Washington State Health Care Authority

Health Technology Assessment

Hip Resurfacing (Re-Review)

Final Evidence Report: Appendices

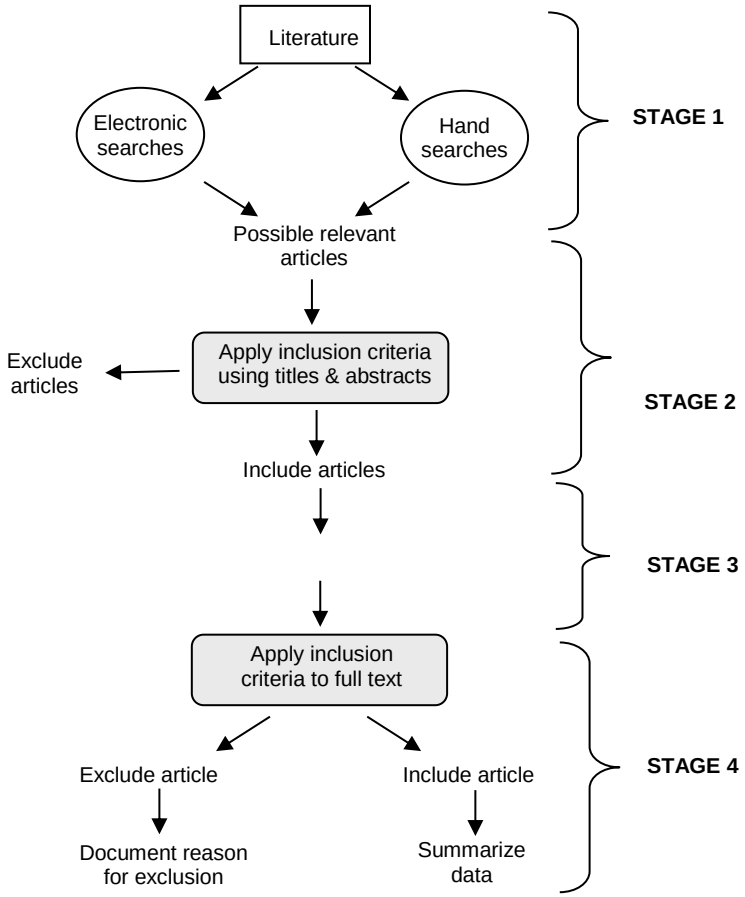
October 14, 2013

Health Technology Assessment Program (HTA)
Washington State Health Care Authority
PO Box 42712
Olympia, WA 98504-2712
(360) 725-5126
hta.hca.wa.gov

shtap@hca.wa.gov

Table of Contents

Appendix A. ALGORITHM FOR ARTICLE SELECTION



Appendix B. SEARCH STRATEGIES

Database: MEDLINE

Key Questions 1, 2, 4, 5

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**
10	SAFE* OR COMPLICATION* OR REVISE* OR REVISION* OR “ADVERSE EVENTS”
11	**#7 AND #10**
12	"cost utility analysis" OR "cost benefit analysis" OR "cost minimization analysis" OR "cost" OR "cost effectiveness analysis" or "Cost-Benefit Analysis"[Mesh]
13	**#7 AND #12**

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

Key Question 3

1	("Arthroplasty, Replacement, Hip"[Mesh] OR "HIP REPLACEMENT" OR "TOTAL HIP") AND "HIP RESURFACING" AND (REVISE*[TI] OR Revision[TI]) AND (OUTCOME*[TI] OR RESULT*[TI] OR FOLLOW-UP[TI] OR SUBSEQUENT[TI] OR FAIL*[TI])
2	("Hip resurfacing" OR "total hip arthroplasty") AND "outcome* of revised hip resurfacing" OR re-revis*
3	#1 OR #2

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
ORIGINAL REPORT	
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.	No clinical outcomes reported
McGrath MS, Desser DR, Ulrich SD, Seyler TM, Marker DR, Mont MA. Total hip resurfacing in patients who are sixty years of age or older. J Bone Joint Surg Am 2008;90 Suppl 3:27-31.	Total HR in patients 60 and older
McGrath MS, Marker DR, Seyler TM, Ulrich SD, Mont MA. Surface replacement is comparable to primary total hip arthroplasty. Clin Orthop Relat Res 2009;467:94-100.	Revision surgery, not primary HR
Mont MA, Rajadhyaksha AD, Hungerford DS. Outcomes of limited femoral resurfacing arthroplasty compared with total hip arthroplasty for osteonecrosis of the femoral head. J Arthroplasty 2001;16:134-9.	Gait only, no other clinical outcomes reported
Mont MA, Seyler TM, Ragland PS, Starr R, Erhart J, Bhave A. Gait analysis of patients with resurfacing hip arthroplasty compared with hip osteoarthritis and standard total hip arthroplasty. J Arthroplasty 2007;22:100-8.	Hemi resurfacing
Le Duff MJ, Wisk LE, Amstutz HC. Range of motion after stemmed total hip arthroplasty and hip resurfacing - a clinical study. Bull NYU Hosp Jt Dis 2009;67:177-81.	Not all THAs are primary
Marker DR, Strimbu K, McGrath MS, Zywiel MG, Mont MA. Resurfacing versus conventional total hip arthroplasty - review of comparative clinical and basic science studies. Bull NYU Hosp Jt Dis 2009;67:120-7	Review with no primary data
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 2009	Data reported previously (Stulberg (2008))
UPDATE REPORT	
Malviya A, Ramaskandhan JR, Bowman R, et al. What advantage is there to be gained using large modular metal-on-metal bearings in routine primary hip replacement? A preliminary report of a prospective randomised controlled trial. J Bone Joint Surg Br 2011;93:1602-9.	Study of prosthetic head size
Marker DR, Zywiel MG, Johnson AJ, Seyler TM, Mont MA. Are component positioning and prosthesis size associated with hip resurfacing failure? BMC Musculoskelet Disord 2010;11:227.	Comparator is type of surgical approach

Safety:

Study	Reason for exclusion
Amstutz, H. C., P. A. Campbell, et al. (2004). "Fracture of the neck of the femur after surface arthroplasty of the hip." J Bone Joint Surg Am 86-A(9): 1874-1877.	Descriptive study of femoral neck fractures following HR
Amstutz, H. C., E. P. Su, et al. (2005). "Surface arthroplasty in young patients with hip arthritis secondary to childhood disorders." Orthop Clin North Am 36(2): 223-230, x.	Case-series with short-term F/U
Amstutz, H. C., J. T. Antoniades, et al. (2007). "Results of metal-on-metal hybrid hip resurfacing for Crowe type-I and II developmental dysplasia." J Bone Joint Surg Am 89(2): 339-346.	All hips reported in a later study (Amstutz, Le Duff improved survivorship (2008))
Amstutz HC, Le Duff MJ, Campbell PA, Dorey FJ. The effects of technique changes on aseptic loosening of the femoral component in hip resurfacing. Results of 600 Conserve Plus with a 3 to 9 year follow-up. J Arthroplasty 2007;22:481-9.	Exposure is a change in the way they did surgery
Amstutz, H. C. (2008). "Present state of metal-on-metal hybrid hip resurfacing." J Surg Orthop Adv 17(1): 12-16.	No safety data reported
Back, D. L., R. Dalziel, et al. (2005). "Early results of primary Birmingham hip resurfacings. An independent prospective study of the first 230 hips." J Bone Joint Surg Br 87(3): 324-329.	Case-series with short-term F/U
Beaule, P. E., F. J. Dorey, et al. (2004). "Risk factors affecting outcome of metal-on-metal surface arthroplasty of the hip." Clin Orthop Relat Res(418): 87-93.	Cemented acetabular components
Beaule, P. E., P. Shim, et al. (2009). "Clinical Experience of Ganz Surgical Dislocation Approach for Metal-on-Metal Hip Resurfacing." J Arthroplasty.	Case-series with short-term F/U
Bergeron, S. G., N. M. Desy, et al. (2009). "The early results of metal-on-metal hip resurfacing - a prospective study at a minimum two-year follow-up." Bull NYU Hosp Jt Dis 67(2): 132-134.	Case-series with short-term F/U
Boyd, H. S., S. D. Ulrich, et al. (2007). "Resurfacing for Perthes disease: an alternative to standard hip arthroplasty." Clin Orthop Relat Res 465: 80-85.	Case-series with short-term F/U
Costi, K., D. W. Howie, et al. (2009). "Long-Term Survival and Reason for Revision of Wagner Resurfacing Hip Arthroplasty." J Arthroplasty.	Cemented acetabular components (not modern HR)
Daniel, J., P. B. Pynsent, et al. (2004). "Metal-on-metal resurfacing of the hip in patients under the age of 55 years with osteoarthritis." J Bone Joint Surg Br 86(2): 177-184.	Case-series with short-term F/U
De Smet KA. Belgium experience with metal-on-metal surface arthroplasty. Orthop Clin North Am 2005;36:203-13, ix	Case-series with short-term F/U
Hart, A. J., W. Dandachli, et al. (2009). "Large ball metal on metal hips obscure cup angle measurement on plain radiographs." Hip Int 19(4): 323-329.	Lab study, no clinical data
Hing, C., D. Back, et al. (2007). "Hip resurfacing: indications, results, and conclusions." Instr Course Lect 56: 171-178.	Review with no primary data

Study	Reason for exclusion
Howie, D. W., D. Campbell, et al. (1990). "Wagner resurfacing hip arthroplasty. The results of one hundred consecutive arthroplasties after eight to ten years." J Bone Joint Surg Am 72(5): 708-714.	Long term follow-up on discontinued total HR system (not modern total HR)
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 2005;36:215-22, ix.	Case-series with short-term F/U
Marker, D. R., T. M. Seyler, et al. (2007). "Femoral neck fractures after metal-on-metal total hip resurfacing: a prospective cohort study." J Arthroplasty 22(7 Suppl 3): 66-71.	Risk factor for femoral neck fracture (used for context on learning curve)
McGrath, M. S., D. R. Marker, et al. (2009). "Surface replacement is comparable to primary total hip arthroplasty." Clin Orthop Relat Res 467(1): 94-100.	Data included both total HR and hemi HR
Mont, M. A., T. M. Seyler, et al. (2007). "Effect of changing indications and techniques on total hip resurfacing." Clin Orthop Relat Res 465: 63-70.	Case-series with short-term F/U
Moroni, A., L. Savarino, et al. (2008). "Does ion release differ between hip resurfacing and metal-on-metal THA?" Clin Orthop Relat Res 466(3): 700-707.	Lab study, no clinical data
Naal, F. D., N. A. Maffiuletti, et al. (2007). "Sports after hip resurfacing arthroplasty." American Journal of Sports Medicine 35(5): 705-711.	Case-series with short-term F/U
O'Neill, M., P. E. Beaule, et al. (2009). "Canadian academic experience with metal-on-metal hip resurfacing." Bull NYU Hosp Jt Dis 67(2): 128-131.	Case-series with short-term F/U
Sandri, A., D. Regis, et al. (2009). "Hip resurfacing using the anterolateral Watson-Jones approach in the supine position." Orthopedics 32(6): 406.	Case-series with short-term F/U
Schmalzried, T. P., P. C. Peters, et al. (1996). "Long-duration metal-on-metal total hip arthroplasties with low wear of the articulating surfaces." J Arthroplasty 11(3): 322-331.	Early discontinued total HR systems (not modern HR)
Shimmin, A. J. and D. Back (2005). "Femoral neck fractures following Birmingham hip resurfacing." Journal of Bone and Joint Surgery - Series B 87(4): 463-464.	Physician survey without reports of response rate
Siebel, T., S. Maubach, et al. (2006). "Lessons learned from early clinical experience and results of 300 ASR hip resurfacing implantations." Proc Inst Mech Eng H 220(2): 345-353.	Case-series with short-term F/U
Springer, B. D., S. E. Connelly, et al. (2009). "Cementless femoral components in young patients: review and meta-analysis of total hip arthroplasty and hip resurfacing." J Arthroplasty 24(6 Suppl): 2-8.	Meta-analysis of case series; did not describe how event rate was calculated
Steffen, R. T., H. P. Pandit, et al. (2008). "The five-year results of the Birmingham Hip Resurfacing arthroplasty: an independent series." J Bone Joint Surg Br 90(4): 436-441.	Case-series with short-term F/U
Witzleb, W. C., M. Arnold, et al. (2008). "Birmingham hip resurfacing arthroplasty: Short-term clinical and radiographic outcome." European Journal of Medical Research 13(1): 39-46.	Case-series with short-term F/U

Special populations:

Study	Reason for exclusion
Steffen RT, Foguet PR, Krikler SJ, Gundle R, Beard DJ, Murray DW. Femoral neck fractures after hip resurfacing. J Arthroplasty 2009;24:614-9.	Risk factors for femoral neck fracture

Efficacy and safety of revisions

Study	Reason for exclusion
Jaiswal, A., R. E. Gilbert, et al. (2011). "Function and survival after revision of hip resurfacing." Hip Int **21**(5): 610-615.	Does not report main outcomes of interest
Sandiford, N. A., S. K. Muirhead-Allwood, et al. (2010). "Revision of failed hip resurfacing to total hip arthroplasty rapidly relieves pain and improves function in the early post operative period." J Orthop Surg Res **5**(88): 5-88.	No comparison
Schmitz, M. W., V. J. Busch, et al. (2013). "Long-term results of cemented total hip arthroplasty in patients younger than 30 years and the outcome of subsequent revisions." BMC Musculoskelet Disord **14**(37): 1471-2474.	Does not report main outcomes of interest

Appendix D. CLASS OF EVIDENCE DETERMINATION

Critical appraisal, risk of bias and overall study quality determination

Each study was critically appraised based on a set of general pre-set criteria listed in the Tables below as an initial starting point for identify risk of bias. The resulting worksheets provide insight into overall quality of individual studies.

Definition of the class of evidence and risk of bias for studies on therapy

		Studies of Therapy	
Class	**Bias Risk**	**Study design**	**Criteria**
I	**Low risk:** Study adheres to commonly held tenets of high quality design, execution and avoidance of bias	Good quality RCT	• Random sequence generation • Allocation concealment • Intent-to-treat analysis • Blind or independent assessment for important outcomes • Co-interventions applied equally • F/U rate of 80%+ • Adequate sample size
II	**Moderately low risk:** Study has potential for some bias; study does not meet all criteria for class I, but deficiencies not likely to invalidate results or introduce significant bias	Moderate or poor quality RCT	• Violation of one of the criteria for good quality RCT
		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	**Moderately High risk:** Study has significant flaws in design and/or execution that increase potential for bias that may invalidate study results	Moderate or poor quality cohort	• Violation of any of the criteria for good quality cohort
		Case-control	• Any case-control design
IV	**High risk:** Study has significant potential for bias; lack of comparison group precludes direct assessment of important outcomes	Case series	• Any case series design

* Outcome assessment is independent of healthcare personnel judgment. Reliable data are data such as mortality or re-operation.

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

Definition of the class of evidence and risk of bias for registry studies

		Registry Studies	
Class	**Risk of Bias**	**Study design**	**Criteria**
II	**Moderately low risk:** Study has potential for some bias; does not meet all criteria for class I but deficiencies not likely to invalidate results or introduce significant bias	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 > 85% • Controlling for possible confounding† • Accounting for time at risk‡
III	**Moderately high risk:** Study has flaws in design and/or execution that increase potential for bias that may invalidate study results	Moderate quality registry	• Prospective data from registry designed specifically for conditions evaluated with violation of 2 of the rest of the criteria in level I
IV	**High risk:** Study has significant potential for bias; does not include design features geared toward minimizing bias and/or does not have a comparison group	Poor quality registry	• 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.

Methods for critical appraisal and level of evidence assessment

The method used for assessing the quality of evidence of individual studies as well as the overall quality of evidence incorporates aspects of rating scheme developed by the Oxford Centre for Evidence-based Medicine, [Oxford Centre for Evidence-based Medicine Levels of Evidence 2009] precepts outlined by the Grades of Recommendation Assessment, Development and Evaluation (GRADE) Working Group,[1] and recommendations made by the Agency for Healthcare Research and Quality (AHRQ).[4] Taking into account features of methodological quality and important sources of bias combines epidemiologic principles with characteristics of study design.

Determination of Overall Strength of Evidence (Overall quality of evidence)

After individual article evaluation, the overall body of evidence with respect to each outcome is determined based on precepts outlined by the Grades of Recommendation Assessment, Development and Evaluation (GRADE) Working Group[1] and recommendations made by the Agency for Healthcare Research and Quality (AHRQ).[4] Qualitative analysis is performed considering AHRQ required and additional domains [OWENS].
The initial strength of the overall body of evidence was considered HIGH for RCTs and LOW for observational studies. The body of evidence may be downgraded one or two levels based on the following criteria: (1) risk of bias (study limitations), (2) inconsistency of results, (3) indirectness of evidence, (4) imprecision of the effect estimates (e.g., wide confidence intervals) or (4) failure to provide am *a priori* statement of subgroup analyses. The body of evidence may be upgraded one or two levels based on the following criteria: (1) large magnitude of effect or (2) dose-response gradient (3) if all plausible biases would decrease the magnitude of an apparent effect. The final overall strength of the body of literature expresses our confidence in the estimate of effect and the impact that further research may have on the results. Interpretation of the strength of evidence categories, based on the AHRQ Methods Guide are as follows:

High - Very confident that effect size estimates lie close to the true effect for this outcome; there are few or no deficiencies in the body of evidence; we believe the findings are stable

Moderate – Moderately confident that effect size estimates lie close to the true effect for this outcome; some deficiencies in the body of evidence; we believe the findings are likely to be stable but some doubt remains

Low – Limited confidence that effect size estimates lie close to the true effect for this outcome; major or numerous deficiencies in the body of evidence; we believe that additional evidence is needed before concluding that findings are stable or the estimate is close to the true effect

Insufficient – We have no evidence, are unable to estimate an effect or have no confidence in the effect estimate for this outcome; No available evidence or the body of evidence has unacceptable deficiencies precluding judgment.

Similar methods for determining the overall quality (strength) of evidence related to economic studies or administrative studies have not been reported.

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.[3] QHES embodies the primary components relevant for critical appraisal of economic studies.[2,3] 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[3] **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. SUMMARY OF STUDIES REPORTING FROM REGISTRIES

Author (Year)	Registry Study Period	Intervention	Outcomes	Results	Conflict of Interest
Garbuz (2009)	Australian 2005–2008	MoM HR (n = 4800) MoM THA (n = 5600)	• Revision	HR: 70/1000 (7.0%) THR: 140/2000 (7.0%)	The institution of one or more of the authors has received funding from Zimmer, Inc. (Warsaw, IN)
Amanatullah (2010)	Australian	MoM HR THA	• Revision	• SARI* < 3, 3% revision rate • SARI > 3, 11% revision rate • Hip resurfacing arthroplasty survivorship at 3 and 5 years is equivalent to THA. • 0.02% revision rate in men younger than 55 years with osteoarthritis	The authors of this article have received no benefits, funding, or support in conjunction with this report.
Corten (2010)	2008 Australian, 2008 English (and Welsh), 2006 and 2007 Swedish	MoM HR THA	• Revision Reasons for revision: • Aseptic loosening • Fracture • Dislocation	English Registry: • HR revision rate: 1.6% (1 yr), 1.8% (3 yr) • THR revision rate: 0.3% (1 yr) Australian Registry: • HR revision rate: 3.7% (5 yr), 4.6% (7 yr) • THA revision rate: 2.7% (5 yr), 3.4% (7 yr) Swedish Registry: • HR revision rate: 3.4% (yr NR) Resurfacing was associated with an overall increased failure rate in comparison to THA. Reason for revision: • HR: aseptic loosening (31%), fracture (31%) • THR: aseptic loosening (46%), dislocation (19%)	One the authors (SJM) is a consultant with DePuy.

* Amanatullah (2010): surface arthroplasty risk index = SARI, (SARI is calculated by summing assigned numbers for specific risks: 1 for previous surgery, 2 if less than 82 kg, 1 for high activity, and 2 for femoral cysts greater than 1 cm.)

Author (Year)	Registry Study Period	Intervention	Outcomes	Results	Conflict of Interest
Corten (2011)	2008 Australian, 2006 and 2007 Swedish, 2008 English (and Welsh)	MoM HR THA	• Revision • Re-revision Reasons for revision: • Femoral neck fracture • aseptic loosening	• Femoral neck fractures have a prevalence of 1.0% to 5.6% • Aseptic loosening has a prevalence of 1.0% to 2.0% • HRA had a three- to fivefold increased risk for revision in comparison to THA • HRA revision rate: 3.7% (5 yr), 5.3% (8 yr) • THA revision rate: 2.8% (5 yr), 4% (8 yr) • HRA re-revision rate: 8.4% (3 yr), 11% (5-yr) • THA re-revision rate: 8.2% (3 yr)	The authors declare no conflicting interests.
Macpherson (2011)	2009 Australian, 2010 England and Wales	MoM HR THA	• Revision • Survivorship	3 yr Revision Rate for patients <55:* HR: UK registry: 3.7% (men), 5.7% (women) Austrailian registry: 1.9% (men), 3.6% (women) THR: UK registry: 2.5% (men), 2.5% (women) Austrailian registry: 1.8% (men), 2.7% (women) Survivorship: 95.9% (mean follow-up 4.2 years) • Revision risk for HR is 1.6 times higher than for cemented THA in England and Wales	The authors declare that they have no conflict of interest and have not received any outside funding or grants.

* Macpherson (2011): See supplementary table for revision rates broken down by registry, age and sex

Author (Year)	Registry Study Period	Intervention	Outcomes	Results	Conflict of Interest
Cuckler (2011)	Australian (2010)	MoM HR THA	• Revision	5 year revision rate: HR: 3.8% THA: 2.8% 9 year cumulative revision rate: HR: 7.2% THA: 5.4%	NR
de Steiger (2011)	Australian (1999-2008)	THR MoM HR: -ASR XL Acetabular System -ASR Hip Resurfacing System -other hip resurfacing	• Revision	5-year revision rates: THR revision rate:* 3.4% (95% CI, 3.3% to 3.5%) ASR XL Acetabular revision rate: 9.3% (7.3% to 11.9%) ASR HR revision rate: 10.9% (8.7% to 13.6%) Other hip resurfacing: 4.0% (3.7% to 4.5%)	None of the authors received payments or services, either directly or indirectly (i.e., via his or her institution), from a third party in support of any aspect of this work. One or more of the authors, or his or her institution, has had a financial relationship, in the thirty-six months prior to submission of this work, with an entity in the biomedical arena that could be perceived to influence or have the potential to influence what is written in this work. Also, one or more of the authors has had another relationship, or has engaged in another activity, that could be perceived to influence or have the potential to influence what is written in this work.

* See supplementary table for revision rates broken down by registry, age and sex

Author (Year)	Registry Study Period	Intervention	Outcomes	Results	Conflict of Interest
Seppanen (2012)	England and Wales, Nordic Arthroplasty Register (2001-2009)	THR HR	• Revision • Survival	There was no statistically significant difference in revision risk between HRAs and THAs (RR = 0.93, 95% CI: 0.78-1.10). Female patients had about double the revision risk of male patients (RR = 2.0, CI: 1.4-2.7)*	NR
Smith, Deippe (2012)	National Joint Registry for England and Wales (NJR) (2003-2011)	HR Stemmed THR	• Revision • Survival	5-year Revision Rate: HR: 5.2% (95% CI 4.9–5.5) THR: 2.8% (2.7–2.9)*	We declare that we have no conflicts of interest.
Naal (2011)	Australian Registry	Durom resurfacing implant	• Revision	-5-year cumulative revision rate of 6.7% -11 revisions in 100 hips required at 5 years -Survival rate of 88.2% -Women had an even higher revision rate of 17%	Each author certifies that he or she has no commercial associations (eg, consultancies, stock ownership, equity interest, patent/licensing arrangements, etc) that might pose a conflict of interest in connection with the submitted article

* See supplementary table for revision rates broken down by age, sex and resurfacing type.
* Smith, Deippe (2012): See supplementary tables for a breakdown of revision rates by sex and brand.

Author (Year)	Registry Study Period	Intervention	Outcomes	Results	Conflict of Interest
Jameson (2012)	England and Wales (2003-2010)	MoM HR Brands (BHR,* ASR,* Adept, Conserve, Cormet, Durom Mitch, Recap)	• Revision	5 year revision rate: 3.59% (1,003 / 27,971) Women have greater risk of revision than men (hazard ratio = 1.30, p = 0.007) Brands with significantly greater risk of revision than BHR:** ASR: HR = 2.82, p < 0.001 Conserve: HR = 2.03, p < 0.001 Cormet: HR = 1.43, p = 0.001 Durom: HR = 1.67, p < 0.001 Recap: HR = 1.58, p = 0.007	The author or one or more of the authors have received or will receive benefits for personal or professional use from a commercial party related directly or indirectly to the subject of this article.
McMinn (2012)	England and Wales (2003-2011)	Stemmed THR Birmingham Hip Resurfacing	• Revision • Mortality	Revision rates (8 years): All men: Revision: 159/8,352 (1.9%) Death: 93/8,352 (1.1%) Men <55 years of age: Revision: 60/3,560 (1.7%) Death: 10/3,560 (0.3%)	Two authors were designers of Birmingham hip resurfacing and were shareholders in Midland Medical Technologies before the company was sold to Smith and Nephew in 2004. One author is an unpaid consultant to Smith and Nephew Orthopaedics UK. The McMinn Centre receives institutional research funding for specific research projects but no funding has been received for the research or preparation of this manuscript.

* Jameson (2012): Birmingham Hip Resurfacing (BHR), Articular Surface Replacement (ASR)
** Jameson (2012): See supplementary table for revision rates broken down by year for each brand of hip resurfacing.

Author (Year)	Registry Study Period	Intervention	Outcomes	Results	Conflict of Interest
Prosser (2010)	Australian Registry	HR THR	• Revision	Total HR: 437/12,093 (3.6%) 8 year revision rate: HR: 5.3 (4.6–6.2) THR: 4.0 (3.8–4.2)*	No competing interests declared.

* Prosser (2010): See supplementary tables for revision rates broken down by resurfacing brand 1-8 years.

Appendix F. LEVEL OF EVIDENCE FOR COMPARATIVE STUDIES.

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

Methodological principle	Vendittoli (2006)	Rama (2009)	Vendittoli, Ganapathi (2010)	Vendittoli, Roy (2010)	Garbuz (2010)	Lavigne (2010)	Smolders (2010)	Smolders (2011)	Jensen (2011)	Costa (2012)
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	**II**	**II**	**II**	**II**	**II**	**II**	**II**	**II**	**II**	**I**

Methodological principle	Fowble (2009)	Li (2009)	Li (2008)	Mont (2009)	Pattyn (2008)	Pollard (2006) Baker (2011)	Stulberg (2008, 10)	Vail (2006)	Zywiel (2009)	Costa (2011)	Killampalli (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**	**III**	**III**	**III**	**III**	**III**	**III**	**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

‡ 70% for 10 year follow-up (Baker 2011)

§87% for 2 year follow-up, 45% for 3 year follow-up

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.

QHES

Study: Bozic 2010	Points	Notes:
1. Was the study objective presented in a clear, specific, and measurable manner?	**7**	*Use ICER to evaluate CE*
2. Were the perspective of the analysis (societal, third-party payer, etc.) and reasons for its selection stated?	**4**	*Healthcare provider*
3. Were variable estimates used in the analysis from the best available source (ie, randomized controlled trial - best, expert opinion - worst)?	**8**	*AOA- gave rationale for doing so*
4. If estimates came from a subgroup analysis, were the groups prespecified at the beginning of the study?	**1**	*Not applicable*
5. Was uncertainty handled by (1) statistical analysis to address random events, (2) sensitivity analysis to cover a range of assumptions?	**9**	*Used one-way, two-way and probabilistic sensitivity analysis to measure impact of data sources and other important variables*
6. Was incremental analysis performed between alternatives for resources and costs?	**6**	*Table 3*
7. Was the methodology for data abstraction (including the value of health states and other benefits) stated?	**5**	*Derived from literature review.*

Study: Bozic 2010	Points	Notes:
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**	*Used a 30-year time horizon. Discounted at 5%.*
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**	*Evaluated different data sources*
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**	*Markov decision model and structure clearly defined.*
13. Were the choice of economic model, main assumptions, and limitations of the study stated and justified?	**7**	*Assumptions given, and limitations discussed*
14. Did the author(s) explicitly discuss direction and magnitude of potential biases?	**6**	*Discussed with limitations*
15. Were the conclusions/recommendations of the study justified and based on the study results?	**8**	*Conclusion more evidence is necessary*
16. Was there a statement disclosing the source of funding for the study?	**3**	*Orthopedic Research and Education Foundation*
TOTAL POINTS	**100**	

Study: Edlin 2012	Points	Notes:
1. Was the study objective presented in a clear, specific, and measurable manner?	**7**	*Compare CE of CAS and CEA using simulation*
2. Were the perspective of the analysis (societal, third-party payer, etc.) and reasons for its selection stated?	**4**	*NHS*
3. Were variable estimates used in the analysis from the best available source (ie, randomized controlled trial - best, expert opinion - worst)?		*Small sample size, single-center source*
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**	*Used one-way sensitivity analysis*
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**	*EQ-5D*
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?		*Only 1-year Discounted at 3%*
9. Was the measurement of costs appropriate and the methodology for the estimation of quantities and unit costs clearly described?	**8**	*Cost estimates obtained from hospital*
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**	*Described details of calculations and justified*
13. Were the choice of economic model, main assumptions, and limitations of the study stated and justified?	**7**	*Assumptions given, and limitations discussed Minimal justifications were provided*
14. Did the author(s) explicitly discuss direction and magnitude of potential biases?	**6**	*Discusses limitations and makes comparisons to other studies*
15. Were the conclusions/recommendations of the study justified and based on the study results?	**8**	*Conclusion suggested need for more evidence*
16. Was there a statement disclosing the source of funding for the study?	**3**	
TOTAL POINTS	**85**	

Study: Buckland 2008	Points	Notes:
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)?	**0**	Heavy reliance on expert opinion, including for costs
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?	**0**	Appeared to only do threshold analysis for few variables
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**	Detailed description of model inputs
13. Were the choice of economic model, main assumptions, and limitations of the study stated and justified?	**0**	Model choice not stated
14. Did the author(s) explicitly discuss direction and magnitude of potential biases?	**0**	Little/no discussion
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?	**0**	
TOTAL POINTS	**67**	

Study: McKenzie	Points	Notes:
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**	Well-conducted study sponsored by NICE

Note: We did not do QHES on Vale as the original submission is not published in the Vale HTA--also several criteria were prespecified by NICE to submitters that are consistent with QHES (statement of perspective/rationale, research question) that might influence results.

Appendix G. Individual Study Results from Randomized Clinical Trials and Comparative Observational Studies.

Randomized Controlled Trials

Author	Outcome	F/U	HR	THA	P-value
FUNCTIONAL	WOMAC	Pre	52.7 (15.4)	54.4 (18.3)	.548
		1 yr	8.0 (13.2)	10.2 (10.7)	
Vendittoli/Rama		2 yr	5.7 (8.6)	9.0 (11.9)	.007
(2006/2009)	Thigh pain	1 yr	4%	5%	.661
and		2 yr	2%	4%	.337
Vendittoli,	Hip pain	1 yr	31%	41%	.124
Ganapathi (2010)		2 yr	14%	20%	.265
	Perception:				
	Feels like natural hip	2 yr	53%	58%	.471
	Artificial, no limitation		20%	21%	
	Artificial, min or sig limitation		27%	20%	
	Merle d'Aubigne	Pre	10.8 (2.8)	10.4 (2.5)	.35
		1 yr	16.7	16.6	NS
		2 yr	17.5 (1.3)	17.5 (1.3)	.88
	UCLA Activity	1 yr	6.3	7.1	0.037
		2 yr	7.5 (1.8)	7.1 (1.6)	0.094
	Hop Test:				
	Very easy	2 yr	78%	70%	0.268
	Easy		17%	17%	
	Difficult/Impossible		4%	13%	
Lavigne (2010)	WOMAC	Pre	46.5 (14.9)	54.3 (14.5)	NS
		Post	3.0 (8.4)	2.7 (8.5)	NS
	SF-36 Mental	Pre	34.3 (8.1)	35.1 (7.2)	NS
		Post	51.9 (7.2)	52.1 (10.9)	NS
	SF-36 Physical	Pre	47.7 (10.1)	46.8 (12.1)	NS
		Post	55.2 (5.1)	53.3 (8.7)	NS
	Feels like natural hip	Post	62%	58%	0.775
	Artificial, no limitation		21%	29%	
	Artificial, min or sig limitation		17%	13%	
	Timed up and go (seconds)	Pre	7.60 (1.70)	8.00 (1.04)	NS
		Post	6.73 (1.00)	7.07 (0.78)	NS
	Step (seconds)	Post	18.12 (3.57)	15.00 (3.10)	< .05
	Merle D'Aubigne	Pre	11.0 (2.8)	10.5 (2.3)	NS
		Post	17.9 (0.4)	18.0 (0.0)	NS
	UCLA Activity	Post	8.0 (1.5)	8.3 (1.7)	NS
Garbuz (2010)	WOMAC Pain	Pre	48.94	52.36	

Author	Outcome	F/U	HR	THA	P-value
		Post	91.52	90.00	0.696
	WOMAC Stiffness	Pre	47.07	42.86	
		Post	85.61	83.13	0.546
	WOMAC Function	Pre	52.15	53.71	
		Post	90.65	91.07	0.905
	WOMAC Global	Pre	51.06	52.61	
		Post	90.41	90.19	0.950
	SF-36 Physical	Pre	32.73	33.60	
		Post	51.23	51.29	0.979
	SF-36 Mental	Pre	46.65	50.65	
		Post	53.88	55.14	0.555
	UCLA Activity	Pre	4.9	4.7	
		Post	6.8	6.3	0.240
Smolders (2011)	SF-12	Pre	88 (59-112)	79 (55-113)	< .05
		Post	110 (69-117)	110 (51-133)	NS
	Oxford Hip Score	Pre	34 (20-46)	37 (21-44)	NS
		Post	13 (12-34)	16 (12-37)	< .05
	VAS Satisfaction	Pre	89 (49-100)	82 (10-100)	NS
		Post	92 (37-100)	89 (15-100)	NS
	Harris Hip	Pre	57 (28-77)	53 (25-82)	NS
		Post	96 (63-100)	95 (47-100)	NS
	UCLA Activity	Pre	5 (2-10)	4 (2-8)	< .05
		Post	8 (5-10)	7 (2-10)	< .05
Costa (2012)	Oxford Hip Score	Pre	19.1 (8.0)	19.6 (7.8)	
		Post	40.4 (37.9-42.9)	38.2 (35.3-41.0)	
	EQ-5D Score	Pre	0.33 (0.34)	0.36 (0.33)	
		Post	0.796 (.721-.870)	0.719 (.636-.802)	
	EQ-5D VAS score	Pre	56.7 (22.9)	57.8 (24.3)	
		Post	78.2 (73.9)	76.2 (70.8-81.5)	
	Disability Rating Index	Pre	57.0 (16.5)	57.9 (18.2)	
		Post	27.7 (21.7-33.7)	34.8 (28.4-41.2)	
	Paffenbarger physical activity	Pre	8.53 (11.40)	6.55 (6.72)	
		Post	15.01 (10.15-19.87)	13.85 (10.90-17.80)	
	Harris Hip Score	Pre	48.6 (14.2)	50.1 (13.5)	
		Post	88.4 (84.4-92.4)	82.3 (77.2-87.5)	
RADIOGRAPHIC					
Vendittoli/Rama (2006/2009)	Acetabular vertical angle (°)	1 yr	47.3 (31-64)	45.3 (30-55)	0.05
	Leg Length discrepancy (mm)	Pre	-1.6 (-14.6-4.0)	-1.3 (-15.5-9.7)	NS
		1 yr	-1.8 (-9.9-5.9)	3.0 (-6.0-12.9)	<.001

Author	Outcome	F/U	HR	THA	P-value
	Leg length discrepancy correction (mm)	1 yr	0.1 (-5.8-5.5)	1.8 (-12.3-10.7)	0.001
Vendittoli, Roy (2010)	Femoral Offset (mm)		33.4 (10.2-46.6)	37.9 (13.9-53.2)	<0.001
	Acetabular component abd angle (°)		46.6 (31.2-61.0)	45.1 (34.5-55.7)	0.186
	Femoral component valgus angle (°)		141.5 (121-154)	NA	
Vendittoli, Ganapathi (2010)	Loosening of acetabular components		0 (0%)	0 (0%)	
	Femoral neck narrowing (>10%)		1 (0.9%)	0 (0%)	
	Migration or radiolucent line (loose)		2 (1.8%)	0 (0%)	
Lavigne (2010)	Femoral offset difference (mm)		-3.3 (4.8)	0.9 (6.3)	0.013
	N (5) of patients with femoral offset within ± 4 mm		14 (58%)	9 (38%)	0.248
	Leg length inequality (mm)		-0.4 (2.8)	-0.1 (4.3)	0.782
	N (%) of patients with legth length inequality within ± 4 mm		21 (88%)	17 (71%)	0.286
Smolders (2010)	BMD (hip-femoral next) (g/cm^2)	Pre	1.87 (0.32)	1.90 (0.39)	NS
		Post	1.97 (0.40)	1.53 (0.40)	<0.001
	Bone mineral density (hip) ratio (%)	Pre	100.00	100.00	NS
		Post	105.2 (9.7)	82.1 (14.6)	<0.001
Jensen (2011)	Total offset (mm)	Pre	68.0 (8.7)	68.6 (8.9)	
		Post	67.4 (9.4)	67.5 (8.5)	
	Femoral offset (mm)	Pre	34.1 (4.9)	33.1 (5.9)	
		Post	33.9 (6.3)	36.8 (7.8)	
	Cup offset (mm)	Pre	33.5 (5.5)	34.9 (5.1)	
		Post	34.1 (4.3)	29.9 (3.3)	
MOTION/STRENGTH					
Lavigne (2010)	Functional reach (cm)	Pre	37.2 (5.2)	36.1 (3.9)	
		Post	39.2 (5.8)	34.6 (4.3)	<.05
	Hip flexor strength ratio(%)	Pre	77.0 (16.4)	81.3 (27.6)	
		Post	91.5 (15.3)	92.1 (7.5)	NS
	Abductor strength ratio (%)	Pre	82.7 (22.1)	82.1 (17.4)	
		Post	92.6 (9.9	89.4 (16.2)	NS
Jensen (2011)	Knee extension (Nm)	Pre	110	108	
		Post	119	136	< .05
	Knee flexion (Nm)	Pre	48	52	
		Post	56	63	NS
	Hip Adduction (Nm)	Pre	112	111	

Author	Outcome	F/U	HR	THA	P-value
		Post	117	124	NS
	Hip Abduction (Nm)	Pre	94	108	
		Post	129	152	< .05
	Hip Extension (Nm)	Pre	152	140	
		Post	183	216	0.06
	Hip Flexion (Nm)	Pre	99	101	
		Post	115	124	NS
GAIT					
Lavigne (2010)	Normal walking speed (m/sec)	Pre	1.19 (0.29)	1.03 (0.20)	0.036
		Post	1.44 (0.19)	1.46 (0.18)	NS
	Step length (m)	Pre	0.64 (0.08)	0.58 (0.06)	0.038
		Post	0.68 (0.07)	0.69 (0.06)	NS
	Cadence (steps/min)	Pre	110.1 (16.2)	106.8 (11.5)	NS
		Post	125.6 (7.5)	126.2 (8.7)	NS
	Fast walking speed (m/sec)	Pre	1.58 (0.29)	1.50 (0.22)	NS
		Post	1.82 (0.24)	1.73 (0.18)	NS
	Postural balance	Pre	113.8 (32.9)	124.8 (20.7)	NS
		Post	108.1 (20.8)	112.3 (24.0)	NS
Vendittoli, Ganapathi (2010)	Step test	Post			
	Very easy		78%	70%	0.268
	Easy		17%	17%	
	Difficult		2%	5%	
	Impossible		2%	8%	
SAFETY					
Vendittoli (2006)	Revision	1 yr	2/103 (1.9%)	1/102 (1.0%)	
	Loosening of femoral head		2 (1.9%)	0 (0%)	
	Intraop conversion to THA		2 (1.9%)	NA	
	Intraop conversion to different type of fixation or component		2 (1.9%)	1 (1.0%)	
	Dislocation		0 (0%)	3 (2.9%)	
	Deep infection		0 (0%)	2 (2.0%)	
	Intraoperative acetabular fissure		2 (1.9%)	0 (0%)	
	Intraoperative proximal femoral fissure		0 (0%)	4 (3.9%)	
	Deep vein thrombosis		2 (1.9%)	2 (2.0%)	
	Sciatic neurapraxia		1 (1.0%)	2 (2.0%)	
Rama (2009)	Heterotopic ossification		44 (42.7%)	30 (29.4%)	
Vendittili, Ganapathi (2010)	Revision		4/109 (3.7%)	2/100 (2.0%)	
	Intra-operative acetabular fracture (uneventful)		2 (1.8%)	0 (0%)	

Author	Outcome	F/U	HR	THA	P-value
	Intra-operative proximal femoral fracture (uneventful)		0 (0%)	4 (4.0%)	0.038
	Deep vein thrombosis (clinically symptomatic)		1 (0.9%)	3 (3.0%)	
	Neurapraxia (sciatic)		1 (0.9%)	2 (2.0%)	
	Deep Infection		0 (0%)	5 (5.0%0	0.02
	Early without recurrence		0 (0%)	4 (4.0%)	
	Chronic		0 (0%)	1 (1.0%)	
	Dislocation		0 (0%)	4 (4.0%)	0.038
	Simple, without recurrence		0 (0%)	2 (2.0%)	
	Recurrent dislocation		0 (0%)	2 (2.0%)	
	Femoral aseptic loosening		6 (5.5%)	0 (0%)	0.017
	Symptomatic leg length discrepancy		0 (0%)	1 (1.0%)	
	Symptomatic femoro-acetabular impingement		2 (1.8%)	0 (0%)	
	Symptomatic heterotopic ossification		2 (1.8%)	0 (0%)	
Vendittoli, Roy (2010)	Whole blood chromium levels (µg/L)	Pre	0.88 (0.40-2.20)	1.03 (0.40-5.00)	0.314
		1 yr	1.56 (0.40-5.50)	1.50 (0.40-2.90)	0.775
		2 yr	1.58 (0.40-3.70)	1.62 (0.80-5.70)	0.819
	Whole blood cobalt levels (µg/L)	Pre	0.16 (0.06-1.05)	0.15 (0.06-0.42)	0.897
		1 yr	0.67 (0.23-2.09)	0.81 (0.23-2.10)	0.074
		2 yr	0.67 (0.20-2.89)	0.94 (0.24-4.89)	0.207
	Whole blood titanium levels (µg/L)	Pre	0.58 (0.10-1.50)	0.57 (0.10-1.50)	0.967
		1 yr	3.05 (1.00-8.40)	1.83 (0.90-4.60)	< 0.0001
		2 yr	1.87 (0.40-4.90)	1.30 (0.35-2.40)	0.001
Garbuz (2010)	Serum cobalt levels (µg/L)	Pre	0.13	0.11	0.565
		1 yr	0.51	5.09	0.000
		2 yr	0.54 (0.4-0.7)	5.38 (3.5-7.2)	
		Pre	0.15	0.20	0.608
		1 yr	0.81	2.14	0.023
		2 yr	0.84 (0.7-1.1)	2.88 (1.1-4.0)	
Lavigne (2010)	Acetabular/femoral loosening		0 (0%)	0 (0%)	
	Femoral calcar cracks		0 (0%)	3 (12.5%)	
	Obturator artery damage		1 (4.2%)	0 (0%)	
	Myocardial infarction		1 (4.2%)	0 (0%)	
Smolders (2011)	Recurrent dislocation		0 (0%)	3 (9.1%)	
	Early deep infection		0 (0%)	2 (6.1%)	

Author	Outcome	F/U	HR	THA	P-value
	Aseptic loosening from avascular necrosis		1 (2.6%)	0 (0%)	
	Whole blood cobalt levels (µg/L)	Pre	0.1 (0.1-0.8)	0.1 (0.1-0.6)	NS00.1
		1 yr	1.25 (0.6-8.3)	1.0 (0.1-4.2)	0.1
		2 yr	1.2 (0.5-2.2)	0.9 (0.1-2.7)	NS
	Whole blood chromium levels (µg/L)	Pre	0.1 (0.1-1.4)	0.1 (0.1-0.1)	< .05
		1 yr	1.0 (0.1-6.1)	0.5 (0.1-2.0)	< .05
		2 yr	1.2 (0.1-10)	0.5 (0.1-2.1)	
Costa (2012)	Deep infection		0 (0%)	2 (3.0%)	0.497
	Deep vein thrombosis		4 (6.7%)	0 (0%)	0.049
	Superficial wound complications		2 (3.3%)	9 (13.6%)	0.057
	Dislocation		1 (1.7%)	1 (1.5%)	1.000
	Other		4 (6.7%)	6 (9.1%)	0.747
	Total		11 (18.3%)	18 (27.3%)	0.291

Observational Studies

Author (Year)	Outcome		HR	THA	P-value
Fowble (2009)	Harris Hip Score	Pre	46 (9)	52 (11)	0.005
	Harris Hip Score	Post	97 (4)	96 (7)	0.4
Mean F/U: 2.9 yrs	UCLA Activity	Pre	4.2 (1.1)	3.6 (1.4)	0.02
	UCLA Activity	Post	8.2 (1.6)	5.9 (1.7)	0.0001
	SF-12 Mental	Pre	44.2 (12.8)	35.2 (15.8)	NS
	SF-12 Mental	Post	54.6 (6.7)	52.5 (9.1)	NS
	SF-12 Physical	Pre	33.6 (8.4)	25.8 (1.6)	NS
	SF-12 Physical	Post	53.6 (5.9)	47.0 (13.1)	0.002
	Function (HHS-(pain+deformity+ROM)	Pre	27.3 (8.3)	29.9 (7.4)	NS
	Function (HHS-(pain+deformity+ROM)	Post	46.4 (1.4)	44.9 (3.3)	0.007
	Pain				All pain:
	No pain	Pre	0 (0%)	0 (0%)	0.0001
	No pain	Post	28 (57%)	32 (80%)	0.007
	Slight pain	Pre	0 (0%)	0 (0%)	
	Slight pain	Post	18 (37%)	6 (15%)	
	Mild pain	Pre	0 (0%)	0 (0%)	
	Mild pain	Post	3 (6%)	0 (0%)	
	Moderate pain	Pre	3 (6%)	17 (42%)	
	Moderate pain	Post	0 (0%)	2 (5%)	
	Marked pain	Pre	47 (94%)	23 (58%)	
	Marked pain	Post	0 (0%)	0 (0%)	
Li (2009)	Harris Hip Score	Pre	50.6 (6.1)	50.3 (6.0)	NR
	Harris Hip Score	Post	91.0 (3.4)	89.7 (3.3)	NR
Mean F/U: NR	UCLA Activity	Pre	2.4 (1.0)	2.5 (1.2)	NR
	UCLA Activity	Post	6.1 (0.7)	3.6 (0.7)	NR
	Hip pain (VAS)	Pre	4.3 (2.2)	3.8 (2.8)	NR
	Hip pain (VAS)	Post	0.9 (0.9)	0.7 (0.9)	NR
Li (2008)	Harris Hip Score	Pre	NR	NR	
Mean F/U: 2.2 yrs	Harris Hip Score	Post	93	91	NS
Mont (2009)	Harris Hip Score	Pre	39 (24-60)	39 (24-56)	NS
	Harris Hip Score	Post	90 (50-100)	91 (62-100)	NS
Mean F/U: 3.3 yrs	Activity Score	Pre	3 (0-15)	2 (0-6)	0.01
	Activity Score	Post	11.5 (0-32)	7 (0-21)	0.0004
	Change in activity score	Pre	NR	NR	
	Change in activity score	Post	8 (0-17)	5 (0-15)	0.0004
	Satisfaction score	Pre	NA	NA	

Author (Year)	Outcome		HR	THA	P-value
	Satisfaction score	Post	9.2 (2-10)	8.8 (0-10)	NS
	Pain score	Pre	NR	NR	
	Pain score	Post	1.4 (0-6)	1.6 (0-9)	NS
Pattyn (2008)	Harris Hip Score	Pre	<50	<50	
	Harris Hip Score	Post	97.9	92.1	NR
F/U: 3-6yrs	Activity	No Pre data			
	Activities of daily living	Post	38.2%	59.1%	NR
	Independent	Post	1.0%	9.6%	NR
	Dependent	Post	0.0%	0.9%	NR
	Strenuous	Post	60.7%	30.4%	NR
Pollard (2006)	Oxford Hip Score	Pre	NR	NR	
	Oxford Hip Score	Post	15.9 (12-42)	18.5 (12-41)	NS
Mean F/U: 5.9 yrs	UCLA Activity	Pre	9.0 (6-10)	8.9 (6-10)	NR
	UCLA Activity	Post	8.4 (4-10)	6.8 (3-10)	< .001
	EQ-5D (QoL)	Pre	NR	NR	
	EQ-5D (QoL)	Post	0.9 (.08-1.0)	0.78 (.06-1.0)	0.003
	EQ-VAS (QoL)	Pre	NR	NR	
	EQ-VAS (QoL)	Post	82.3 (20-100)	69.3 (15-100)	0.001
	Patient activities:	No Pre data			
	Running	Post	31 (58.5%)	7 (13.2%)	< .001
	Played a sport	Post	39 (73.6%)	17 (32.1%)	< .001
	Performed heavy manual work	Post	32 (60.4%)	20 (37.7%)	0.049
Stulberg (2008)	Harris Hip Score	Pre	50.1 (11.6)	49.7 (11.3)	NS
	Harris Hip Score	Post	96.7 (7.5)	96.2 (7.7)	NS
F/U: >2 yrs	Composite Clinical Success	Pre	NA	NA	
	Composite Clinical Success	Post	251 (86%)	224 (87.5%)	NR
Vail (2006)	Harris Hip Score	Pre	48.5	42	< .001
	Harris Hip Score	Post	98.1	92.6	NS
Mean F/U: 3 yrs	Harris Hip Score - Pain	Pre	11.3	10.9	NR
	Harris Hip Score - Pain	Post	42.9	41.8	NS
	Harris Hip Score - Function	Pre	28.8	23.2	< .001
	Harris Hip Score - Function	Post	46.2	42.1	NS
	Activity	Pre	8.5	7.7	0.035
	Activity	Post	14	12.7	0.028
Zywiel (2009)	Harris Hip Score	Pre	52 (28-71)	49 (20-69)	NS
	Harris Hip Score	Post	91 (32-100)	90 (50-100)	NS
Mean F/U: 3.6 yrs	Activity Score	Pre	2.1 (0-6.0)	2.3 (0-6.0)	NS
	Activity Score	Post	10.0 (1.0-27.5)	5.3 (0-12.0)	< .001
	Satisfaction Score	Pre	NA	NA	
	Satisfaction Score	Post	9.1 (5-10)	9.1 (2-10)	NS

Author (Year)	Outcome		HR	THA	P-value
	Pain Score	Pre	NR	NR	
	Pain Score	Post	1.3 (0-10)	1.2 (0-5)	NS
Baker (2011)	Oxford Hip Score	Post	16.6 (12-46)	19.1 (12-43)	NS
	UCLA Activity	Pre	9.0 (6-10)	8.9 (6-10)	NS
Mean F/U:	UCLA Activity	Post	8.6 (2-10)	6.65 (3-10)	< 0.0001
HR: 9 yrs	EQ-VAS (QoL)	Post	82% (30-100%)	65.6% (9-97%)	0.009
THA: 10.7 yrs	EQ-5D	Post	0.84 (-0.18-1.00)	0.78 (0.06-1.00)	NS
	Activity Scores				
	Running: No trouble	Post	30% (13/43)	6% (2/36)	0.003
	Very little trouble	Post	19% (8/43)	6% (2/36)	
	Moderate trouble	Post	5% (2/43)	3% (1/36)	
	Extreme trouble	Post	7% (3/43)	11% (4/36)	
	Impossible	Post	5% (2/43)	11% (4/36)	
	Not attempted	Post	35% (15/43)	64% (23/36)	
	Sports: No trouble	Post	40% (17/43)	31% (11/36)	0.004
	Very little trouble	Post	26% (11/43)	8% (3/36)	
	Moderate trouble	Post	12% (5/43)	0%	
	Extreme trouble	Post	0%	8% (3/36)	
	Impossible	Post	5% (2/43)	6% (2/36)	
	Not attempted	Post	19% (8/43)	47% (17/36)	
	Heavy manual labor: No trouble	Post	40% (17/43)	17% (6/36)	0.042
	Very little trouble	Post	26% (11/43)	14% (5/36)	
	Moderate trouble	Post	5% (2/43)	17% (6/36)	
	Extreme trouble	Post	2% (1/43)	8% (3/36)	
	Impossible	Post	2% (1/43)	0%	
	Not attempted	Post	26% (11/43)	44% (16/36)	
	Patient Satisfaction:				
	Delighted	Post	67% (29/43)	58% (21/36)	0.484
	Pleased	Post	19% (8/43)	11% (4/36)	
	Satisfied	Post	0%	19% (7/36)	
	A little disappointed	Post	12% (5/43)	11% (4/36)	
	Very disappointed	Post	2% (1/43)	0%	
Costa (2011)	Harris Hip Score	Pre	52 (31-86)	39 (28-62)	0.001
Mean F/U: 2.5 yrs	Harris Hip Score	Post	97 (77-100)	96 (71-100)	0.237
Killampalli (2009)	Oxford Hip Score	Pre	44.4 (31-57)	46.1 (16-60)	NR
	Oxford Hip Score	Post	16.6 (12-31)	18.8 (12-45)	0.82
Mean F/U: 5 yrs	UCLA Activity	Pre	4.2 (1-8)	3.4 (1-7)	NR
	UCLA Activity	Post	6.7 (3-10)	5.8 (3-10)	0.60

Author (Year)	Outcome		HR	THA	P-value
Stulberg (2010)	Harris Hip Score (mean)	Pre	50.1 (11.6)	49.7 (11.3)	0.233
	"Excellent" category (90-100 pts)	Pre	0% (0/337)	0.4% (1/252)	0.001
	"Good" (80-89)	Pre	0% (0/337)	0% (0/252)	
	"Fair" (70-79)	Pre	0.3% (1/337)	3.6% (9/252)	
	"Poor" (<70)	Pre	99.7% (336/337)	96.0% (242/252)	
	"Excellent" category (90-100 pts)	2 yr	91.3% (240/263)	91.1% (225/247)	0.933
	"Good" (80-89)	2 yr	5.3% (14/263)	4.5% (11/247)	
	"Fair" (70-79)	2 yr	0.8% (2/263)	2.0% (5/247)	
	"Poor" (<70)	2 yr	2.7% (7/263)	2.4% (6/247)	
	"Excellent" category (90-100 pts)	3 yr	87.5% (70/80)	88.8% (166/187)	0.75
	"Good" (80-89)	3 yr	5.0% (4/80)	6.4% (12/187)	
	"Fair" (70-79)	3 yr	6.3% (5/80)	2.1% (4/187)	
	"Poor" (<70)	3 yr	1.3% (1/80)	2.7% (5/187)	

Appendix H. CLINICAL PEER REVIEWERS

Reviewer	Areas of expertise
1. Jason S. Weisstein, MD, MPH Eisenhower Medical Center Rancho Mirage, CA 92270	• Orthopedic Surgeon • Director, Total Joint Replacement Surgery Eisenhower Med Center • Director, Total Joint Replacement Surgery Boca Raton Regional Hospital • AAOS Musculoskeletal Tumors and Diseases Evaluation Subcommittee • Manuscript reviewer *Clinical Orthopaedics and Related Research* (hip/knee reconstruction topics) • Manuscript reviewer, *Orthopedics* (hip/knee reconstruction topics)
2. Howard A. Chansky, MD Seattle WA, University of Washington, Seattle, Washington	• Orthopedic Surgeon • Professor, vice chair and chief of orthopedics and sports medicine, UW • Hip and knee surgery, particularly complex joint replacement, as well as bone and soft tissue tumors. • He practices primarily at UW Medical Center and the VA Puget Sound Health Care System.
3. Creighton Tubb, LTC, MD	• Orthopedic Surgeon • Director, Total Joint Service, Brooke Army Medical Center, San Antonio, Texas

References

1. Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. BMJ 2004;328:1490.
2. Chiou CF, Hay JW, Wallace JF, et al. Development and validation of a grading system for the quality of cost-effectiveness studies. Med Care 2003;41:32-44.
3. 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 2003;9:53-61.
4. 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). In: Agency for Healthcare Research and Quality, Rockville, MD; 2002.