip replacements at a mean of seven years. *J Bone Joint Surg Am.* Jun 2006;88(6):1173-1182.

97. Willert HG, Buchhorn GH, Fayyazi A, et al. Metal-on-metal bearings and hypersensitivity in patients with artificial hip joints. A clinical and histomorphological study. *J Bone Joint Surg Am.* Jan 2005;87(1):28-36.

98. Hallab NJ, Anderson S, Caicedo M, Brasher A, Mikecz K, Jacobs JJ. Effects of soluble metals on human peri-implant cells. *J Biomed Mater Res A.* Jul 1 2005;74(1):124-140.

99. Daley B, Doherty AT, Fairman B, Case CP. Wear debris from hip or knee replacements causes chromosomal damage in human cells in tissue culture. *J Bone Joint Surg Br.* May 2004;86(4):598-606.

100. Soloviev A, Schwarz EM, Darowish M, O'Keefe RJ. Sphingomyelinase mediates macrophage activation by titanium particles independent of phagocytosis: a role for free radicals, NFkappaB, and TNFalpha. *J Orthop Res.* Nov 2005;23(6):1258-1265.

101. Huk OL, Catelas I, Mwale F, Antoniou J, Zukor DJ, Petit A. Induction of apoptosis and necrosis by metal ions in vitro. *J Arthroplasty.* Dec 2004;19(8 Suppl 3):84-87.

102. Langton DJ, Jameson SS, Joyce TJ, Webb J, Nargol AV. The effect of component size and orientation on the concentrations of metal ions after resurfacing arthroplasty of the hip. *J Bone Joint Surg Br.* Sep 2008;90(9):1143-1151.

103. Ziaee H, Daniel J, Datta AK, Blunt S, McMinn DJ. Transplacental transfer of cobalt and chromium in patients with metal-on-metal hip arthroplasty: a controlled study. *J Bone Joint Surg Br.* Mar 2007;89(3):301-305.

104. Daniel J, Ziaee H, Pradhan C, Pynsent PB, McMinn DJ. Blood and urine metal ion levels in young and active patients after Birmingham hip resurfacing arthroplasty: four-year results of a prospective longitudinal study. *J Bone Joint Surg Br.* Feb 2007;89(2):169-173.

105. Vendittoli PA, Mottard S, Roy AG, Dupont C, Lavigne M. Chromium and cobalt ion release following the Durom high carbon content, forged metal-on-metal surface replacement of the hip. *J Bone Joint Surg Br.* Apr 2007;89(4):441-448.

106. Allan DG, Trammell R, Dyrstad B, Barnhart B, Milbrandt JC. Serum cobalt and chromium elevations following hip resurfacing with the Cormet 2000 device. *J Surg Orthop Adv.* Spring 2007;16(1):12-18.

107. Daniel J, Ziaee H, Pradhan C, McMinn DJ. Six-year results of a prospective study of metal ion levels in young patients with metal-on-metal hip resurfacings. *J Bone Joint Surg Br.* Feb 2009;91(2):176-179.

108. Witzleb WC, Ziegler J, Krummenauer F, Neumeister V, Guenther KP. Exposure to chromium, cobalt and molybdenum from metal-on-metal total hip replacement and hip resurfacing arthroplasty. *Acta Orthop.* Oct 2006;77(5):697-705.

109. Antoniou J, Zukor DJ, Mwale F, Minarik W, Petit A, Huk OL. Metal ion levels in the blood of patients after hip resurfacing: a comparison between twenty-eight and thirty-six-millimeter-head metal-on-metal prostheses. *J Bone Joint Surg Am.* Aug 2008;90 Suppl 3:142-148.

110. Moroni A, Savarino L, Cadossi M, Baldini N, Giannini S. Does ion release differ between hip resurfacing and metal-on-metal THA? *Clin Orthop Relat Res.* Mar 2008;466(3):700-707.

111. Vyskocil A, Viau C. Assessment of molybdenum toxicity in humans. *J Appl Toxicol.* May-Jun 1999;19(3):185-192.

112. Davies AP, Sood A, Lewis AC, Newson R, Learmonth ID, Case CP. Metal-specific differences in levels of DNA damage caused by synovial fluid recovered at revision arthroplasty. *J Bone Joint Surg Br.* Oct 2005;87(10):1439-1444.

113. Doherty AT, Howell RT, Ellis LA, et al. Increased chromosome translocations and aneuploidy in peripheral blood lymphocytes of patients having revision arthroplasty of the hip. *J Bone Joint Surg Br.* Sep 2001;83(7):1075-1081.

114. Ladon D, Doherty A, Newson R, Turner J, Bhamra M, Case CP. Changes in metal levels and chromosome aberrations in the peripheral blood of patients after metal-on-metal hip arthroplasty. *J Arthroplasty.* Dec 2004;19(8 Suppl 3):78-83.

115. Park YS, Moon YW, Lim SJ, Yang JM, Ahn G, Choi YL. Early osteolysis following second-generation metal-on-metal hip replacement. *J Bone Joint Surg Am.* Jul 2005;87(7):1515-1521.

116. Hart AJ, Skinner JA, Winship P, et al. Circulating levels of cobalt and chromium from metal-on-metal hip replacement are associated with CD8+ T-cell lymphopenia. *J Bone Joint Surg Br.* Jun 2009;91(6):835-842.

117. Ollivere B, Darrah C, Barker T, Nolan J, Porteous MJ. Early clinical failure of the Birmingham metal-on-metal hip resurfacing is associated with metallosis and soft-tissue necrosis. *J Bone Joint Surg Br.* Aug 2009;91(8):1025-1030.

118. McGregor DB, Baan RA, Partensky C, Rice JM, Wilbourn JD. Evaluation of the carcinogenic risks to humans associated with surgical implants and other foreign bodies - a report of an IARC Monographs Programme Meeting. International Agency for Research on Cancer. *Eur J Cancer.* Feb 2000;36(3):307-313.

119. MacDonald SJ, Brodner W, Jacobs JJ. A consensus paper on metal ions in metal-on-metal hip arthroplasties. *J Arthroplasty.* Dec 2004;19(8 Suppl 3):12-16.

120. Cole P, Rodu B. Epidemiologic studies of chrome and cancer mortality: a series of meta-analyses. *Regul Toxicol Pharmacol.* Dec 2005;43(3):225-231.

121. Tharani R, Dorey FJ, Schmalzried TP. The risk of cancer following total hip or knee arthroplasty. *J Bone Joint Surg Am.* May 2001;83-A(5):774-780.

122. Visuri T, Pukkala E. *Does metal-on-metal total hip prosthesis have influence on cancer?: a long term follow-up study.*: In: Reiker C, Oberholzer S, Wyss U, eds. *World tribology forum in arthroplasty*. Seattle: Hans Huber, 181-88.; 2001.

123. MacDonald SJ. Can a safe level for metal ions in patients with metal-on-metal total hip arthroplasties be determined? *J Arthroplasty.* Dec 2004;19(8 Suppl 3):71-77.

124. Pandit H, Glyn-Jones S, McLardy-Smith P, et al. Pseudotumours associated with metal-on-metal hip resurfacings. *J Bone Joint Surg Br.* Jul 2008;90(7):847-851.

125. De Haan R, Pattyn C, Gill HS, Murray DW, Campbell PA, De Smet K. Correlation between inclination of the acetabular component and metal ion levels in metal-on-metal hip resurfacing replacement. *J Bone Joint Surg Br.* Oct 2008;90(10):1291-1297.

126. McBryde CW, Shears E, O'Hara JN, Pynsent PB. Metal-on-metal hip resurfacing in developmental dysplasia: a case-control study. *J Bone Joint Surg Br.* Jun 2008;90(6):708-714.

127. Mont MA, Seyler TM, Marker DR, Marulanda GA, Delanois RE. Use of metal-on-metal total hip resurfacing for the treatment of osteonecrosis of the femoral head. *J Bone Joint Surg Am.* Nov 2006;88 Suppl 3:90-97.

128. Beaule PE, Dorey FJ, LeDuff M, Gruen T, Amstutz HC. Risk factors affecting outcome of metal-on-metal surface arthroplasty of the hip. *Clin Orthop Relat Res.* Jan 2004(418):87-93.

129. Le Duff MJ, Amstutz HC, Dorey FJ. Metal-on-metal hip resurfacing for obese patients. *J Bone Joint Surg Am.* Dec 2007;89(12):2705-2711.

130. McKenzie L, Vale L, Stearns S, McCormack K. Metal on metal hip resurfacing arthroplasty. An economic analysis. *Eur J Health Econ.* 2003;4(2):122-129.

131. Ofman JJ, Sullivan SD, Neumann PJ, et al. Examining the value and quality of health economic analyses: implications of utilizing the QHES. *J Manag Care Pharm.* Jan-Feb 2003;9(1):53-61.

132. Chiou CF, Hay JW, Wallace JF, et al. Development and validation of a grading system for the quality of cost-effectiveness studies. *Med Care.* Jan 2003;41(1):32-44.

Glossary of Terms

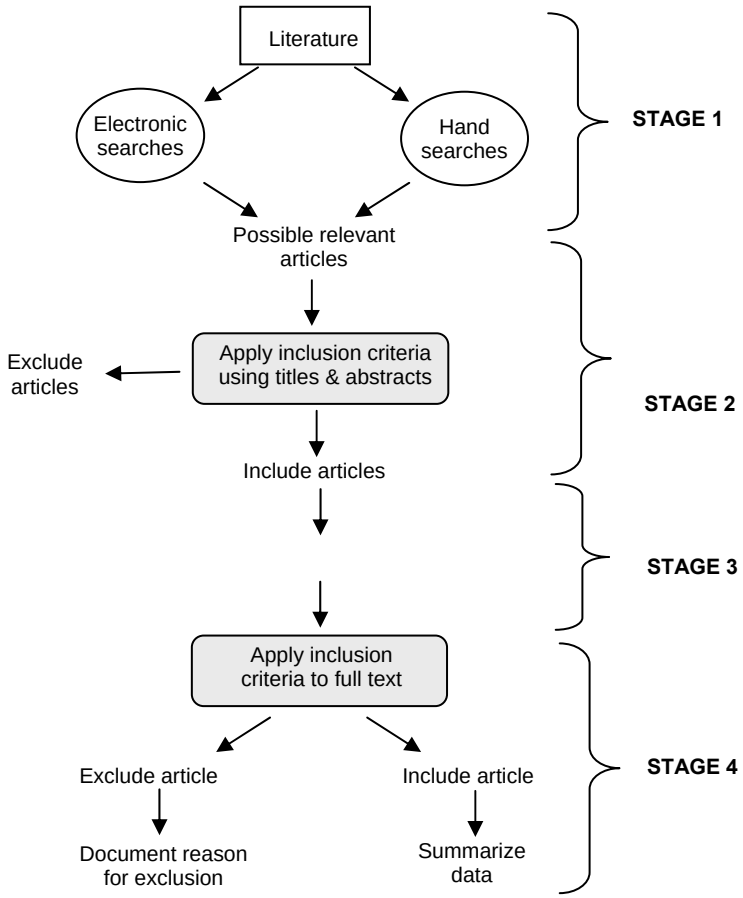
Ankylosing spondylitis (AS)	A chronic, inflammatory arthritis that affects the joints of spine and the sacroiliac joint of the pelvis and causes eventual fusion of the spine. Though genetics play a role, its cause is largely unknown. AS causes pain and stiffness of low back and hip, progressing to the neck and chest.
Anteversion	The tipping forward of an entire organ or part. In this report it is used to describe acetabular component positioning.
Apoptosis	A natural process of self-destruction in certain cells that is determined by the genes and can be initiated by a stimulus or by removal of a repressor agent. Also called *programmed cell death*.
Articular Surfacing Replacement (ASR) Device	A metal-on-metal prosthesis that is manufactured by DePuy Orthopaedics, Inc., Warsaw, IN. The ASR is not FDA-approved. It is currently marketed in Canada, Europe, India, and Australia.
Avascular necrosis (AVN)	Bone death due to temporary or permanent cessation of blood flow to the bone.
Birmingham Hip Resurfacing (BHR) System	The first FDA-approved (May 2006) hip resurfacing system available for use in the US; it is also manufactured globally. It is a metal-on-metal prosthesis composed of high carbon and cobalt-chromium alloy; the acetabular component has a hydroxyapatite coating. The BHR is manufactured by Smith & Nephew, Inc., Memphis TN.
Body mass index (BMI)	A measurement that has replaced weight as the preferred determinant of obesity. The BMI can be calculated (in English units) as 703.1 times a person's weight in pounds divided by the square of the person's height in inches.
Conserve Plus Hip Resurfacing	A metal-on-metal prosthesis composed of cobalt-chromium alloy. It is manufactured by Wright Medical Technology, Inc., Arlington, TN. It is currently being marketed in Europe and Asia and is awaiting FDA approval in the US.
Cormet Hip Resurfacing System	Approved for use in the US by the FDA in July 2007. The Cormet is a metal-on-metal prosthesis composed of cobalt-chromium alloy; the acetabular component has a bi-coating of plasma sprayed titanium and hydroxyapatite. It is manufactured by Stryker/Corin Medical, Ltd., USA, Tampa, FL.
Cytotoxicity	The degree to which an agent possesses a specific destructive

	action on certain cells. Most often used to describe decomposition of cells by immune mechanisms.
Dislocation	Displacement of the bone.
Durom Hip Resurfacing	A metal-on-metal prosthesis composed of a wrought-forged high carbon and cobalt-chromium alloy; the acetabular component has a coating of pure sprayed titanium. It is manufactured by Zimmer, Inc., Swindon, UK. The Durom is not FDA-approved but is marketed in North America outside of the US and in the UK.
Dysplasia of the hip	A hereditary disease that, in its more severe form, can eventually cause crippling lameness and painful arthritis of the hip. The term dysplasia refers to an abnormality in maturation of cells within a tissue.
Endocytosis	A process of cellular ingestion by which the plasma membrane folds inward to bring substances into the cell.
European Quality of Life (EQ-5D) measure	A generic, patient-reported outcome measures that assesses mobility, self-care, usual activity, pain, and anxiety/depression.
Harris Hip Score (HHS)	A disease-specific, clinician-reported outcome measure that assesses a patient's pain, function, deformity, and range of motion, and classifies their overall hip function as excellent, good, fair, or poor based on the sum of all domain scores ranging 0 (poor) to 100 (excellent).
Heterotopic ossification (HO)	Unwanted bone growth around an implant that causes pain and reduces range of motion.
Hydroxyapatite	A naturally occurring mineral form of calcium apatite and is commonly used as a coating to promote bone ingrowth into prosthetic implants.
Macrophage	A large white blood cell, found primarily in the bloodstream and connective tissue, that helps the body fight off infections by ingesting the disease-causing organism. They are usually immobile but become actively mobile when stimulated by inflammation.
Merle D'Aubigne hip score	A disease-specific, clinician-reported outcome measure that assesses a patient's pain, mobility, and walking ability and classifies their overall hip function as very good, good, medium, fair, or poor based on the sum of all domain scores ranging from 0 (poor) to 12 (very good).
Mont Activity measure	A disease-specific, patient-reported outcome measure that

	assesses each activity that the patient regularly performs and interprets patients as "low-activity" or "high-activity" based on an established scoring system.
Osteopenia	A condition where bone density is lower than normal. It is considered by many doctors to be a precursor to osteoporosis.
Osteoporosis	The thinning of bone tissue and loss of bone density over time, that leads to an increased risk of fracture, even after minimal trauma.
Osteoarthritis (OA)	Also referred to as degenerative joint disease (DJD), OA is a non-inflammatory, progressive disorder of the joints caused by gradual loss of cartilage and resulting in the development of bony spurs and cysts. It is caused by "wear and tear" on the joint and most commonly affects the knee and hip.
Osteolysis	Dissolution of bony tissues; applied especially to the removal or loss of the calcium of bone.
Osteonecrosis	Destruction and death of bone tissue due to ischemia (disruption of the blood supply), infection, malignant disease, or trauma.
Osteophyte	Unwanted bone growth.
Oxford hip score	A disease-specific, patient-reported outcome measure comprised of 12 questions (1–5 points each) concerning the perception of pain and function. The higher the score, the lower the function.
Oxidative stress	Any of various pathologic changes seen in living organisms in response to excessive levels of cytotoxic oxidizing agents and free radicals, which are generated by various stressors in the environment (e.g., tobacco, alcohol, toxic metals, quinones).
Perthes disease	A degenerative disease of the hip joint, where growth/loss of bone mass leads to some degree of collapse of the hip joint and to deformity of the ball of the femur and the surface of the hip socket. It is typically found in young children, and it can lead to osteoarthritis in adults.
Phagocytosis	A process by which a white blood cell envelopes and digests debris and microorganisms to remove them from the blood.
Quality-adjusted life years (QALYs)	A way of measuring both the quality and the quantity of life lived, as a means of quantifying in benefit of a medical intervention. They are based on the number of years of life that would be added by the intervention.
Rheumatoid arthritis	A chronic, systemic disease that affects the lining of peripheral

(RA)	joints. It causes inflammatory responses, which destroy the articular cartilage and the tissues around the joints, causing joint deformity.
Sepsis	A serious medical condition characterized by a whole-body inflammatory state and infection due to the overwhelming presence of pathogenic organisms in the bloodstream.
Short Form 36 health survey questionnaire (SF-36)	A generic, patient-reported outcome measure comprised of 8 subscales with various #'s of items: physical functioning, role limitation due to physical health problems, bodily pain, general health, vitality, social functioning, role limitations due to emotional problems, and mental health. Each subscale is scored separately (0–100 points); a total score is not used. The lower the score, the greater the disability.
University of California-Los Angeles (UCLA) activity scale	A disease-specific, patient-reported outcome measure that classifies a patient's activity level on a scale from 1 ("bedridden") to 10 ("unrestricted").
Visiual Analogue Scale (VAS) for pain	A generic, patient-reported outcome measure in which a patient rates their level of pain on a scale from 0 (no pain) to 10 (worst pain imaginable).
Western Ontario and McMaster Universities OA index (WOMAC)	A disease-specific, patient-reported outcome measure assessing pain, stiffness, and physical function. The higher the total score, the greater the disability.

Appendix A. ALGORITHM FOR ARTICLE SELECTION



Appendix B. SEARCH STRATEGIES

Database: MEDLINE

1	("Surface replacement arthroplasty" AND HIP) OR "hip resurfacing" OR ((MoM OR "METAL ON METAL") AND HIP)
2	(Hip[TI] AND (Resurfacing[TI] OR Metal-On-Metal[TI] OR Birmingham OR Conserve Plus OR Wagner Resurfacing)
3	"Finite Element Analysis"[Mesh] OR Engineer*
4	"Case Reports "[Publication Type] OR cadaver OR IN VITRO
5	#1 OR #2
6	#5 NOT (#3 OR #4)
7	limit English/abstracts
8	("Comparative Study "[Publication Type] OR "Clinical Trials, Phase III as Topic"[Mesh])
9	**#7 AND #8**
1	("Surface replacement" AND HIP[TI]) OR (hip[TI] AND resurfacing*[TI])
2	"Finite Element Analysis"[Mesh] OR Engineer*
3	"Case Reports "[Publication Type] OR cadaver OR IN VITRO
4	#1 NOT (#2 OR #3)
5	**limit English/abstracts**
6	SAFE* OR COMPLICATION*
7	**#5 AND #6**

Database: EMBASE

1	(("surface replacement arthroplasty" and hip) or "hip resurfacing" or ((mom or "metal on metal") and hip)).mp.
2	(Hip and (Resurfacing or Metal-On-Metal or Birmingham or Conserve Plus or Wagner Resurfacing)).mp.
3	("Finite Element Analysis" or Engineer).mp.
4	1 or 2
5	limit 4 to abstracts
6	limit 5 to (human and (article or report or "review"))
7	comparative study/ or clinical trial/
8	6 and 7
9	perioperative complication/ or peroperative complication/ or postoperative complication/ or complication/ or safety.mp.
10	6 and 9
11	"cost utility analysis"/ or "cost benefit analysis"/ or "cost minimization analysis"/ or "cost"/ or "cost effectiveness analysis"/
12	6 and 11

Parallel strategies were used to search the Cochrane Library and others listed below. Keyword searches were conducted in the other listed resources.

Electronic Database Searches
The following databases have been searched for relevant information:
Agency for Healthcare Research and Quality (AHRQ)
Cumulative Index to Nursing and Allied Health (CINAHL)
Cochrane Database of Systematic Reviews (through 2009, Issue 2)
Cochrane Registry of Clinical Trials (CENTRAL) (through 2009, Issue 2)
Cochrane Review Methodology Database (through 2009, Issue 2)
Computer Retrieval of Information on Scientific Projects (CRISP)
Database of Reviews of Effectiveness (Cochrane Library) (through 2009, Issue 2)
EMBASE (1985 through July 23, 2009)
PubMed (1975 through July 23, 2009)
Informational Network of Agencies for Health Technology Assessment (INAHTA)
NHS Economic Evaluation Database (Cochrane Library through 2009, Issue 2)
HSTAT (Health Services/Technology Assessment Text)
EconLIT

Additional Economics, Clinical Guideline and Gray Literature Databases
AHRQ- Healthcare Cost and Utilization Project
Canadian Agency for Drugs and Technologies in Health
Centers for Medicare and Medicaid Services (CMS)
Food and Drug Administration (FDA)
Google

Institute for Clinical Systems Improvement (ICSI)
National Guideline Clearinghouse

Appendix C. EXCLUDED ARTICLES

Exclude at full-text review

Efficacy/ effectiveness:

Study	Reason for exclusion
1. Vendittoli (2006 (prospective…)	No clinical outcomes reported
2. McGrath (2008)	Total HR in patients 60 and older
3. McGrath (2009)	Revision surgery, not primary HR
4. Mont (2007) (Gait)	Gait only, no other clinical outcomes reported
5. Mont (2001)	Hemi resurfacing
6. Le Duff	Not all THAs are primary
7. Marker (2009)	Review with no primary data
8. Stulberg (2009)	Data reported previously (Stulberg (2008))

Safety:

Study	**Reason for exclusion**
1. Amstutz 2004 (fracture...)	Descriptive study of femoral neck fractures following HR
2. Amstutz 2005 (surface…)	Case-series with short-term F/U
3. Amstutz 2007	All hips reported in a later study (Amstutz, Le Duff improved survivorship (2008))
4. Amstutz 2007 (effects tech change)	Exposure is a change in the way they did surgery
5. Amstutz 2008 (present state…)	No safety data reported
6. Back 2005	Case-series with short-term F/U
7. Beaule 2004 (MoM…)	Cemented acetabular components
8. Beaule 2009	Case-series with short-term F/U
9. Bergeron 2009	Case-series with short-term F/U
10. Boyd 2007	Case-series with short-term F/U
11. Costi 2009	Cemented acetabular components (not modern total HR)
12. Daniel 2004	Case-series with short-term F/U
13. De Smet	Case-series with short-term F/U
14. Hart 2009	Lab study, no clinical data
15. Hing 2007	Review with no primary data
16. Howie 1990	Long term follow-up on discontinued total HR system (not modern total HR)
17. Lilikakis 2005	Case-series with short-term F/U
18. Marker 2007	Risk factor for femoral neck fracture (used for context on learning curve)
19. McGrath 2009	Data included both total HR and hemi HR
20. Mont-Seyler 2007 Effect	Case-series with short-term F/U
21. Moroni 2008	Lab study, no clinical data
22. Naal 2009 (sports)	Case-series with short-term F/U
23. O'Neill 2009	Case-series with short-term F/U
24. Sandri 2009	Case-series with short-term F/U
25. Schmalzried 1996	Early discontinued total HR systems (not modern HR)
26. Shimmin 2005	Physician survey without reports of response rate
27. Siebel 2006	Case-series with short-term F/U
28. Springer (meta)	Meta-analysis of case series; did not describe how event rate was calculated
29. Steffen 2008	Case-series with short-term F/U
30. Witzleb 2008	Case-series with short-term F/U

Special populations:

Study	Reason for exclusion
1. Steffen 2009	Risk factors for femoral neck fracture

Appendix D. LEVEL OF EVIDENCE DETERMINATION

Each study was rated against pre-set criteria that resulted in an evidence rating (Level of Evidence I, II, III, or IV) and presented in a table. For therapeutic and prognostic articles, the criteria are listed in the Table below.

Definition of the different levels of evidence for articles on therapy and prognosis

	Studies of Therapy			Studies of Prognosis	
Level	**Study design**	**Criteria**		**Study design**	**Criteria**
I	Good quality RCT	• Concealment • Blind or independent assessment for important outcomes • Co-interventions applied equally • F/U rate of 80%+ • Adequate sample size		Good quality cohort	• Prospective design • Patients at similar point in the course of their disease or treatment • F/U rate of 80%+ • Patients followed long enough for outcomes to occur • Controlling for extraneous prognostic factors*
II	Moderate or poor quality RCT	• Violation of any of the criteria for good quality RCT		Moderate quality cohort	• Prospective design, with violation of one of the other criteria for good quality cohort study • Retrospective design, meeting all the rest of the criteria in level I
	Good quality cohort	• Blind or independent assessment in a prospective study, or use of reliable data* in a retrospective study • Co-interventions applied equally • F/U rate of 80%+ • Adequate sample size • Controlling for possible confounding†			
III	Moderate or poor quality cohort	• Violation of any of the criteria for good quality cohort		Poor quality cohort	• Prospective design with violation of 2 or more criteria for good quality cohort, or • Retrospective design with violation of 1 or more criteria for good quality cohort
	Case-control	• Any case-control design		Case-control	• Any case-control design
IV	Case series	• Any case series design		Case series	• Any case series design

* Authors must provide a description of robust baseline characteristics, and control for those that are unequally distributed between treatment groups.

Studies from Registries	
Study design	**Criteria**
Good quality registry	• Designed specifically for conditions evaluated • Includes prospective data only • Validation of completeness and quality of data • Patients followed long enough for outcomes to occur • Independent outcome assessment* • Complete follow-up of $\geq$ 85% • Controlling for possible confounding† • Accounting for time at risk‡
Moderate quality cohort	• Prospective data from registry designed specifically for conditions evaluated with violation of 2 of the rest of the criteria in level I
Poor quality cohort	• Prospective data from registry designed specifically for conditions evaluated with violation of 3 or more of the rest of the criteria in level I • Retrospective data or data from a registry not designed specifically for conditions evaluated

* Outcome assessment is independent of healthcare personnel judgment. Some examples include patient reported outcomes, death, and reoperation.
† Authors must provide a description of robust baseline characteristics, and control for those that are unequally distributed between treatment groups.
‡ Equal follow-up times or for unequal follow-up times, accounting for time at risk.

Determination of Overall Strength of Evidence
Following the assessment of the quality of each individual study included in the report, an overall "strength of evidence for the relevant question or topic is determined. Methods for determining the overall strength of evidence for diagnostic studies are variable across the literature and are most applicable to evaluation of therapeutic studies.

SRI's method incorporates the primary domains of quality (LoE), quantity of studies and consistency of results across studies as described by AHRQ.[63]

The following definitions are used by SRI to determine whether or not the body of evidence meets the criteria for each domain:

Domain	Definition/Criterion
Quality	• At least 80% of the studies are LoE I or II
Quantity	• There are at least three studies which are adequately powered to answer the study question
Consistency	• Study results would lead to a similar conclusion (similar values, in the same direction) in at least 70% of the studies

Based on the criteria described above, the possible scenarios that would be encountered are described below. Each scenario is ranked according to the impact that future research is likely to have on both the overall estimates of an effect and the confidence in the estimate. This ranking describes the overall "Strength of Evidence" (SoE) for the body of literature on a specific topic. The method and descriptions of overall strength are adapted for diagnostic studies from system described by the GRADE Working Group[62] for the development of clinical guidelines.

SoE	Description	Further Research Impact	Domain Criterion Met		
			Quality	Quantity	Consistency
1	High	Very unlikely to change confidence in effect estimate	+	+	+
2	Moderate	Likely to have an important impact on confidence in estimate and *may* change the estimate	+	-	+
			+	+	-
3	Low	Very likely to have an important impact on confidence in estimate and *likely* to change the estimate	+	-	-
			-	+	+
4	Very Low	Any effect estimate is uncertain	-	+	-
			-	-	+
			-	-	-

Assessment of Economic Studies

Full formal economic analyses evaluate both costs and clinical outcomes of two or more alternative interventions. The four primary types are cost minimization analysis (CMA), cost-utility analysis (CUA), cost-effectiveness analysis (CEA), and cost-benefit analyses (CBA). Each employs different methodologies, potentially complicating critical appraisal, but some common criteria can be assessed across studies.

No standard, universally accepted method of critical appraisal of economic analyses is currently in use. A number of checklists [Canadian, BMJ, AMA] are available to facilitate critique of such studies. The Quality of Health Economic Studies (QHES) instrument developed by Ofman, et al.[131] QHES embodies the primary components relevant for critical appraisal of economic studies.[131, 132] It also incorporates a weighted scoring process and which was used as one factor to assess included economic studies. This tool has not yet undergone extensive evaluation for broader use but provides a valuable starting point for critique.

In addition to assessment of criteria in the QHES, other factors are important in critical appraisal of studies from an epidemiologic perspective to assist in evaluation of generalizability and potential sources of study bias.

Such factors include:

- Are the interventions applied to similar populations (eg, with respect to age, gender, medical conditions, etc)? To what extent are the populations for each intervention comparable and are differences considered or accounted for? To what extent are population characteristics consistent with "real world" applications of the comparators?
- Are the sample sizes adequate so as to provide a reasonable representation of individuals to whom the technology would be applied?
- What types of studies form the basis for the data used in the analyses? Data (eg, complication rates) from randomized controlled trials or well-conducted, methodologically rigorous cohort studies for data collection are generally of highest quality compared with case series or studies with historical cohorts.
- Were the interventions applied in a comparable manner (eg, similar protocols, follow-up procedures, evaluation of outcomes, etc)?
- How were the data and/or patients selected or sampled (eg, a random selection of claims for the intervention from a given year/source or all claims)? What specific inclusion/exclusion criteria or processes were used?
- Were the outcomes and consequences of the interventions being compared comparable for each? (eg, were all of the relevant consequences/complications for each intervention considered or do they primarily reflect those for one intervention?)

Assessment of the overall strength of evidence for formal economic analyses does not appear to be documented in the literature. For the purposes of this HTA, overall strength was determined by:

- Quality of the individual studies: Where the majority of quality indicators described in the QHES met and were the methods related to patient/claim selection, patient population considerations and other factors listed above consistent with a high quality design?
- Number of formal analyses (3 or more)
- Consistency of findings and conclusions from analyses across studies.

QHES Instrument[131] **Study** _____

Questions	Points	Yes	No
1. Was the study objective presented in a clear, specific, and measurable manner?	7		
2. Were the perspective of the analysis (societal, third-party payer, etc.) and reasons for its selection stated?	4		
3. Were variable estimates used in the analysis from the best available source (ie, randomized controlled trial - best, expert opinion - worst)?	8		
4. If estimates came from a subgroup analysis, were the groups prespecified at the beginning of the study?	1		
5. Was uncertainty handled by (1) statistical analysis to address random events, (2) sensitivity analysis to cover a range of assumptions?	9		
6. Was incremental analysis performed between alternatives for resources and costs?	6		
7. Was the methodology for data abstraction (including the value of health states and other benefits) stated?	5		
8. Did the analytic horizon allow time for all relevant and important outcomes? Were benefits and costs that went beyond 1 year discounted (3% to 5%) and justification given for the discount rate?	7		
9. Was the measurement of costs appropriate and the methodology for the estimation of quantities and unit costs clearly described?	8		
10. Were the primary outcome measure(s) for the economic evaluation clearly stated and did they include the major short-term, long-term and negative outcomes included?	6		
11. Were the health outcomes measures/scales valid and reliable? If previously tested valid and reliable measures were not available, was justification given for the measures/scales used?	7		
12. Were the economic model (including structure), study methods and analysis, and the components of the numerator and denominator displayed in a clear, transparent manner?	8		
13. Were the choice of economic model, main assumptions, and limitations of the study stated and justified?	7		
14. Did the author(s) explicitly discuss direction and magnitude of potential biases?	6		
15. Were the conclusions/recommendations of the study justified and based on the study results?	8		
16. Was there a statement disclosing the source of funding for the study?	3		
TOTAL POINTS	**100**		

Appendix E. LEVEL OF EVIDENCE FOR COMPARATIVE STUDIES.

Methodological quality of therapeutic studies evaluating efficacy or effectiveness following hip resurfacing.

Methodological principle	Fowble (2009)	Garbuz (2009)	Lavigne (2009)	Li (2009)	Li (2008)	Mont (2009)
Study design Randomized controlled trial Cohort study Case-series	✓	✓	✓	✓	✓	✓
Statement of concealed allocation*			✓			
Intention to treat*			✓			
Independent or blind assessment		✓	✓			✓
Cointerventions applied equally	✓	✓	✓	✓	✓	
Complete follow-up of ≥ 85%	✓		✓	✓	✓	✓
Adequate sample size	✓	✓				
Controlling for possible confounding†		✓	✓			
Evidence class	**III**	**II**	**II**	**III**	**III**	**III**

Methodological principle	Pattyn (2008)	Pollard (2006)	Rama (2009)	Vendittoli (2006)	Stulberg (2008)	Vail (2006)	Zywiel (2009)
Study design Randomized controlled trial Cohort study Case-series	✓	✓		✓	✓	✓	✓
Statement of concealed allocation*							
Intention to treat*							
Independent or blind assessment		✓					✓
Cointerventions applied equally	✓	✓		✓	✓	✓	✓
Complete follow-up of ≥ 85%	✓	✓		✓	✓		
Adequate sample size	✓	✓		✓		✓	✓
Controlling for possible confounding†						✓	✓
Evidence class	**III**	**III**	**II**	**II**	**III**	**III**	**III**

* Applies to RCTs only.
† Authors must provide a description of robust baseline characteristics, and control for those that are unequally distributed between treatment groups.

Methodological quality of registry studies assessing hip resurfacing.

Methodological principle	Australia Registry	Swedish Registry	UK Registry
Designed specifically for conditions evaluated	✓	✓	✓
Includes prospective data only	✓	✓	✓
Validation of completeness and quality of data	✓		✓
Patients followed long enough for outcomes to occur	✓	✓	✓
Independent outcome assessment*	✓	✓	✓
Complete follow-up of ≥ 85%	✓	✓	✓
Controlling for possible confounding†	✓	✓	✓
Accounting for time at risk‡	✓	✓	✓
Evidence class	**II**	**II**	**II**

* Outcome assessment is independent of healthcare personnel judgment. Some examples include patient reported outcomes, death, and reoperation.

† Authors must provide a description of robust baseline characteristics, and control for those that are unequally distributed between treatment groups.
‡ Equal follow-up times or for unequal follow-up times, accounting for time at risk.

Methodological quality of prognostic studies assessing factors associated with outcome following hip resurfacing.

Methodological principle	Amstutz (2004)	Ball (2007)	Beaule (2004)	Grammatopoulos (2009)	Le Duff (2007)	McBryde (2008)	Mont (2006)	Ollivere (2009)
Study design Prospective cohort study Retrospective cohort study Case-control study Case-series	✓	✓	✓	✓ ✓	✓	✓	✓	✓
Patients at similar point in the course of their disease or treatment	✓	✓	✓	✓	✓	✓	✓	✓
Patients followed long enough for outcomes to occur	✓	✓	✓	✓	✓	✓	✓	✓
Complete follow-up of ≥ 85%	✓		✓		✓	✓	✓	✓
Controlling for extraneous prognostic factors*	✓		✓				✓	
Evidence class	**II**	**III**	**II**	**III**	**III**	**III**	**II**	**III**

* Authors must provide a description of robust baseline characteristics, and control for those that are unequally distributed between treatment groups.

Appendix F. CLINICAL PEER REVIEWERS

Reviewer	Areas of expertise
Seth S. Leopold, MD Professor & Vice Chair University of Washington School of Medicine Department of Orthopaedics	• Orthopedic surgeon • Assistant Professor, University of Washington School of Medicine, Department of Orthopaedic • Instructor, evidence-based Orthopaedic
Jason S. Weisstein, MD, MPH, FACS Assistant Professor University of Washington School of Medicine Department of Orthopaedics	• Orthopedic surgeon • Assistant Professor, University of Washington School of Medicine, Department of Orthopaedics and Sports • Joint Reconstruction Service – primary and revision hip and knee reconstruction utilizing the latest techniques. Specialty trained in hip resurfacing, quadriceps sparing, and mini incision total knee arthroplasty. • Musculoskeletal Oncology Service – resection and reconstruction surgery for benign and malignant bone tumors, treatment of metastatic disease, avascular necrosis, and orthopaedic conditions affecting cancer patients. • Master of Public Health in Epidemiology •
Paul A. Manner, MD, FRCSC Assistant Professor University of Washington School of Medicine Department of Orthopaedics	• Orthopedic surgeon • Assistant Professor, University of Washington School of Medicine, Department of Orthopaedics and Sports • Adult reconstruction and arthroplasty