NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Overview

Total hip replacement and resurfacing arthroplasty for treating pain or disability resulting from end stage arthritis of the hip (review of technology appraisal guidance 2 and 44)

This overview is a summary of:

- the evidence and views submitted by the manufacturers, the consultees and their nominated clinical specialists and patient experts and
- the assessment report.

It highlights key issues for discussion at the first Appraisal Committee meeting and should be read with the full supporting documents for this appraisal.

.

Key issues for consideration

Population

- The Assessment Group defined the population for whom both resurfacing arthroplasty and total hip replacement is considered appropriate as people in the National Joint Registry (NJR) who received resurfacing arthroplasty between 2003 and 2012. This population had a mean age of 55.8, and 35% were women. Does the population of people who received resurfacing over the period of NJR follow-up reflect the characteristics of people receiving primary hip replacement in current practice?
- Technology appraisal 44 recommended metal on metal hip resurfacing arthroplasty for younger people who would otherwise receive and are likely to outlive a conventional primary total hip replacement. It further recommended that the activity level of potential recipients be taken into account when deciding whether resurfacing arthroplasty or total hip replacement was the best option for the patient. The Assessment Group stated that the National Joint Registry did not

include data on the activity levels of people in the registry and that as there was no published data on how activity levels may differ between people for whom resurfacing is considered suitable or unsuitable.

    - To what extent is a person's activity level taken into account when deciding whether resurfacing is a treatment option?
    - Can age and gender be used as a proxy for activity levels?
    - Are there additional patient characteristics that are taken into account when deciding whether resurfacing is a treatment option?

- The Assessment Group defined the population for whom resurfacing arthroplasty is not considered appropriate as people in the NJR who received total hip replacement (5 categories) between 2003 and 2012. This population had a mean age of 71.6, and 63.5% were women. Does the population of people who received these categories of total hip replacement over the period of NJR follow up reflect the characteristics of those receiving primary total hip replacement prostheses for whom resurfacing is considered not suitable in current practice?

Classification of prostheses

- The Assessment Group selected the 4 most commonly used combinations of bearing surface and fixation methods since the start of the NJR and an additional category of total hip replacement suggested by its clinical adviser as being clinically relevant
    - Have the most appropriate categories of total hip replacement prostheses been compared?
    - Do the proportions of people receiving each category of total hip replacement prostheses obtained from the up to 9 years of NJR data reflect the proportions of people who received each category in current practice?

Clinical practice

- What are the key determining factors in choosing between different total hip replacement prostheses in clinical practice (cemented, cementless or hybrid, polyethylene on metal, ceramic on ceramic, polyethylene on ceramic)?

    - To what extent would a patient's age and gender influence the decision of which type of total hip replacement prostheses to use in clinical practice?
    - Do other patient characteristics such as activity level and bone quality influence the choice of total hip replacement prosthesis?
- To what extent does the choice of prosthesis depend on factors other than patient characteristics (such as surgeon training and clinical setting)?

Clinical effectiveness

- Is revision rate the key outcome in determining effectiveness of resurfacing arthroplasty and total hip replacement prostheses?
- The Assessment Group noted that the RCTs identified in its systematic review reported outcomes such as osteolysis, aseptic loosening and femoral head penetration. Do these outcomes reflect the likelihood of revision?
- Are there any outcomes which would not be expected to lead to revision which are important to consider separately?
- Limited evidence was presented on adverse events and complication rates across prosthesis type. Would adverse events be expected to be dependent on prosthesis type?

Revision rates

- Extrapolation of revision rates:
    - Is the risk of revision expected to increase or decrease with time after surgery?
    - Would there be any expected difference in how the risk of revision changes over time for any groups of people such as older or younger people?
- Would the revision rates for resurfacing products based on the NJR data be expected to differ if the recalled resurfacing (RS) products were excluded?
- Are revision rates dependent on surgeon experience with a particular type of prosthesis?

- Technology appraisal guidance 2 recommended a bench mark of 10% at 10 years. The Assessment Group found that for the categories of total hip replacement prostheses it assessed, the revision rates were much lower than this benchmark. However, the prostheses for resurfacing arthroplasty did not meet this benchmark. What is an appropriate benchmark of the best prostheses for hip replacement?

Cost effectiveness

- The Assessment Group defined a lifetime horizon as 80 years, but also presented results for a shorter time horizon of 10 years. What is an appropriate time horizon to capture differences in costs and QALYs between resurfacing and total hip replacement and for different categories of hip replacement?
- In the Assessment Group's model people who require a revision enter the revision total hip replacement (THR) health state which is associated with a utility decrement.
    - How long would people be expected to stay in this health state before having revision surgery?
    - Which indications for revision would require immediate surgery?
    - How long are waiting lists for non-emergency revisions?
- The costs of prostheses used in the base case were averages of list prices from a sample of 5 manufacturers for total hip replacement and 3 manufacturers for resurfacing prostheses provided by the National Supply Chain. It is standard for discounts to be offered on dependent on volume purchased or through local negotiations. How can the Committee consider the availability of discounts provided in clinical practice?
- The differences in QALYs between different types of THR are small (around 0.002). Are different types of THR equivalent in terms of QALYs?

1 Background: clinical need and practice

1.1 Arthritis is a general term describing pain and inflammation within a joint and is a leading cause of pain and disability in the UK and worldwide. Arthritis can have many causes of which the most common is osteoarthritis, defined by a loss of cartilage within the joint and accompanying changes in the associated bone. It is estimated that over 2.8 million patients in the UK have osteoarthritis. Osteoarthritis is more common in women than men and its prevalence and incidence increases with age. Estimates of age-standardised incidence rates of hip osteoarthritis amongst women and men in Europe are about 53.2 and 38.1 per 100,000 respectively.

1.2 Rheumatoid arthritis is an autoimmune disease commonly affecting the synovial lining of joints causing inflammation of joints and is the second most common form of arthritis with approximately 400,000 people affected in the UK. Rheumatoid arthritis is approximately three times more common in women than men, and approximately 10-40% of rheumatoid arthritis manifest within the hip. Onset of rheumatoid arthritis can occur at any age but is typically between the ages of 40 and 50 years.

1.3 Symptoms of hip arthritis include pain, stiffness and limited daily activities such as walking, climbing stairs, performing household tasks. The diagnosis of osteoarthritis of the hip is usually based on history and clinical examination with particular assessment of joint pain, deformity, and reduced range of movement. Plain radiographs of the hip are used to identify and stage osteoarthritis.

1.4 NICE Clinical Guideline 59 on the care and management of osteoarthritis in adults states that non-surgical treatments should

be offered as initial treatments for people with osteoarthritis. The clinical guideline also recommends that referral for joint replacement surgery should be considered for people with osteoarthritis who experience pain, stiffness and reduced function that have a substantial impact on their quality of life and are refractory to non-surgical management such as exercise and manual therapy, and pain management.

1.5 People undergoing elective surgery of the hip may receive either a total replacement of the damaged hip (total hip replacement [THR] or a hip resurfacing arthoplasty (RS [although this is not suitable for all people]). NICE technology appraisal 2, published in 2000, provides recommendations on the revision rate benchmarks that total hip replacement prostheses should meet in order to be recommended for routine use in the NHS. The guidance states that the best prostheses should demonstrate a 'benchmark' revision rate of 10% or less at 10 years or, as a minimum, a three year revision rate consistent with this benchmark.

1.6 NICE technology appraisal 44, published in 2002, recommends metal on metal hip (MoM) resurfacing arthroplasty as one option for people with advanced hip disease who would otherwise receive and are likely to outlive a conventional primary total hip replacement. Following research recommendations in NICE technology appraisals 2 and 44, the National Joint Registry (NJR) was set up by the Department of Health and Welsh Assembly Government for the mandatory collection of information on all hip, knee, ankle, elbow and shoulder replacement operations from NHS organisations and to monitor the performance of joint replacement implants.

1.7 In addition to recommending the initiation of a registry, technology appraisal guidance 2 recognised to need for a central organization

to collate information on whether prostheses met the benchmarks. The Orthopaedic Data Evaluation Panel (ODEP), part of the NHS Supply Chain, was established to provide an independent assessment of clinical evidence, submitted by suppliers, on the compliance of their implants for total hip replacement and resurfacing arthroplasty with the NICE benchmarks for safety and effectiveness. For 10 year benchmark products (those with a lower than 10% revision rate at 10 years) ODEP places products in one of 4 categories:

- Level A- strong evidence that meets NICE guidance
- Level B- reasonable evidence
- Level C weak evidence
- Unacceptable evidence

For products that do not have 10 years of data, ODEP looks at the bench marks of 3% at 3 years, 5% at 5 years and 7% at 7 years and again these are split into whether there exists acceptable, weak or unacceptable evidence for the product meeting NICE guidance. As of March 2011, ODEP ratings had been given to 38% of available brands of femoral stems and 41% of available brands of acetabular cups used in primary procedures in England.

1.8 Adverse events associated with hip replacement surgery itself or complications can occur over the longer term from THR and hip RS and failure of THR and RS may occur because of complications at the time of, or after, surgery. The following complications may lead to the hip replacement revision surgery: implant instability, dislocation, aseptic loosening, osteolysis (bone reabsorption), implant failure and infection. In 2010/11 there were 77,800 primary hip procedures (including at independent hospitals) and 9,200 hip revision adverse events and revision. Revision surgery is a key element of the current service expenditure and unit costs of revision are generally higher than for primary surgery.

2 The technologies

2.1 In total hip replacement (THR) surgery, the acetabulum (hip socket of the pelvis) is replaced with a cup with or without a liner and monoblock (single piece) metal stem and head or a two piece (modular) femoral component consisting of a metal stem with a metal, ceramic or ceramicised metal head is inserted into the proximal femur (top of the thigh bone). THR prostheses vary by materials used and the resultant articulation (bearing) surfaces which include: Ceramic on Ceramic (CoC), Ceramic on Polyethylene (CoP); Metal on Polyethylene (MoP); Metal on Metal (MoM) and Ceramicised metal on Polyethylene. THR also vary by the fixation method used for each component. Some THR are fixed into position using cement (hereafter referred to as cemented), other types of THR are designed to be used without cement, initially inserted using press fit fixation and are designed that natural bone growth over time will secure the prosthesis in place (hereafter referred to as cementless). Cementless prostheses rely on initial press fit fixation followed by natural bone growth and may further vary by surface structure or coating of the surface that is in contact with the bone. Some are termed hybrid in which the femoral component is cemented and the cup is cementless, in reverse hybrid THR the femoral component is cementless and the cup is cemented. THR may also vary by the femoral head size; a large head is defined as being greater or equal to 36 mm.

2.2 Hip resurfacing arthroplasty (RS) involves the removal and replacement of the surface of the femoral head with a metal hollow hemisphere, which fits into a metal cup which locates in the acetabulum. All RS prostheses currently on the market are MoM and are cemented. As with THR prostheses, RS prostheses may also vary by the femoral head size. Patient selection for RS depends on various factors including: patient characteristics,

surgeon choice and surgeon experience using a particular class of prosthesis. It is reported to be an option predominantly suited for younger, active men.

2.3 There are over 20 manufacturers of prostheses for hip replacement (THR and RS) and some manufacturers produce multiple brands of components. Data from the National Joint Registry (NJR) has shown that choice of prosthesis has varied over time. For example in 2005, 9% of hip replacements were RS but by 2011 this had decreased to 2%. Since 2004 there has been a trend for a reduction in the proportion of cemented total hip procedures (77% of hip replacements were cemented in 2004, and around 50% in 2010) and an increase in the proportion of cementless procedures (a 4% increase in recent years). For THR carried out during 2010/11, MoP was the most commonly used bearing surface. MoM was the least commonly used.

2.4 All THR and RS must be granted a CE (Conformité Européene) mark before they are used in the UK. The safety assessment required for this is performed by Notified Bodies who inform the Medicines and Healthcare Regulatory Agency (MHRA). The MHRA monitor the safety of devices used in clinical practice. In 2010, the MHRA issued an alert on all MoM hip replacement prostheses following reports of soft tissue reactions, which may be associated with unexplained pain. In June 2012, the MHRA released an updated alert that MoM implants (THR or RS) may wear at an accelerated rate in some people. The MHRA stated that people with MoM prostheses require monitoring for soft tissue damage resulting from reactions to debris from these implants. For symptomatic patients with any type of MoM THR or RS, a blood metal measurement and imaging of the joint is recommended. The MHRA has issued alerts for recall of the following hip replacement prosthesis brands: the R3 metal liner component of the R3

acetabular system (Smith and Nephew alert issued 2012); The ASR system which consists of ASR acetabular cups for hip resurfacing arthroplasty or total hip replacement, ASR surface replacement heads for hip resurfacing arthroplasty, ASR XL femoral heads for total hip replacement (DePuy Synthes alert issued 2010).

2.5 Typically, the price of hip replacement prostheses is dependent on the volume ordered and locally negotiated discounts. For this appraisal the Assessment Group selected the 4 most commonly used combinations of bearing surface and fixation methods in THR prostheses since the start of the NJR and an additional category suggested by its clinical adviser as being clinically relevant (this selection is described in section 4.26). The average list prices from a sample of the manufacturers of commonly used prosthesis combinations and required accessories (such as cement and screws) within these categories and for RS were obtained from the NHS supply chain. The average costs of the 5 categories of THR selected by the Assessment Group and RS were: Cemented polyethylene on metal (£1,557.38), cementless polyethylene on metal (£3,015.60); cementless ceramic on ceramic (£3,886.80); hybrid polyethylene on metal (£2,649.78); cemented polyethylene on ceramic (£1,995.98); RS (£2,672). A full breakdown of how these average costs were derived is given in tables 74, page 230 (RS prosthesis costs), 79 and 80 page 233-234 (total hip prosthesis costs and cement pack costs) of the assessment report.

3 Remit and decision problem(s)

3.1 The remit from the Department of Health for this appraisal was: To appraise the clinical and cost effectiveness of total hip replacement and surface replacement within their CE marked indications for the

treatment of pain or disability resulting from end stage arthritis of the hip.

Population	
Final scope issued by NICE	People with pain or disability resulting from arthritis of the hip for which non-surgical management has failed.
Additional comments or specifications in the Assessment Group's protocol	People with pain or disability resulting from end stage arthritis of the hip for whom non-surgical management has failed
Intervention	
Final scope issued by NICE	1. Primary total hip replacement 2. Primary hip resurfacing arthroplasty
Additional comments or specifications in the Assessment Group's protocol	1.Elective primary total hip replacement 2. Primary hip resurfacing arthroplasty

3.2 The population addressed in the assessment report is restricted to 'end stage arthritis' following discussion with NICE. The scope for this appraisal defines end stage arthritis of the hip as arthritis for which non-surgical management has failed.

3.3 The intervention addressed in the assessment report is restricted to 'elective' primary total hip replacement. This means that people who have received a trauma to their hip requiring hip replacement surgery are not included in this appraisal. This was agreed following discussion with NICE.

	Comparators
Final scope issued by NICE	Different types of primary total hip replacement and hip resurfacing arthroplasty will be compared with each other for people in whom both procedures are suitable. Different types of primary total hip replacement will be compared with each other for people in whom hip resurfacing arthroplasty is not suitable. The different types of hip replacement that will be considered separately are dependent on the available evidence, but may include: •Hip replacements with components made from different materials (metal, ceramic, polyethylene, ceramicised metal) • Cemented, cementless or hybrid prostheses • Prostheses with differing femoral head size • Prostheses with differing revision rates
Additional comments or specifications in the Assessment Group's protocol	Different types of primary total hip replacement and hip resurfacing arthroplasty for people in whom both procedures are suitable Different types of primary total hip replacement compared with each other for people in whom hip resurfacing arthroplasty is not suitable

3.4 The final scope issued by NICE did not define the different types of total hip replacement that should be compared in the economic evaluation. The Assessment Group therefore reviewed the evidence in order to determine relevant categories to be compared. This is described in section 4.26 of this overview.

Outcomes	
Final scope issued by NICE	• Functional result • Pain • Bone conservation • Revision rates • Radiosteriometric analysis to assess prosthesis movement • Dislocation rates • Adverse effects of treatment (peri- and post- procedural) including degradation. products where appropriate • Health related quality of life • Mortality
Additional comments or specifications in the Assessment Group's protocol	• Function • Pain • Bone conservation • Revision rates (device failure/ revision rates/time to revision) • Radiosteriometric analysis (to assess prosthesis movement) • Radiological result • Dislocation rates • Adverse events include peri- and post-procedural complications (e.g. infection, nerve palsy, dislocation rates, femoral neck fracture, • Metallosis, muscle weakness) and metal and other degradation • Products • Health related quality of life • Mortality
Economic evaluation	
Final scope issued by NICE	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective.
Additional comments or specifications in the Assessment Group protocol	Cost-effectiveness outcomes include mean differences in costs and clinical effectiveness measures or utility measures; incremental cost-effectiveness ratio (ICER), uncertainty measures, ceiling willingness-to-pay ratios and probabilities from cost effectiveness acceptability curves (CEACs).

4 Clinical-effectiveness evidence

4.1 The Assessment Group conducted a systematic review of randomised controlled trials (RCTs) and systematic reviews; a systematic review of registry studies of hip replacement procedures and a retrospective cohort analysis of individual patient data from the National Joint Registry (NJR). The Assessment Group considered the retrospective cohort analysis of the NJR the most appropriate and applicable approach to define the populations in the final scope issued by NICE (that is people for whom resurfacing arthroplasty is or is not considered suitable), and to identify relevant categories of THR prostheses to be compared. The Assessment Group stated that it allowed a comprehensive analysis of revision rates used in England and Wales. The Assessment Group commented that the data from the RCTs and systematic reviews had limitations in methodological quality, inconsistency and incomplete coverage of all outcomes and prosthesis types.

Clinical effectiveness evidence from RCTs and systematic reviews

4.2 Literature on Total Hip Replacement (THR) or hip resurfacing arthroplasty (RS) was searched from 2002 onwards (date of publication of NICE technology appraisal 44). Studies were excluded that assessed indications for hip replacement other than end stage arthritis of the hip or had assessed revision surgery rather than primary surgery. Owing to the volume of studies identified, the Assessment Group performed a further round of filtering in which it excluded studies that were published before 2008 (unless it was a companion paper to an included study) or included less than 100 people.

4.3 The Assessment Group identified 16 randomised controlled trials (RCTs) and 8 systematic reviews. It noted that there were a further 20 on-going clinical trials. Thirteen of the 16 RCTs and 5 of the

8 systematic reviews were studies that compared different types of THR with each other; 3 RCTs and 3 systematic reviews compared RS with THR. The Assessment Group assessed the risk of bias and methodological quality of each included study. For comparison it established a decision rule to determine whether evidence could be considered conclusive or non- conclusive based on the precision, consistency and clinical relevance of the effect size of the data. For further details of the decision rule, see pages 73-74 of the assessment report.

4.4 Table 1 provides a summary of the RCTs identified and critiqued by the Assessment Group. Studies in bold text show the publications from which the results were extracted, studies in non-bold text are companion publications from the same study. Numbers in brackets are the reference numbers from the assessment report.

Table 1: Summary of RCTs identified and critiqued by the Assessment Group

Comparison	Maximum follow up across studies and outcomes (years)	Reference
THR vs. RS	6	1)**Costa** 2012(127) (Achten 2010 (104) 2) **Garbuz** 2010 (128) 3) **Vendittoli** 2010(129) (Vendittoli 2006 (130); Girard 2006 (131); Rama 2009 (132); Vendittoli (133)
Cup fixation 1+2) Cemented vs. cementless	10	1) **Bjorgul** 2010 (107) (Bjorgul 2010) 2) **Angadi** 2012 (109)
Cup liner bearing surface 1+2) Cross linked polyethylene vs. non-cross linked polyethylene	10	1) **McCalden** 2009 (238) 2) **Engh** 2012 (110) Engh 2006
Cup shell design 1) porous coated vs. arc-deposited hydroxyapatite-coated	10	1) **Capello** 2008 (112) (D'Antonio 2005, D'Antonio 2003 Mesko 2011)
Cup/stem fixation 1) Cemented vs. cementless	7	1) **Corten** 2011 (116) (Laupacis 2002, Bourne 2010 Corten 2011)
Femoral head size 36mm vs.28mm	1	1) **Howie** 2012 (120)
Femoral head bearing Oxinium vs. cobalt chronium	2	1) **Lewis** 2008 (121)
Femoral head on cup liner bearing surface 1+2) ceramic-on-ceramic vs. metal-on-polyethylene vs. ceramic on polyethylene 3) cobalt chromium/ oxinium-on-polyethylene vs. cobalt chromium/oxonium-on-cross linked polyethylene	2-5	1) **Amanatullah** 2011 (122) 2) **Capello** 2008 (112) (D'Antonio 2005, D'Antonio 2003 Mesko 2011) 3) **Kadar** 2011(123)
Stem composition Cobalt chromium vs. titanium	5	1) **Healy** 2009 (124)
Stem design Short metaphyseal fitting vs. conventional diaphyseal filling	3	1) **Kim** 2011 (125)
Stem fixation Cemented vs. cementless	20	1) **Kim** 2011 (126)

4.5 Three RCTs compared the effectiveness of THR with RS. One RCT compared metal on metal (MoM) RS with large head MoM THR; 1 study compared MoM RS with MoM THR and the final study did not specify the bearing surface of the THR prostheses being compared with the MoM RS. Two of the RCTs were undertaken in Canada

and the other in the UK. A total of 422 patients were randomised to treatment across the 3 RCTs (ranging from 104 to 192). Length of follow-up of the trials ranged from 1 to 6 years. The mean age of the patients recruited to the trials ranged from 50 to 56 years. The proportion of women recruited across the trials ranged from 10.5% to 41%. The proportion of people who had a diagnosis of primary osteoarthritis varied between 33% and 95% across the 3 RCTs.

4.6 Three systematic reviews compared the effectiveness of THR with RS with respect to function, risk of revision, complications and mortality. Evidence in the systematic reviews was synthesized from both RCTs and non-RCTs. Two systematic reviews assessed all types of THR compared with RS, 1 systematic review compared cementless THR prostheses with RS prostheses. Two systematic reviews specified that the comparison of RS prostheses with THR prostheses was in younger patients (one had an eligibility criteria of under 65 years, the other stated that the average age of the study populations was 55 years). Two of the systematic reviews included RCTs that the Assessment Group had critiqued separately, 1 included 2 RCTs and the other 1 RCT.

4.7 Out of the 3 RCTs comparing THR with RS the only outcome where there was a conclusive result showing a difference was infection rates. The Assessment Group's meta-analysis of 2 RCTs that had assessed this outcome indicated that at 12 to 56 months after surgery, patients who received THR were at increased risk of infection compared with those who had received RS (pooled odds ratio [OR]=7.94, 95% CI 1.78 to 35.40). There was conclusive evidence showing no difference between THR and RS with 3 measures of function (Oxford Hip Score, Western Ontario McMaster Osteoarthritis Index [WOMAC] score, McMaster University osteoarthritis). For the remaining function measures reported across these RCTs (University of California Los Angeles

[UCLA], Harris Hip Score) the results were inconclusive. The available results for Quality of life (SF-36, EQ-5D), revision and complications (deep vein thrombosis, implant dislocation, superficial wound complications, aseptic loosening) were inconclusive. No data was presented on mortality or pain (although the quality of life and function measures include a pain component) or failure rates. For further details of the results of the trials see pages 114- 118 of the assessment report. For further details of function measures see page 30 of the assessment report.

4.8 The systematic reviews showed conclusive evidence of a lower revision rate with THR than RS (2 systematic reviews reported this outcome; 1 of these systematic reviews provided a meta-analysis of 4 RCTs which compared risk of revision for patients receiving THR compared with RS reported a pooled relative risk (RR) estimate for revision of 2.60 (95% CI 1.31 to 5.15) over a 1 to 10 year follow up). Two systematic reviews conclusively found a higher relative risk of component loosening with RS prostheses compared with THR prostheses (RR 3.00, 95% CI 1.11 to 8.50 and RR 4.96, 95% CI 1.82 to 13.50). One systematic review found numerically lower, but not statistically significant dislocation rate with RS compared with THR (RR 0.25, 95% CI 0.05 to 1.21); the other systematic review found that dislocation was statistically significantly lower with RS than THR (RR 0.20, 95% CI0.10 to 0.05,) The Assessment Group considered the evidence on function reported in the 3 systematic results to be inconclusive as there was a lack of pooled mean difference estimates and inconsistent results for the function scores. None of the comparisons of function, mortality, failure rates and infection were considered conclusive. None of the systematic reviews reported health related quality of life. For further details of the results from the systematic reviews see pages 119 – 122 of the assessment report.

4.9 Of the 13 trials identified by the Assessment Group comparing different types of THR, only one was conducted in the UK. The RCTs compared the composition, design, bearing surface, fixation method or component size (see table 1 for a summary of the comparison in each trial) of the THRs. The number of people in each RCT ranged from 100 to 557. The length of follow-up ranged from 3 months to 20 years. The mean age of the patients across the RCTs ranged from 45 to 72 years. Eight of the RCTs included people with a mean age greater than 60 years and the remaining 5 trials included people whose mean age was below 60 years. The proportion of women recruited across the trials ranged from 24% to 73%. Reported outcomes across the RCTs varied. The outcomes reported were: functional scores, risk of revision, femoral head penetration rate (a measure of prosthesis movement), implant dislocation rate, osteolysis (a long term complication which results in bone loss, aseptic loosening (which follows osteolysis), femoral fracture, infection, health related quality of life and mortality. None of the RCTs reported pain scores or bone conservation which were listed as outcomes in the final scope issued by NICE although it should note that pain may be captured in the function or quality of life measures.

Function

4.10 Twelve of 13 trials comparing different types of THR reported function as an outcome measure. The function scores used varied and were measured at different post-procedure follow-ups. For all function outcomes across the RCTs, the results either showed no difference or were inconclusive. For further details of the results from the trials, see pages 82-90 of the assessment report.

Revision

4.11 Data on revision rates was reported in 10 RCTs. One RCT which compared different cup liner bearing surfaces (cross- linked

polyethylene compared with non cross-linked polyethylene cup liners) suggested a statistically significantly reduced risk of revision in patients who received cross-linked polyethylene cup liners relative to non-cross linked polyethylene cup liners (RR 0.18, 95% CI 0.04 to 0.78). The results from the remaining RCTs that assessed this outcome suggested that the difference in risk of revision between the different types of THR were inconclusive. For further details of the results from the trials, see pages 92 to 93 of the assessment report.

Prosthesis movement: femoral head penetration rates

4.12 Three RCTs reported femoral head penetration rates. The first compared the yearly femoral head penetration rates of 5 bearing surfaces of head material on cup liner. The results indicated that over 2 years, the bearings using cross linked polyethylene and the steel on non-cross linked polyethylene combination had a statistically significantly lower femoral head penetration rates than the cobalt chromium-on-polyethylene combination or the oxinium-on-polyethylene combination. The other 2 RCTs compared a non cross-linked polyethylene cup liner with a cross linked polyethylene liner at 5 and 10 years. Cross linked polyethylene was associated with a statistically significantly lower femoral head penetration rate than non cross linked polyethylene ($p<0.001$). For further details of the results, see table 25, on page 97 of the assessment report.

Implant dislocation

4.13 Seven RCTs reported data on the relative risk of implant dislocation between THR types. Four of the studies had inconclusive results, Two RCTs with 10 years follow up found a decreased risk of dislocation with a cemented cup compared with a cementless cup. The pooled odds ratio for these 2 studies was 0.34 (95% CI 0.13 to 0.89). An RCT that compared dislocation rates in prostheses with a large femoral head (36 mm) compared with a small femoral head

(26 mm) showed that over 1 year, fewer dislocations occurred with the larger femoral head size than for the smaller femoral head size (RR 0.17; 95% CI 0.04 to 0.78). For further details of the results, see table 26 on page 98 of the assessment report.

Osteolysis and aseptic loosening

4.14 Seven RCTs reported data on osteolysis (bone reabsorption). Only one RCT presented a conclusive result. Over 10 years fewer people who had received a THR comprising of a ceramic head on a ceramic cup liner bearing had osteolysis (3/222) than people who had received a THR comprising of a metal head on a polyethylene liner (15/106, [RR 0.10, {95% CI 0.02 to 0.32]). None of the 5 RCTs that had assessed aseptic loosening were considered conclusive by the Assessment Group. For further details of the results, see tables 28 and 30, on pages 100 and 102 of the assessment report.

Other complications

4.15 Seven RCTs reported data for the following complications: femoral fractures, infection and deep vein thrombosis. The Assessment Group considered all of the results to be inconclusive. For further details of the results, see tables 31(femoral fracture), 32 (infection) and 33 (deep vein thrombosis) on pages 102 to 103 of the assessment report.

Health related Quality of Life

4.16 Only 3 RCTs reported any comparative data on measures of health related quality of life for different types of THR. All 3 RCTs had used the short form health survey 12 (SF-12) questionnaire. One RCT which compared cup liner bearing surfaces (cross linked compared with non-cross linked polyethylene cup liner bearing surfaces) reported that there was no difference in health related quality of life (on mental and physical subscales of SF-12) between the two groups of patients at a follow-up of 1-5 years. In the other

2 RCTs (one comparing different head surfaces [oxinium compared with cobalt chromium] and the other comparing different femoral head-on-cup liner articulations [ceramic-on-ceramic compared with ceramic -on- polyethylene]) reported that there was no statistically significant difference in health related quality of life (on mental and physical subscales of the SF-12) between the different groups of patients. The Assessment Group commented that it considered the evidence from these 2 RCTs to be inconclusive. For further details of the results, see table 20, on page 91 of the assessment report.

Mortality

4.17 Data on mortality were reported in 6 RCTs with the longest follow up of 10 years. The Assessment Group considered the data from all of the RCTs to be inconclusive. For further details of the results, see table 24, on page 95 of the assessment report.

4.18 The primary focus of the 5 systematic reviews evaluating the clinical effectiveness of different types of THR was the comparison of different cup fixation methods (cemented compared with cementless), and materials used for implant articulation on post-operative clinical function scores and revision rates. The Assessment Group considered the majority of evidence from these systematic reviews to be inconclusive as they had either reported only a narrative syntheses; had used inappropriate pooling methods or had reported inconsistent summary findings. The only conclusive result identified by the Assessment Group was that there was no difference in the risk of revision between 2 different articulations (zirconia [a type of ceramic] head-on-polyethylene compared with a non-zirconia on polyethylene) (Risk difference 0.02, 95% CI -0.01 to 0.06).

Clinical effectiveness evidence from registry studies

4.19 The Assessment Group performed a search for registry studies of THR or RS for people with end stage arthritis of the hip. It excluded studies with fewer than 1000 people in the registry study at the time of publication and in which hip and knee replacement data was not reported separately. It also excluded studies that were not the most recent report in a publication series or were not the most recent annual report. It identified 8 registry studies that had reported on RS prostheses and 30 studies that had reported on THR.

4.20 Three of the 8 studies that reported on RS and THR used data from England and Wales, 1 used data from the Finnish Joint Registry, 1 from the Australian Joint Registry, 1 combined data from Australia, England and Wales and Sweden (multinational), 1 reported data from the Nordic Arthroplasty Registry (data from Norway, Denmark and Sweden) and 1 compiled data from the UK, Australia, Asia and USA. The registry studies reported data for different outcomes at different follow up times. The comparisons reported also differed.

4.21 The majority of studies that compared revision rates found that RS had a higher revision rate than THR. Analysis of registry data from various countries suggested the risk of revision with RS varied by country and one study demonstrated revision rates were lower if undertaken in specialist centres. The registry studies also demonstrated that the revision rates for RS were higher for women than for men: one analysis of the NJR data for England and Wales demonstrated that women had a 30% greater risk of revision with RS than men (hazard ratio 1.30; 99% CI 1.01 to 1.76). A Nordic registry study also found a greater relative risk of revision for RS prostheses than THR with a greater discrepancy in women (for women [younger than 50 years] RS compared with THR, RR 4.7; [95% CI 2.6 to 8.5]; and for men [younger than 50 years] RS compared with THR, RR 1.9, [95% CI 1.0 to 3.9]). A further

analysis of the NJR suggested that although RS always had higher revision rates than THR in women, in men RS prostheses with a larger head size (54 mm) had similar predicted 5 year revision rates to THR prostheses, but the proportion of men who received large head RS was only 23% of all male RS recipients.

4.22 The Assessment Group identified 22 registry studies that reported on THR (without separate analysis of RS) which included data from many countries (NJR [England and Wales]; Denmark; Sweden; Australia; Italy (regional), Finland; Slovakia; Norway; USA; Multinational 1(including data from 9 registries: England and Wales, Italy [regional], Australia, Denmark, New Zealand, Portugal, Norway, Slovakia, Sweden) Multinational 2 (including data from Australia, England and Wales, New Zealand) and the Nordic registry(Denmark, Sweden, Norway).

4.23 Two registry studies assessed whether there was an association between femoral head size and outcome for THR and demonstrated that the relationship between femoral head size and revision rate was dependent on bearing surface. One study of NJR data from England and Wales (Smith et al 2012) showed that the failure of MoM THR increased as the femoral head size increased. In men aged 60 years, the cumulative incidence of revision at 5 years was 3.2% (95% CI 2.5 to 4.1) with a 28 mm femoral head component whereas with a large 56 mm femoral head component the cumulative incidence of revision was 5.1%, (95% CI 4.2 to 6.2). The 5 year cumulative revision rates in women were 6.1% with a 46 mm MoM THR. A study that combined data from England and Wales, Australia and New Zealand registries also reported an increased revision rate associated with larger femoral head sizes when MoM was the bearing surface. Conversely, for bearing surfaces other than MoM a large femoral head size was associated with a lower risk of revision compared with smaller femoral heads.

For example Smith et al reported that that in men aged 60 years, the 5 year revision rate with a 28 mm ceramic on ceramic THR was 3.3% (95% CI 2.6 to 4.1), whereas the equivalent revision rate for men who had received larger 40 mm ceramic on ceramic THR prostheses was 2.0% (95% CI 1.5 to 2.7).

4.24 Analysis of the NJR (McMinn et al 2012) showed, at 8 years follow up a higher mortality rate for patients undergoing cemented compared with cementless THR (adjusted hazard ratio 1.11, 95% CI 1.07 to 1.16).

4.25 The Assessment Group presented data on revision rates up to 19 years from the Swedish Registry for all hip replacements grouped together but noted that these revision rates may be affected by devices and practices that are no longer used. This suggested that there were differing revision rates based on age. For people younger than 50 years of age at 19 years follow up 60.2% of women and 62.6% of men had not had a revision, for people aged between 50 and 59 years, 73.7% of women and 67.2% of men had not had a revision, for people aged between 60 and 75 years 87.2% of women and 80.5% of men had not had a revision and for people over 75 years, 94.8% of women and 92.1% of men had not had a revision (see figure 72 on page 301 of the assessment report).

Assessment Group analysis of individual patient data from National Joint Registry

Assessment Group's determination of categories of total hip replacement

4.26 The final scope issued by NICE stated that the different types of THR to be considered separately are dependent on the available evidence. The Assessment Group obtained individual patient data

from the NJR that included data from 2003 to September 2012. In total 31,222 RS operations had been performed and 387694 THR operations (of which 387,667 had complete records with usable data). The Assessment Group assessed the THR types listed in its NJR dataset to identify the most frequently used categories of THR prostheses which it supplemented with advice from its clinical specialists on other relevant categories. The Assessment Group stated that the database categorised the key components that make up a THR prosthesis: cup component group; cup component type; cup composition; cup fixation; cup implant type; head component type; head composition; liner component type; liner composition; stem component type; stem fixation and stem implant type. The Assessment Group identified 7 mutually exclusive categories of THR. Of these, it selected the 4 most frequently used combinations and a further combination of a cemented stem with a ceramic head articulating with a cemented polyethylene cup which its clinical specialist stated is often used in younger high demand (more active) patients because of its low wear characteristics. The 5 categories of THR prosthesis accounted for 239,089 (62%) of 387,667 THR in the NJR with usable records. The categories were (numbers in brackets are the number of procedures recorded in the NJR with each combination):

- Category A: Cemented polyethylene cup on a metal head (cemented stem) CePoM (125,285),
- Category B: Cementless HA (hydroxyapatite) coated metal cup (with a polyethylene liner) on a metal head CeLPoM (37,874),
- Category C: Cementless HA coated metal cup (polyethylene liner) on ceramic head CeLCoC (34,754),
- Category D: Cementless HA coated metal cup (polyethylene liner) on metal head (cemented stem) HyPoM (28,471),

- Category E: Cemented polyethylene cup on ceramic head (cemented stem) CePoC (12,705).

4.27 The Assessment Group analysed the characteristics of people in the NJR data set including the ratio of men to women, the average age and the age range of people who had received RS, all types of THR prostheses, and each of the 5 categories of THR in the NJR. Table 2 summarises the patient characteristics in the Assessment Group's NJR dataset and categories. The age distributions are represented graphically in figures 1 and 2 showing the degree of overlap between the populations.

.

Table 2: Age and gender of RS and THR recipients (modified from tables 56 and 57, on pages 185 and 188 of the assessment report)

Population	Number	% women	Mean age (SD)
All RS recipients	31,222	29.9	55.0 (8.6)
All THR recipients	386,556	61.4	69.5 (10.3)
All THR female recipients	237,436	100	70.2 (10.3)
All THR male recipients	149,120	0	68.45 (10.3)
THR Categories A to Erecipients	239,089	63.5	71.6 (9.6)
All CAT A recipients	125,285	66.9	74.6 (7.9)
All CAT B recipients	37,874	60.2	71.5 (8.7)
All CAT C recipients	34,754	55.4	61.6 (9.9)
All CAT D recipients	28,471	64.2	73.0 (8.3)
All CAT E recipients	12,705	60.1	66.2 (9.6)
RS propensity matched pop*	26,643	35.0	55.83 (8.3)
THR propensity matched pop*	26,643	35.0	55.83 (8.3)

Figure 1 Kernel density diagram of the distribution of recipients of hip resurfacing arthroplasty and total hip replacement (figure 24, page 185 of the assessment report)

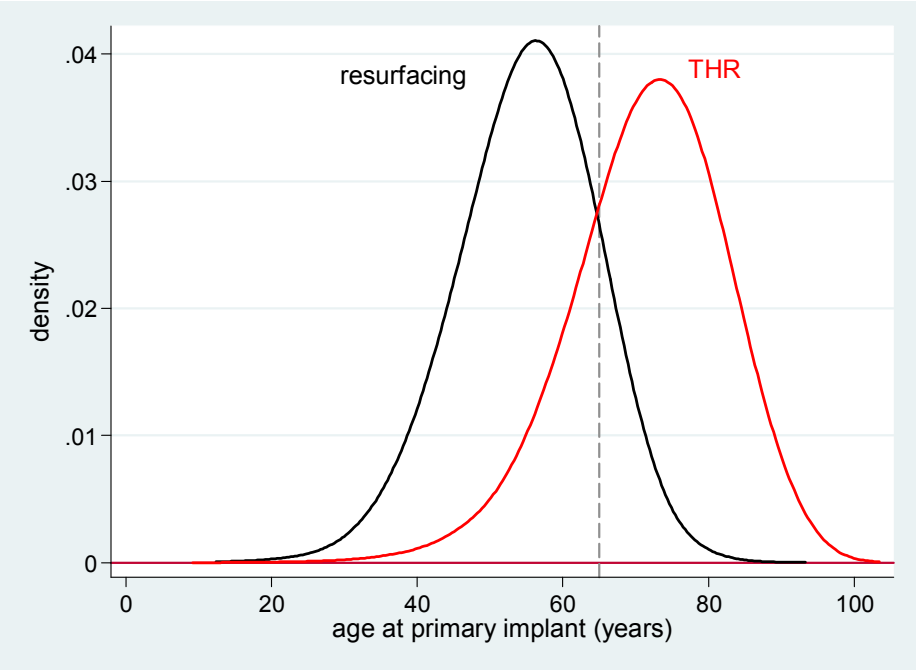
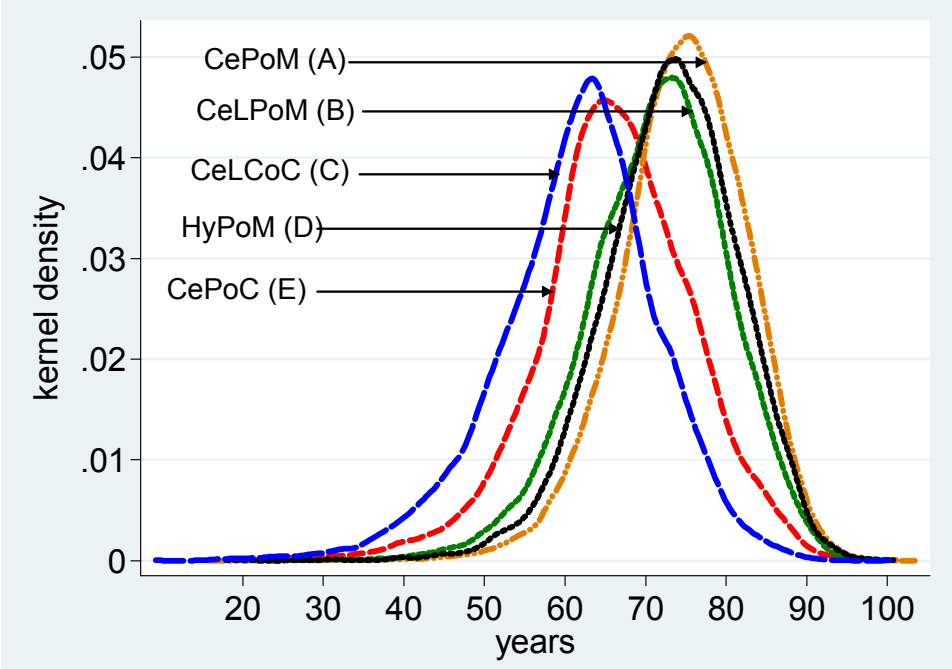


Figure 2 Kernel density plots of age at primary implant for category A to E, THR prostheses (figure 46, page 206 of the assessment report)



Assessment Groups determination of populations for whom hip resurfacing arthroplasty is considered suitable or unsuitable

4.28 The Assessment Group assessed how to determine for which population RS is suitable. It noted that NICE technology appraisal

guidance 44, recommended resurfacing arthroplasty for people with advanced hip disease who would otherwise receive and are likely to outlive a conventional primary total hip replacement. The Assessment Group stated that the guidance noted that the evidence (available at the time for TA 44) was principally for people younger than 65 years. The Assessment Group stated that clinical opinion is that RS is offered as a treatment option mainly to relatively active younger people, while THR is the predominant option for less active older people. The Assessment Group noted that the NJR did not categorise people according to activity levels. In the absence of data on activity levels the Assessment Group defined suitability for RS based on the characteristics of the people in the NJR who had received RS and to which it propensity matched (based on age and gender) a population who had received one of the 5 categories of THR it had selected (see section 4.26). Further details of the methods of propensity matching are given on page 176 of the assessment report. For the comparison of different types of THR for people in whom RS was not suitable, the Assessment Group noted that although it can be assumed that THR would also be suitable for those people for whom RS was suitable the reverse is less likely (that is, RS may not be suitable for all people for whom THR is suitable). The Assessment Group also noted that most people who received THR in the NJR were over 65 years which is consistent with clinical opinion that older people are more likely to receive THR than RS. However, the Assessment Group suggested that as higher revision rates have been observed with RS than with THR fewer younger people may be considered to be candidates for both procedures in the future. As a result of this the Assessment Group compared the 5 THR categories across the whole population that received them in the NJR (irrespective of age and gender) in its base case. The effect of age and gender was explored in subgroup analyses (see

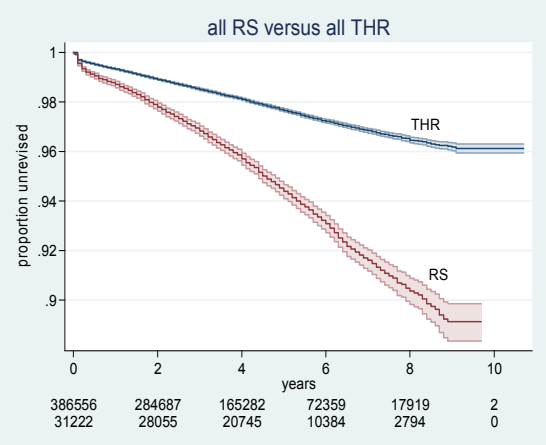
sections 4.38 and 4.39 of overview for revision rates in subgroups, 6.19 and 6.20 for subgroup cost effectiveness results).

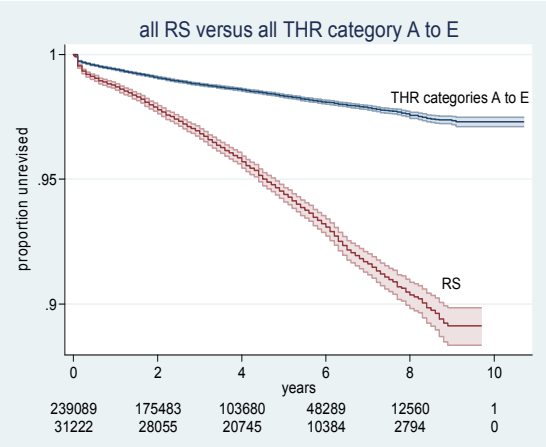
Assessment Group analysis of revision rates of prostheses in the NJR

4.29 The Assessment Group analysed revision rates using the available data from the NJR (maximum follow up of 9 years). In order to remain consistent with all previous economic analyses of hip replacement technologies, it used Kaplan Meier estimates rather than Competing Risks estimates of revision rates. As the registry data included people with different lengths of follow up, the Assessment Group measured time to revision for each person in the registry over the length of time of follow up each person had. People were censored if they died during their follow up period as they were no longer at risk of revision. The Assessment Group presented plots of the proportion of people who had not received a revision up to 10 years.

4.30 The Assessment Group found, consistent with previous published analysis of the NJR (see section 4.21 of overview), that the revision rate for RS over 9 years of follow up was about 3 times higher than for THR (all the types of THR prostheses recorded in the NJR). The difference was even larger when the comparison with RS was restricted to a combination of the 5 commonly used THR (prosthesis categories A to E). The Kaplan Meier plots of time to revision are shown in figure 3 below.

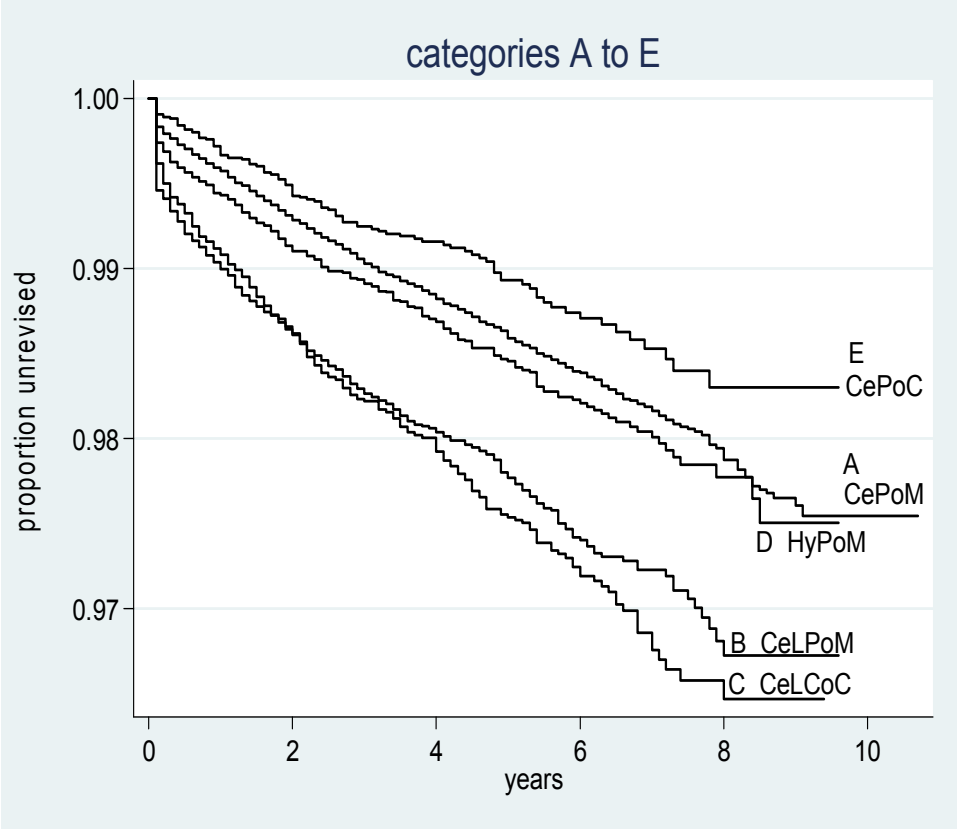
Figure 3: KM plots of time to revision A) all RS vs. all THR B) all RS versus all THR category A to E, (Figures 30 and 32, pages 193 and 194 assessment report)





4.31 The Assessment Group assessed the time to revision for the 5 categories of THR (A to E) separately using the observed data from all people who had received these prostheses in the NJR with no adjustments for age or gender. The Assessment Group noted that the revision rates for the cementless prostheses (category C [cementless ceramic on ceramic] and category B [cementless polyethylene on metal]) were higher than the cemented prostheses (Category E [cemented polyethylene on ceramic]; and Category A [Cemented polyethylene on metal]). Observed time to revision for the 5 THR prosthesis categories are shown in figure 4 below.

Figure 4 observed time to revision for 5 THR prosthesis categories (figure 41, page 201 assessment report)



4.32 The Assessment Group presented data from the NJR for men and women separately (see figure 5 below). Revision rates for RS were higher for women (18% at 9 years) than for men (7% at 9 years). The Assessment Group also presented data for each THR category separately for men and women. The Assessment Group noted that revision was more frequent for men than women when the 5 categories of prosthesis were combined, but that this was least pronounced for the cemented polyethylene on ceramic prosthesis (category E) (see figure 6).

Figure 5 time to revision all RS and all THR patients (categories A to E) by gender (figure 33, page 194 of assessment report) Blue = THR, red= RS

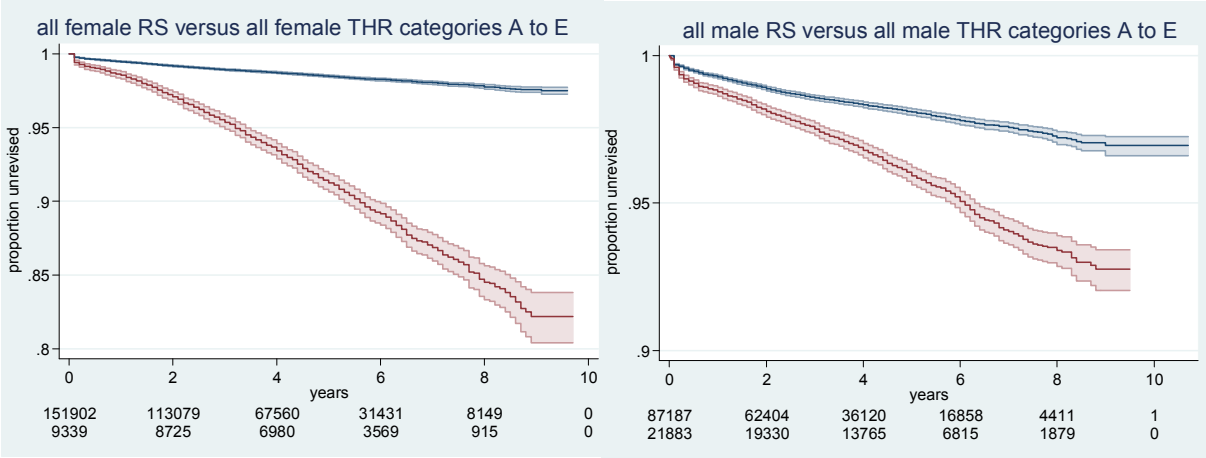
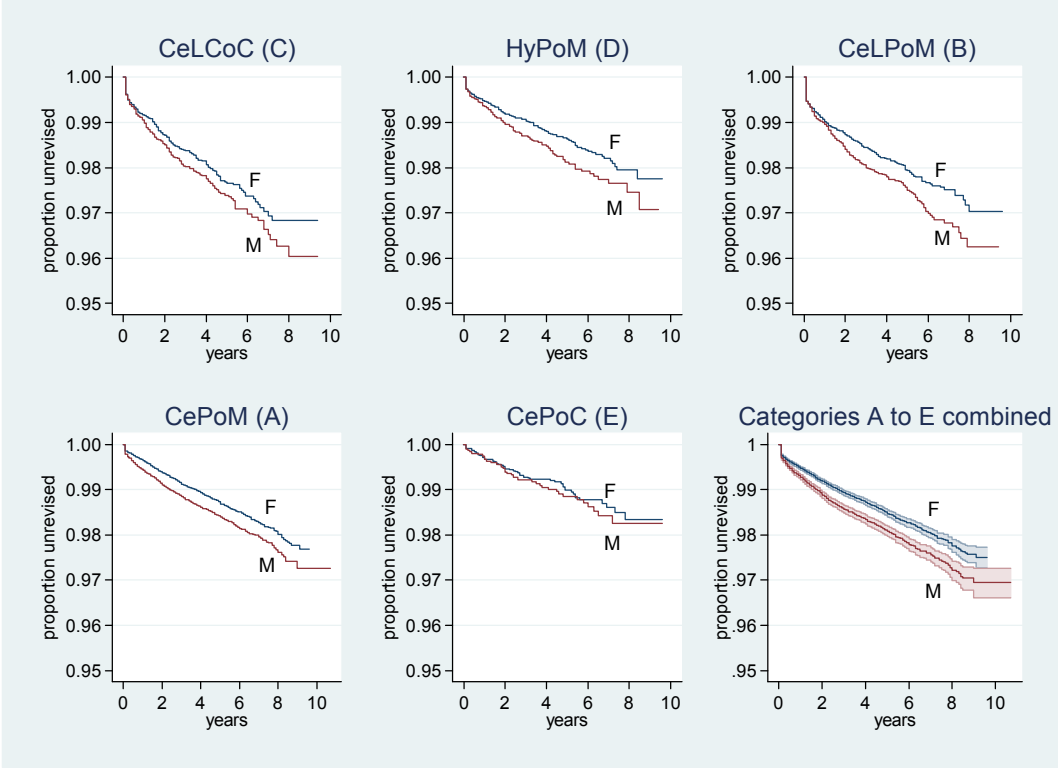


Figure 6 time to revision for each THR category by gender (figure 45 page 206 of assessment report) Blue= women, red = men



4.33 The Assessment Group considered the factors that would determine whether a person had revision surgery and suggested that this included both failure of the prosthesis and a person's suitability for revision surgery. It concluded that for younger people the hazard of revision was expected to increase over time as the risk of outliving the prosthesis increases and that there would be an increasing likelihood of revision because of wear to the prosthesis

particularly if the person was active. The Assessment Group concluded that for older people the hazard of revision may decrease over time because the risks of surgery may outweigh the benefit of revision surgery given the shorter life expectancy at the point of failure of the hip prosthesis. In the economic base case, the Assessment Group applied a bath tub model (U shaped model which allows for an initial decreasing hazard of revision with increasing hazard longer term), to extrapolate from the NJR data over the long term as it was the best fit to the NJR data and as it considered an increasing hazard of revision in the extrapolated period plausible. In the sensitivity analysis, the Assessment Group used the Log normal model (which gave a decreasing hazard of revision with time) as this was the second best fitting distribution and allowed for a decreasing hazard in the extrapolated period.

4.34 For the population for whom RS was suitable, the Assessment Group presented bath tub modeled percentage revision at 10, 20 and 30 years for the RS population and the THR combined categories A to E propensity matched to the RS population based on age and gender. The results are presented in table 3 below.

Table 3: modelled long term revision rates for population for whom hip resurfacing arthroplasty is suitable (modified from Table 60, page 199 of assessment report)

RS vs. THR (RS suitable)			
	RS	THR 5 categories propensity matched to RS	
Mean age Median age IQ range	56 54 49-59	56 54 49-59	
% female	35	35	
Intervention	10 years	20 years	30 years
RS	17.2%	48.3%	76.3%
THR 5 categories Propensity matched to RS	4.6%	12.9%	24.6%

4.35 For the population for whom RS was not suitable, the Assessment Group presented bath tub modeled percentage revision at 10, 20

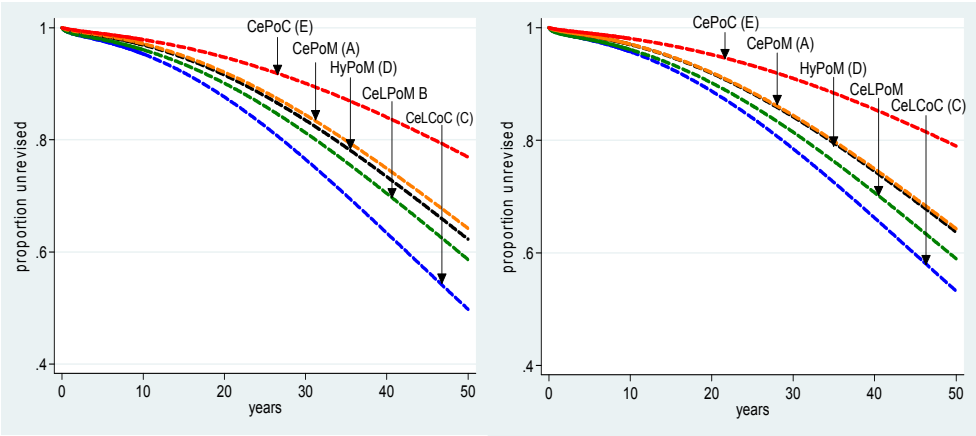
and 30 years for each of the 5 THR prosthesis categories separately. The results are presented in table 4 below.

Table 4 modelled revision rates for population for whom RS is unsuitable (modified from table 62, page 205 assessment report)

THR vs THR (RS unsuitable)					
	THR A CePoM	THR B CeLPoM	THR C CeLCoC	THR D HyPoM	THR E CePoC
Mean age Median age (IQ range)	74.6 74.9 70-80	71.5 72 66-78	61.6 62.3 56-68	73.0 73.4 68-79	66.2 66.3 61-73
% female	66.9	60.2	55.4	64.2	60.1
10 years	2.8 %	3.9 %	4.6%	3.0%	2.1%
20 years	7.9	9.9%	12.3%	8.4%	5.2%
30 years	15.6%	18.7%	23.5%	16.5%	9.9%

4.36 The Assessment Group noted differences in the age distributions and proportions of men and women across the THR categories A-E. As age may, and gender has been shown to, affect revision rates, the Assessment Group repeated its modeling of revision rates for each category of THR in which it controlled for age and gender (that is age and gender were included as covariates). The age and gender coefficients for each of the 5 THR prosthesis categories are presented in table 86, page 240 of the assessment report. The Assessment Group noted that the ranking of the 5 THR prosthesis categories (in terms of proportions of people who had a revision at various time points) was the same regardless of whether or not adjustment for age and gender had been undertaken, however the relative difference in revision rates of the cemented polyethylene on ceramic category (category E) compared with the other THR prosthesis categories was increased following adjustment.

Figure 7 extrapolation of bath tub models of revision for THR categories A to E, left panel is uncontrolled results, right panel is controlled for age and gender (Figure 44 page 204 assessment report)



Assessment Group modelling of longer term revision rates for subgroups

4.37 For the population for whom RS is considered suitable, the Assessment Group explored what the expected longer term revision rates with THR and RS would be in women and men separately. The Assessment Group propensity matched (using age) men who had received the 5 THR categories A to E combined with men who had received RS and the equivalent for women. As expected the predicted revision rates were higher for women who received RS at 10, 20 and 30 year follow up than men. The modeled longer term revision rates for women and men who received THR (all categories combined) were similar across genders and lower than the rates for people who had received RS (see table 5 below).

Table 5 Predicted proportions of people requiring revision in the population for whom RS is suitable (modified from tables 65 and 66, on pages 210 and 212 assessment report)

Intervention	10 years	20 years	30 years
Female (RS suitable) mean age 53.5 years			
RS	23.1%	61.2%	87.6%
THR	4.8%	13.2%	25.2%
Male (RS suitable) mean age 57.1 years			
RS	12.4%	35.6%	61.2%
THR	4.7%	13.2%	25.5%

4.38 The Assessment Group noted that its analysis of the NJR data had demonstrated that revision was more frequent for men than women for all of the 5 categories of THR, although this was least pronounced for the cemented polyethylene on ceramic prosthesis (category E, see section 4.32). The Assessment Group therefore performed a subgroup analysis of revision rates in men and women separately in the population for whom RS was unsuitable. Furthermore the Assessment Group explored revision rates in people over 65 years (the majority of people who receive a THR) and people under 65 years separately. The subgroups considered were:

- Women aged under 65 years
- Men aged under 65 years
- Women aged over 65 years
- Men aged over 65 years

4.39 In order to undertake the 4 subgroup analyses for the population for whom RS is not suitable, the Assessment Group stratified data from the NJR according to age (under 65 years and over 65 years) and by gender. For the 2 subgroups aged over 65 years, the Assessment Group further assumed that the risk of revision would decrease with time and therefore used the log normal distribution to extrapolate from the NJR data. For the 2 subgroups aged less than 65 years, the Assessment Group assumed that the risk of revision would increase in the extrapolated period and therefore used the

bath tub distribution to extrapolate from the NJR data. The estimated revision rates at 10, 20 and 30 years were presented for each category. See tables 6 and 7 below.

Table 6 predicted revision rates for men and women aged over 65 years (tables 67 and 68 pages 213 and 214 assessment report)

THR category	**10 years**	**20 years**	**30 years**
Men > 65 years			
CePoM (A)	2.4%	3.5%	4.4%
CeLPoM (B)	3.6%	4.9%	5.9%
CeLCoC (C)	3.9%	5.5%	6.7%
HyPoM (D)	2.5%	3.7%	4.6%
CePoC (E)	1.9%	2.9%	3.6%
Women > 65 years			
CePoM (A)	2.0%	3.1%	3.9%
CeLPoM (B)	2.8%	3.8%	4.5%
CeLCoC (C)	2.7%	3.7%	4.4%
HyPoM (D)	1.9%	2.7%	3.3%
CePoC (E)	1.4%	2.3%	3.0%

Table 7: predicted revision rates for men and women under 65 years (tables 69 and 70 pages 216 and 217 assessment report)

THR category	**10 years**	**20 years**	**30 years**
Men <65 years			
CePoM (A)	4.2%	10.3%	18.9%
CeLPoM (B)	6.9%	20.7%	39.0%
CeLCoC (C)	5.4%	14.3%	27.0%
HyPoM (D)	5.3%	13.8%	26.0%
CePoC (E)	2.9%	8.5%	19.7%
Women < 65 years			
CePoM (A)	4.7%	14.3%	28.0%
CeLPoM (B)	4.8%	9.4%	13.8%
CeLCoC (C)	5.2%	14.2%	27.1%
HyPoM (D)	4.5%	14.9%	29.7%
CePoC (E)	3.1%	10.0%	20.3%

4.40 The Assessment Group commented on the revision benchmarks that are set out in NICE technology appraisal guidance 2 (that is, the best prostheses demonstrate a revision rate of 10% or less at 10 years). The Assessment Group stated that a new benchmark lower than 10% at 10 years would now appear appropriate for THR

prostheses, but that RS prostheses may still require considerable improvement to meet the 10 % benchmark. See table 8 below.

Table 8 predicted revision rates at 10 years (Table 71, page 218 assessment report)

Intervention	Population	Revision at 10 years
RS	All NJR patients (n 31,222)	14.4
RS	Matched population (n 26,643)	17.2
RS	Female matched (n 9321)	23.1
RS	Male matched (n 17,322)	12.4
THR	Category A to E matched to RS (n 26,643)	4.7
THR	All NJR patients (n 386,566)	5.2
THR	All CePoM (A) (n 125,285)	2.8
THR	All CeLPoM (B) (n 37,874)	3.9
THR	All CeLCoC (C) (n 34,754)	4.7
THR	All HyPoM (D) (n 28,471)	3.0
THR	All CePoC (E) (n 12,705)	2.1

4.41 The Assessment Group undertook analyses of the NJR Patient Reported Outcomes Measures (PROMS) following THR. The first analysis was of the January 2009 to December 2012 dataset. This data contained 117,044 records with full data on EQ-5D-3L and surgery dates but no age-specific utility values by age. The second PROMS dataset contained EQ-5D-3L data for THR by age and gender for the year 2010/2011 (containing 32,577 records for people over 40 years). The first analysis showed that the average EQ-5D score immediately after the hip operation was 0.767 (0.787 for men and 0.753 for women). For people who required further surgery after their first operation the average EQ-5D score was 0.562 (0.575 for men and 0.553 for women). The Assessment Group also presented the EQ-5D values collected by age and gender (40-50; 50-60; 60-70; 70-80; 80-90), and found that men had a slightly higher EQ-5D utility value than women after their hip operation for all age bands. For both men and women people in the age 60-70 age band, valued their health related quality of life following surgery higher than for any other age band.

Table 9 EQ-5D results for all patients by age band and gender who completed the EQ-5D-3L questionnaire after total hip replacement 2010/2011 (modified from table 17 page 180 assessment report)

	All patients	Men	Women
40-50 years Mean (N)	0.726 (794)	0.736 (316)	0.720 (478)
50-60 years Mean (N)	0.753 (4,352)	0.767 (1,883)	0.742 (2,469)
60-70 years Mean (N)	0.779 (11,106)	0.792 (4,758)	0.769 (6,348)
70-80 years Mean (N)	0.764 (12,308)	0.790 (4,841)	0.747 (7,467)
80-90 years Mean (N)	0.721 (4,017)	0.745 (1,234)	0.710 (2,783)

5 Comments from other consultees

5.1 One patient group provided a statement which detailed an account of one patient's experience of diagnosis, total hip replacement surgery and post-operative care and physiotherapy. The patient group presented this history as illustrative of best practice for pre- and post-operative care but also to highlight issues surrounding diagnosis in primary care and referral to hip replacement specialists. The patient group suggested that referral for surgery decisions should continue to be made on an individual basis; multiple joint involvement is common, hip pathology can be referred to knee pain and multiple joint involvement should be considered in the care pathway. It further highlighted that outcomes for differing types of joint replacement vary and care pathways should be considered separately. The patient group stated that patients' value and derive benefit from good quality pre-operative information and that post operative physiotherapy, supplemented by home based

exercise are an important part of the rehabilitative process and are valued by patients.

6 Cost-effectiveness evidence

6.1 The Assessment Group undertook a systematic review to identify economic evaluations of total hip replacement (THR) and hip resurfacing arthroplasty (RS) published between 2002 and 2012. The Assessment Group identified 66 studies of which only 5 investigated RS. The Assessment Group stated that although the studies they identified confirmed that THR and RS are cost-effective interventions for patients with osteoarthritis of the hip (compared with no surgical intervention) there was limited evidence, and no conclusions could be drawn on different types of hip replacement. However from this systematic review, the Assessment Group identified 4 relevant studies to inform its de novo cost effectiveness analysis (Edlin et al 2012; Pennington et al 2013; Vale et al 2002 and Vanhegan et al 2012 [summarized in table 48 pages 144 to 147 assessment report])

Assessment Group model

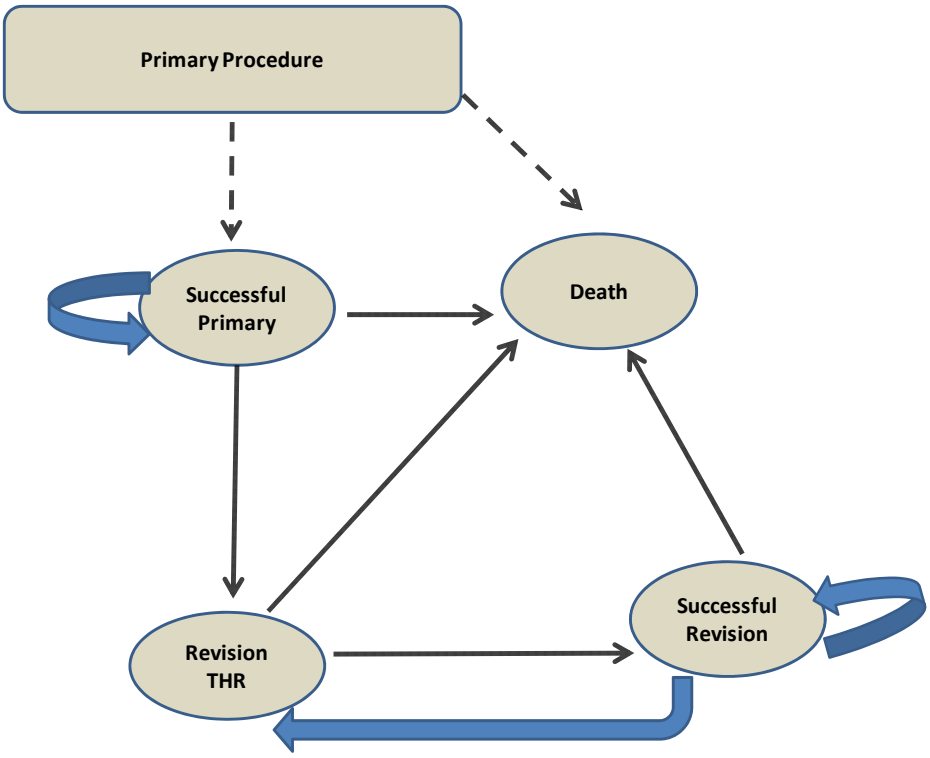
6.2 The Assessment Group developed a Markov model based on Fitzpatrick et al (1998) which the Assessment Group adapted to address the decision problem in this appraisal and up-dated with new data. The model had 4 health states and the cycle length was 1 year. Discounting of 3.5% was applied to both costs and outcomes. The analysis was from the perspective of the NHS and PSS. The Assessment Group reported results for both a lifetime (80 years) and a 10-year time horizon.

6.3 Two cohorts of people entered the model following their primary procedure. One cohort was people for whom RS was suitable and were people in England and Wales who had RS between 2003 and 2012 (propensity matched with people who received THR

categories A-E). The other modeled cohort of people were those for whom RS was unsuitable and were people in England and Wales who had received THR categories A-E between 2003 and 2012.

6.4 People entered the model at the point of their primary procedure (RS or THR) and moved to the either 'successful primary' health state (that is where the initial primary surgery is successful) or death. If primary hip replacement failed people requiring revision moved to the revision total hip replacement state and stayed in that state for 1 cycle (1 year) where all people received THR (rather than RS) as their revision procedure. If revision was successful people moved to the successful revision health state. People in the model could have multiple revisions (no maximum number of revisions). All THR events (time to revision; surgical mortality because of a primary THR, revision THR and re-revision THR; risk of re-revision) were assumed to occur at the beginning of a cycle, mortality due to non-total hip replacement was assumed to occur at the end of a cycle.

Figure 8 structure of Assessment Group model (page 225 assessment report)



6.5 The transition probability between successful primary surgery to revision THR was based on the revision rates calculated and extrapolated from the National Joint Registry (NJR) data (see sections 4.29- 4.36). The transition probability between successful revision and further revision THR was based on revision rates from the New Zealand Joint Registry (risk of re-revision, 0.0326). Surgical mortality was assumed to be 0.5% (based on the NJR annual report [2012]). The Office for National Statistics on death rates in England and Wales were used to determine all-cause mortality by age.

6.6 The utility values applied in the successful primary health state were based on the post operation utility values from the NJR Patient Reported Outcome Measures (PROMs) database. The utility values were grouped by 10-year age band and gender (see section 4.41, table 9) and were adjusted for the increasing age of the cohort after every 10 model cycles. The utility values by age and gender were assumed to be equivalent for people who

received RS or each of the 5 THR categories in the base case. The utility value in the revision THR health state was 0.5624 and was the same for all prosthesis types, age and gender. The utility value for successful revision was assumed to be the same as the utility value for successful primary surgery.

6.7 Costs in the model included the cost of the surgery, prostheses costs, post-operative hospital costs and follow-up costs. The cost of surgery was assumed to be the same for both THR and RS and included the cost of theatre overheads, theatre staff and X-rays. The costs were based on Vale et al (2002) but were updated to 2011/2012 prices using the projected health service cost index. The overall cost of surgery per patient was £2,805. The total costs of primary THR surgery are shown in table 10 below

Table 10 total cost of surgery (table 76, page 231 assessment report)

Resource use	1996 prices		2011/2012 prices
	Primary THR (units)	Total cost (£)	Total cost (£)
Theatre overheads	134	655	1,799
Theatre staff	-	232	637
Number of x-rays	6	134	368
Total cost per patient			£2,805

6.8 The costs of prostheses used for RS and the 5 THR prosthesis categories were obtained from the NHS supply chain. The NHS supply chain provided the average list price costs for 5 manufacturers of the 5 THR categories identified by the Assessment Group and 3 manufacturers of RS prostheses and associated accessories. The Assessment Group further obtained advice from its clinical specialists on which accessories were typically used in clinical practice. For the comparison of RS with THR for people for whom hip RS is suitable, the Assessment Group combined the 5 THR prosthesis categories and weighted the average of these based on the frequency of their use (based on the NJR data). For cemented prostheses there was an additional cost

of cement and cement preparation accessories at a total cost of £203.10 for prostheses in which both the stem and cup were cemented and £163.90 for prostheses in which only the stem was cemented. See table 11 below.

Table 11: Costs of RS and THR prostheses

Prosthesis type	Cost
RS	£2,672
A-E combined	£2,571
A cemented polyethylene cup on metal head	£1,557.38
B cementless HA coated metal cup (polyethylene liner) on metal head (cementless stem)	£3,015.60
C cementless HA coated metal cup (ceramic liner)on ceramic head	£3,868.80
D hybrid – cementless HA coated metal cup (polyethylene liner) on metal head (cemented stem)	£2,649.78
E cemented polyethylene cup on ceramic head (cemented stem)	£1,995.98

6.9 Post-operative ward costs were derived from those reported by Edlin et al which reported the costs of RS and THR over 1 year based on data collected in a randomized controlled trial. The average cost per day of hospital stay was estimated at £296. People who had RS stayed on average for 5.5 days and people who had a THR stayed on average for 5.7 days resulting in overall cost of hospital stay of £1,628 for RS and £1,687 for THR. The outpatient costs for follow up post primary THR or RS were also obtained from Edlin et al. The costs of outpatient care, primary and community care, aids and adaptions provided by the NHS and medications such as pain relief and other medications, inflated from 2009/2010 to 2010/2012 prices were £501 over the first 12 months for RS and £394 over the first 12 months for THR. The follow up costs were applied for all other consecutive years across the lifetime of the model. The costs of follow-up for successful primary THR and RS are shown in table 12 below.

Table 12: cost of follow up for successful primary THR and RS (table 78 page 232 assessment report)

Costs	2009/2010 prices (£)		2011/2012 prices (£)	
	Total cost RS	Total cost THR	Total cost RS	Total cost THR
Outpatient	360	276	383	294
Primary/Community	63	49	67	52
Aids & adaptations	21	21	22	22
Medications	27	24	29	26
Total cost			501	394

6.10 The cost of revision was assumed to be the same for both total THR and RS; however the Assessment Group noted that data from the cohort study of 305 successive revisions following THR in 286 patients between 1999 and January 2008 (Vanhegan et al) showed that the cost of revision was dependent on the reason for revision. For example, surgery for infection and peri-prosthetic fracture resulted in longer operating times, increased blood loss, increase in complications and a longer length of stay in hospital and was therefore associated with higher costs than other reasons for revision such as aseptic revision and dislocation. The costs of revision reported by Vanhegan et al included the cost of the prostheses, materials, theatre cost use of recovery room, inpatient stay physiotherapy, occupational therapy, pharmacy, radiology and laboratory costs, with costs based on NHS 2007/2008 rates by payment by results. The Assessment Group adjusted these costs for inflation to 2011/2012 prices by applying the projected health service cost index. To derive the "average cost of revision" of £16,517, the Assessment Group weighted the mean cost of revision for aseptic loosening, deep infection, peri-prosthetic fracture and dislocation by the number of people who had experienced each of these reasons for requiring a revision in Vanhegan et al (see table 13 below). The follow up costs from Edlin et al that were applied to follow up costs for primary procedures (£394, see table 12 above) were applied in the successful revision health state.

Table 13 Cost of revision (table 77, page 232 assessment report)

Indication	Number of patients	Mean cost (£) (2007/2008 prices)	Mean cost (£) (2011/2012 prices)
Aseptic loosening	194	11,897	13,226
Deep infection	76	21,937	24,387
Peri-prosthetic fracture	24	18,185	20,216
Dislocation	11	10,893	12,109
Weighted average			£16,517

6.11 In the deterministic base case for the population for whom RS was suitable, THR dominated RS (that is, it was less costly and more effective) over both the 10-year and life time horizons (over 10-year time horizon the incremental cost with RS was £10,641 for 0.1317 fewer QALYs; over a lifetime horizon the incremental cost with RS was £11,490 for 0.0879 fewer QALYs). The probabilistic analysis gave very similar results to the deterministic analysis for both time horizons. For further details of the base case deterministic and probabilistic results for the 10-year and life-time horizons, see table 90, on page 243 of the assessment report.

6.12 In both the deterministic and probabilistic analyses in the population for whom RS was not suitable, over a lifetime-time horizon THR category E (cemented polyethylene on ceramic) dominated all of the other THR categories in the incremental analysis. The Assessment Group also presented results for a 10-year time horizon. In both the deterministic and probabilistic analyses, THR category E (cemented polyethylene on ceramic) was more costly and more effective than category A (cemented polyethylene on metal): deterministic incremental cost effectiveness ratio (ICER) £166,217 per QALY gained (incremental costs £299.00 and incremental QALYs 0.0018); probabilistic ICER £225,225 per QALY gained (incremental costs £326.00 and incremental QALYs 0.0014). THR categories D (Hybrid polyethylene on metal), B (cementless polyethylene on metal) and C (cementless polyethylene on metal) were all dominated. The

Assessment Group commented that the difference in QALYs was negligible between the THR categories A to E and that the probabilistic analyses of costs and effectiveness of all categories overlapped markedly confirming the differences were relatively small. For a summary of the results from the deterministic and probabilistic analyses, see table 91, page 247 of the assessment report.

6.13 The Assessment group performed 3 scenario analyses for the population for whom RS was suitable. One scenario analysis tested assumptions used to determine time to revision, and 2 tested assumptions on prosthesis costs. The scenario analyses were as follows:

- Time to revision: the bathtub model was controlled for age and gender (as the age distributions of the matched populations were somewhat removed from a normal distribution) and results presented for the average population 35% female, aged 55.8 years.
- Cost of prostheses (1): weighted average of the highest costing prostheses for each category from the sampled manufacturers (£3,073 for THR and £2,994 for RS).
- Cost of prostheses (2): assessed the weighted average of the lowest costing prostheses for each category from the sampled manufacturers (£2,180 for THR and £2,487 for RS).

In the base case the Assessment Group used the weighted average list price for the 5 THR prosthesis categories across 5 manufacturers (£2,571) and the average list price of RS prostheses across 3 manufacturers (£2,672) sampled by the NHS supply chain. For both the 10-year and life-time horizons, the 3 sensitivity analyses had a minimal effect on incremental costs and QALYs and the results were consistent with the base-

case as THR continued to dominate RS. For a summary of the results of the 3 sensitivity analyses, see table 106 (bathtub model adjusted for age and gender) on page 271 and table 110 (lowest and highest costs of prostheses) on page 283 of the assessment report.

6.14 The Assessment Group performed 7 sensitivity analyses for the population for whom RS was unsuitable. Three tested assumptions used to determine time to revision, 3 tested assumptions on prostheses costs and 1 on the source of utility value for the successful primary and successful revision health states. The scenario analyses were as follows:

- Time to revision bath tub model controlled for age and gender
- Time to revision log normal model used
- Time to revision log normal model controlled for age and gender was used
- Highest list price for all THR prostheses
- Lowest list price for all THR prostheses
- Reduction in costs of 20% across all prostheses categories
- Post operative utility values from a Swedish cohort rather than from the NJR PROMS data

6.15 In the sensitivity analysis in which the bathtub model was adjusted for age and gender, THR category E continued to dominate all other categories over a life time horizon in both the deterministic and probabilistic results. Using the 10-year time horizon, THR category E remained more costly and more effective than category A (deterministic and probabilistic ICERs £127,420 and £176,776 per QALY gained respectively) THR categories D, B and C continued to be dominated (for further details, see table 107 page 274 of the assessment report). For the 2 scenarios in which the lognormal model was used to extrapolate long term revision rates, THR category E is more costly and more effective than category A

in both the 10-year and life time horizons (deterministic ICER £342,781 per QALY gained for 10-year time horizon, deterministic ICER £442,830 per QALY gained for lifetime horizon for the unadjusted model; deterministic ICER £202,741 per QALY gained for the 10-year time horizon, deterministic ICER £227,031 per QALY gained for the lifetime horizon for the lognormal model adjusted for age and gender). In the lognormal model scenario analyses THR categories D, B and C continued to be dominated in both the deterministic and probabilistic results (see table 108 page 277, table 109 page 280 of the assessment report) In all 3 cost sensitivity analyses (highest and lowest costs of THR and a 20% discount applied to each prosthesis category), THR category E continued to dominate all of the other categories over a life-time horizon (both probabilistic and deterministic results). Over a 10-year time horizon, THR category E remained more costly and more effective than category A (highest price deterministic ICER £190,326 per QALY gained, lowest price deterministic ICER £153,663 per QALY gained, 20% discount deterministic ICER £117,489 per QALY gained). THR categories D, B and C continued to be dominated (see tables 111- 113, pages 286-288 of the assessment report). The incremental costs and QALYs from these scenarios for THR categories A and E are provided in table 14 below.

6.16 In the sensitivity analysis around utility values, the Assessment Group used post-operative utility values taken from 32,396 patients from the Swedish Hip Arthroplasty register using a UK EQ-5D tariff. The utility values from this population were 0.77 for people aged 50-60; 0.80 for people aged 60-70; 0.78 for people aged 70-80 and 0.73 for people aged more than 80 years. Over the lifetime-time horizon, THR category E continued to dominate all other categories of THR prostheses in both the deterministic and probabilistic results. Over the 10-year time horizon the deterministic

and probabilistic ICERs for category E compared with category A were £153,067 and £150,644 per QALY gained respectively, categories D, B and C continued to be dominated (see table 114, page 291of the assessment report). The incremental costs and QALYs from this scenario for THR categories A and E are in table 14 below.

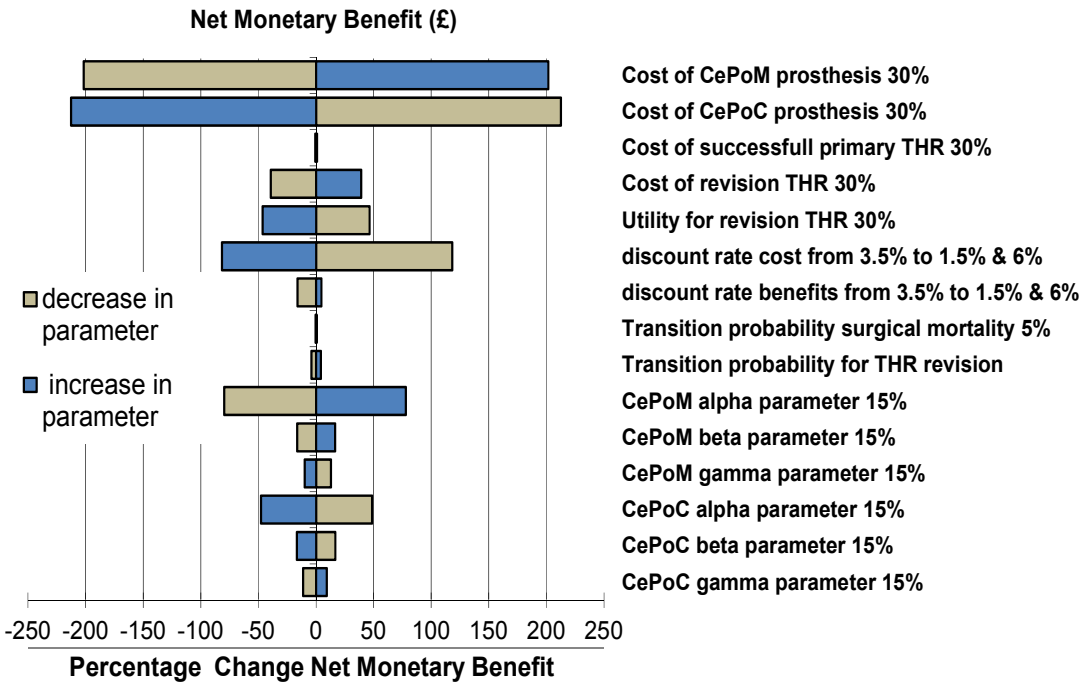
Table 14: Effect of sensitivity analyses on ICER for category A (cemented polyethylene on metal) compared with category E (cemented polyethylene on ceramic) in population for whom RS not considered suitable (data is combined from tables in the assessment report)

Category	Total mean costs £	Total mean QALYs	Comparison	Incremental costs £	Incremental QALYs	ICER (£)
base case: 10-year time horizon (table 91, page 247)						
A	9,444	7.4189	-	-	-	-
E	9,743	7.4207	E v A	299	0.0018	166,217
Base case: life time horizon (table 91, page 247)						
E	14,522	14.7909	-	-	-	-
A	14,801	14.7887	A v E	278	-0.0022	Dominated
bathtub model adjusted for age and gender: 10 year time horizon (table 107 page 274)						
A	9,458	7.4187	-	-	-	-
E	9,731	7.4208	E v A	273	0.0021	127,420
log normal model: 10-year time horizon (table 108, page 277)						
A	9,331	7.4203	-	-	-	-
E	9,690	7.4214	E v A	359	0.0010	342,781
log normal model lifetime horizon (table 108, page 277)						
A	13,476	14.7919	-	-	-	-
E	13,794	14.7926	E v A	318	0.0007	442,830
log normal model adjusted for age and gender: 10 year time horizon (table 109, page 280)						
A	9,349	7.4201	-	-	-	-
E	9,667	7.4217	E vA	318	0.0016	202,741
log normal model adjusted for age and gender lifetime horizon (table 109, page 280)						
A	13,505	14.7917	-	-	-	-
E	13,753	14.7928	E v A	248	0.0011	227,031
highest prostheses costs 10- year horizon (table 111, page 286)						
A	9,675	7.4189	-	-	-	-
E	10,018	7.4207	E v A	343	0.0018	190,326
lowest prostheses costs 10 year time horizon (table 112, page 287)						
A	9,046	7.4189	-	-	-	-
E	9,322	7.4207	E v A	277	0.0018	153,663
price de-escalator of 20% for all THR patients 10 year time horizon (table 113, page 288)						
A	9,132	7.4189	-	-	-	-
E	9,344	7.4207	E v A	212	0.0018	117,489
Swedish utility values used 10 year time horizon (table 114, page 291)						
A	9,444	7.5764	-	-	-	-
E	9,743	7.5783	E v A	299	0.0020	153,067

6.17 The Assessment Group performed further one-way deterministic sensitivity analyses for the comparison of the two cemented THR prosthesis categories E with A in the population for whom RS was

unsuitable. It presented its results in a Tornado diagram of net monetary benefit. Net monetary benefit is the change in QALYs multiplied by a maximum acceptable ICER minus the change in cost. For its analysis the Assessment Group chose a maximum acceptable ICER of £20,000. The Assessment Group noted that the most important factor was cost of the primary prosthesis.

Figure 9 Tornado diagram illustrating sensitivity analysis for net monetary benefit cemented polyethylene on ceramic vs. cemented polyethylene on metal (Figure 71 page 293 assessment report)



6.18 The Assessment Group also performed subgroup analyses for the population for whom RS was suitable in men and women. The Assessment Group presented results for 3 discrete ages within the subgroups separately, applying a weighting to the modeled revision rates for these subgroups (see overview section 4.37) for age for people aged 40, 50 and 60 years (see table 83, page 238 of the assessment report for the age coefficient). For all ages, in both men and women, THR dominated RS over both the 10-year and lifetime horizons horizon. For further details of the deterministic and probabilistic results, see tables 92-93, page 251-252 of the assessment report.

6.19 For people for whom RS was not suitable the Assessment Group presented results for 4 subgroups (men aged over 65, women aged over 65, men aged under 65, and women aged under 65). The Assessment Group presented the results for 5 discrete ages for each subgroup separately. For the under 65 year old subgroups it presented the results for 40, 50 and 60 year olds separately. For the over 65 year old subgroups it presented the results for 70 and 80 year olds. The Assessment Group used the same approach to group people as it had to assess revision rates in subgroups of people for whom RS was not suitable (for the men over 65 subgroup, the modeled cohort had the same characteristics as men over 65 in the NJR who had received the 5 THR categories). It also used the same approach to model revision rates in subgroups (that is,. for people aged over 65 subgroup a lognormal model was used [decreasing hazard of revision over time] whereas for people under the age of 65 subgroup a bathtub model was applied [decreasing hazard of revision over time initially followed by increasing hazard of revision over time]). As the age distribution varied between THR categories and as age may affect revision rates, the Assessment Group controlled for age in each subgroup analysis (that is age was included as a covariate). The Assessment Group applied the age coefficient for each THR prosthesis category to each discrete age (for example, for the age 70 group the revision rates were weighted by the age coefficient for each THR category) in order to produce these results.

6.20 For men and women aged 70 and 80, THR category E was more costly and more effective (QALY difference ranged from 0.0001 and 0.0002) than category A, and categories D, B, C were dominated over the life time horizon. For women under 65 years of age, all categories were dominated by category E. For men aged 40 years, all THR categories were dominated by category A. In men aged 50 or 60 years, category E was more costly and more

effective than category A and categories D, C, B were dominated (for the full results of these subgroup analyses, see tables 96, 100 and 104 on pages 257, 263 and 269 of the assessment report).

Manufacturers' submissions

6.21 Submissions were received from 4 manufacturers (DePuy Synthes, JRI, Smith &Nephew, and Stryker). An economic model was provided by DePuy Synthes.

DePuy Synthes model

6.22 In the model the populations for whom RS was and was not suitable were defined as people who had received RS or THR respectively in the NJR. For the population for RS was suitable the mean age was 55.3 years and 70.9% were men. For the population for whom RS was not suitable the mean age was 70.4 years and 37.5% were men.

6.23 For both the populations for whom RS was, and was not suitable, the manufacturer compared different types of THR prostheses based on methods of fixation: comparing cemented, cementless, hybrid and reverse hybrid. It also assessed 2 of its own brands ([redacted]) however all the results for its own brands were marked as commercial in confidence.by the manufacturer. This overview has focused on the results of the categories rather than the brand results. The manufacturer excluded MoM THR from its analyses stating that THR using this bearing surface are no longer commercially available or used in the UK.

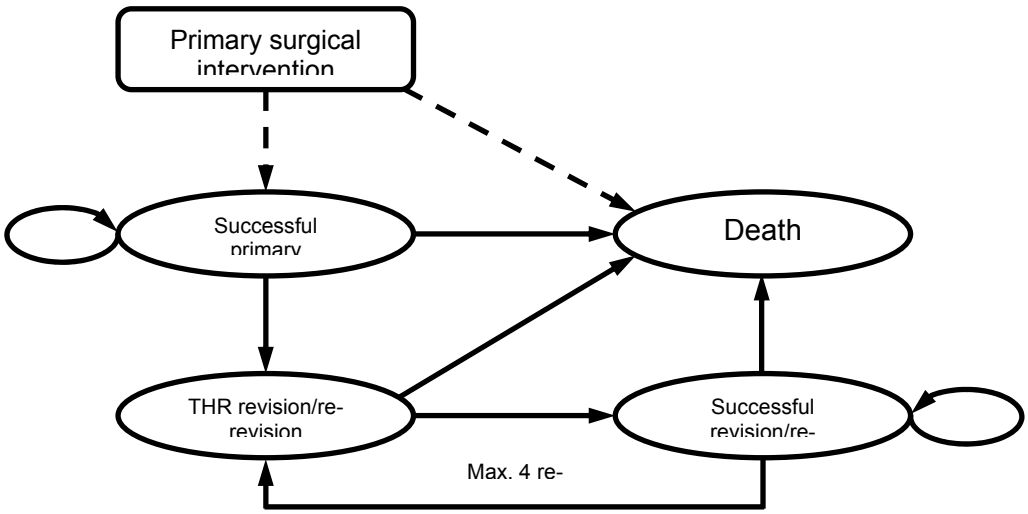
6.24 The manufacturer used a data set of individual patient data (and supplier feedback data for its own prosthesis brands) which covered up to 8 years of follow-up. It filtered the data set to exclude incomplete entries and those in which the indication for surgery

was not osteoarthritis. The manufacturer noted that the NJR cumulative hazard data for revision for each category of prosthesis was not parallel, therefore it fitted individual parametric models for each prosthesis type to extrapolate revision rates from the NJR data. The manufacturer stated that previous models of revision had fitted early and late parametric distributions to extrapolate revision. On consultation with its Advisory Board, the manufacturer classed the reasons for revision observed with its own products as early and later. Reasons for early revision included: dislocation, mismatch, infection, incorrect sizing, and malalignment. Reasons for later revision included prosthesis fracture, lysis, pain, acetabular wear, dissociation liner, soft tissue reaction and other. The manufacturer assessed models that would fit early revisions, late revisions and all revisions combined. The manufacturer stated that a single statistical model was sufficient and it selected the Weibull model (which gave a decreasing hazard over time) as t this had been used in previous evaluations. The manufacturer considered the assumption of a decreasing hazard over time to be realistic for most prosthesis types, although it had concerns regarding this assumption for cemented prostheses as data from the Australian registry had shown that the risk of revision with cemented prostheses increases over time.

6.25 The manufacturer developed a transition state Markov model which had 3 monthly cycles and a lifetime horizon (all patients were assumed to have died by age 100). Costs and outcomes were discounted at 3.5%. The health states in the model were the same as those in the Assessment Group's model but only allowed a maximum of 4 revisions. Like the Assessment Group's model, the rate of re-revision was based on the New Zealand Joint Registry data (rate 0.0331). People stayed in the THR revision/ re-revision health state for 1 cycle. The manufacturer stated that as some patients may receive revision soon after primary surgery, the model

allowed patients to undergo up to 2 interventions in the same cycle. It was assumed that all people would receive the same type of prosthesis in revision surgery. Surgical mortality was assumed to be the same for all prosthesis types (0.5%), and age-and gender adjusted all-cause mortality was applied to all people in the model. Both of these mortality assumptions were the same those applied in the Assessment Group's model. The structure of the manufacturer's economic model is shown in figure 10 below.

Figure 10 Structure of the DePuy Synthes model (figure 16, page 103 of DePuy Synthes submission)



6.26 The model included both the costs of prostheses and surgical costs. The costs of the prostheses were list prices. The total prosthesis costs were: cemented THR £1,029.00; cementless £2,550.50; hybrid £2,011.50; and reverse hybrid £1,568.00 Resurfacing prostheses were assumed to have the same list price as cemented THR (£1,029). For the group "all THR" the manufacturer used a weighted cost (40% cemented, 40% cementless, 17% hybrid; 2% reverse hybrid). Surgery costs were obtained from a micro costing study which included the costs of anaesthetics, surgical consumables, and staff and theatre time. The costs of surgical resources and staff and theatre time were different across prosthesis type. Length of stay was based on NHS reference costs and the length of stay was assumed to be 4.93

days with a unit cost of £295.29. Surgical resource use and costs are shown in table 15 below.

Table 15 surgical resource use costs, staff and theatre time total costs and overall costs of primary procedure (modified from tables 29 and 31 and 33 pages 119-121 DePuy submission). This data is academic in confidence.

	Cemented	Cementless	Hybrid	Reverse-hybrid	All THR	Resurfacing
Total anaesthetic costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total surgical consumables costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Procedure	No. mins	Total cost
Cemented THR	[redacted]	[redacted]
Cementless	[redacted]	[redacted]
Hybrid	[redacted]	[redacted]
Reverse hybrid	[redacted]	[redacted]
All THR	[redacted]	[redacted]
Resurfacing	[redacted]	[redacted]

Prosthesis class	Total costs
RS	[redacted]
Cemented	[redacted]
Cementless	[redacted]
Hybrid	[redacted]
Reverse hybrid	[redacted]
All THR	[redacted]

6.27 Complications were not explicitly modeled. The manufacturer stated that complications that occurred during surgery were reflected in the average cost and health related quality of life associated with each intervention and that prosthesis complications post-surgery were implicitly included within the risk of revision estimates

6.28 The cost of rehabilitation in the first 3 months post-surgery was £467.00 (Payment by Results tariff). The cost of revision was assumed to be double the mean cost of the primary procedure (double the 'All THR' cost) and was £13,399.42. However unlike

the Assessment Group's model, it was not assumed that the cost of revision was dependent on the reason for revision.

6.29 The manufacturer performed a systematic review to identify utility values. For its base case, the manufacturer used utility values from Rolfson et al (2011, Swedish registry) as the study had a large sample size (32,396 patients) and reported the pre-operative and 1 year post-operative utility values .The pre-operative utility value was 0.41, and the post operative utility value was 0.78. A disutility of 0.145 was applied to the post-operative utility value following revision to reflect the lower quality of life associated with a subsequent surgical intervention. This value was based on a separate study (Briggs et al, 2003). The utility values used in the manufacturer's model were different to those used in the Assessment Group's model in which it was assumed that post successful primary surgery and post successful revision utility values would be the same (see section 6.6).

6.30 In the manufacturer's base case for the population for whom RS is suitable, THR (all types combined) dominated RS. The total incremental costs of RS were £2504.31 for 0.106 fewer QALYS. The manufacturer presented an incremental analysis of the results for cemented, cementless, hybrid, reverse hybrid and resurfacing prosthesis categories, alongside the results for 2 of the manufacturer's own products and all THR prostheses combined.

[redacted]. Table 16 presents the results of the incremental analysis for the THR prosthesis categories only (that is, without the results for the manufacturer's own products and all THR prostheses

combined). For the results of the manufacturer's incremental analysis which included the manufacturer's own products and all THR prostheses combined see table 38, page 126 of DePuy Synthes' submission. The manufacturer noted that the range of QALYs generated by the probabilistic analysis overlapped substantially between the THR prosthesis categories and concluded that all categories of THR could be considered equivalent.

Table 16 Base case results patients for whom RS suitable (this table shows incremental results which have been recalculated for categories only. For DePuy brands and all-THR category see table 38 page 126 DePuy Synthes submission. Probabilistic costs and QALYs are from table 44 page 139 DePuy Synthes submission).

Technology	Costs	QALYs	Incremental analysis
			ICER
Cemented	£8,231	11.145	-
Reverse hybrid	£8,570	11.148	Extendedly dominated (reverse hybrid vs. cemented £113,000)
Cementless	£8,743	11.146	Dominated
Hybrid	£8,817	11.167	Hybrid vs. cemented £26,636
Resurfacing	£11,399	11.009	Dominated

Probabilistic				
Technology	Costs		QALYs	
Cemented	£8,240	(6,484 - 10,073)	11.145	(11.08 - 11.21)
Reverse-hybrid	£8,596	(6,740 - 10,450)	11.146	(11.07 - 11.22)
Cementless	£8,747	(7,068 - 10,482)	11.146	(11.08 - 11.21)
Hybrid	£8,826	(7,092 - 10,588)	11.166	(11.1 - 11.23)
Resurfacing	£11,408	(9,138 - 13,830)	11.009	(10.93 - 11.09)

6.31 For the population for whom RS is unsuitable, the manufacturer presented an incremental analysis of the results for cemented, reverse hybrid, cementless hybrid prostheses alongside the results for 2 of the manufacturer's own products and all THR prostheses combined. Table 17 below, shows the results of the incremental analysis for the THR prosthesis categories only (that is, without the results for the manufacturer's own products and all THR prostheses combined). For the results of the manufacturer's incremental analysis which included the manufacturer's own products and all THR prostheses combined see table 39, page 126 of the DePuy

submission. The manufacturer noted that the range of QALYs generated by the probabilistic analysis (10,000 simulations) overlapped substantially between the THR prosthesis categories and concluded that all categories of THR could be considered equivalent.

Table 17 base case for population for whom RS not suitable (this table shows incremental results recalculated for categories only. For DePuy brands and all- THR category see table 39 page 126 DePuy Synthes submission). Probabilistic costs and QALYs are from table 45 page 140 DePuy Synthes submission.

Technology	Costs	QALYs	Incremental analysis
			ICER
Cemented	£7,709	8.811	-
Reverse hybrid	£8,158	8.805	Dominated
Cementless	£8,383	8.799	Dominated
Hybrid	£8,488	8.814	Hybrid vs. cemented £259,667

Probabilistic				
Technology	Costs		QALYs	
Cemented	£7,713	(6,118 - 9,409)	8.811	(8.76 - 8.86)
Reverse-hybrid	£8,171	(6,494 - 9,937)	8.804	(8.75 - 8.85)
Cementless	£8,387	(6,823 - 10,029)	8.799	(8.75 - 8.85)
Hybrid	£8,498	(6,872 - 10,216)	8.814	(8.76 - 8.86)

6.32 The manufacturer conducted the following sensitivity analyses for the population for whom RS was suitable and for the population for whom RS was unsuitable:

- NHS reference costs for hip replacement procedures
- EQ-5D from NJR PROMs rather than Swedish registry data.
- Exponential model rather than Weibull model used to derive transition probabilities for revision that were independent of time.
- Weibull model used to derive transition probabilities for revision stratified by age at primary procedure < 70 years.
- Weibull model used to derive transition probabilities for revision stratified by age at primary procedure < 55 years.

The results of the sensitivity analyses for each THR prosthesis category and RS are shown in table 18.

Table 18 Scenario analyses for populations for whom resurfacing is suitable and unsuitable (modified from table 40 and table 41, pages 129-130 DePuy submission- For results of DePuy Synthes brand and 'all THR' see the DePuy Submission)

Resurfacing suitable												
Technology	Base case		Scenario 1: NHS Reference Costs		Scenario 2: PROMS		Scenario 3: Exponential model		Scenario 4: Age <70 years model		Scenario 5: Age <55 years Model	
	Costs	QALYs	Costs	QALYs	Costs	QALYs	Costs	QALYs	Costs	QALYs	Costs	QALYs
Cemented	£8,231	11.145	£7,642	11.145	£8,231	10.886	£8,476	11.126	£8,330	11.138	£8,454	11.127
Reverse hybrid	£8,570	11.148	£7,620	11.148	£8,570	10.889	£9,011	11.112	£8,595	11.147	£8,570	11.148
Cementless	£8,743	11.146	£7,618	11.146	£8,743	10.886	£9,416	11.090	£8,831	11.138	£8,950	11.128
Hybrid	£8,817	11.167	£7,521	11.167	£8,817	10.907	£9,187	11.137	£8,872	11.163	£8,840	11.167
Resurfacing	£11,399	11.009	£10,087	11.009	£11,399	10.749	£11,560	10.997	£11,418	11.008	£11,569	10.999

Resurfacing not suitable												
Technology	Base case		Scenario 1: NHS Reference Costs		Scenario 2: PROMS		Scenario 3: Exponential model		Scenario 4: Age <70 years model		Scenario 5: Age <55 years model	
	Costs	QALYs	Costs	QALYs	Costs	QALYs	Costs	QALYs	Costs	QALYs	Costs	QALYs
Cemented	£7,709	8.811	£7,321	8.811	£7,709	8.607	£7,823	8.804	£7,779	8.806	£8,209	8.772
Reverse hybrid	£8,158	8.805	£7,354	8.805	£8,158	8.600	£8,416	8.788	£8,323	8.792	£8,158	8.805
Cementless	£8,383	8.799	£7,379	8.799	£8,383	8.595	£8,771	8.773	£8,373	8.801	£8,420	8.798
Hybrid	£8,488	8.814	£7,297	8.814	£8,488	8.609	£8,704	8.800	£8,528	8.811	£8,392	8.823

6.33 The manufacturer performed additional one way sensitivity analyses for the population for whom RS was suitable presenting the results in terms of net monetary benefit, assuming a maximum acceptable ICER of £20,000. There was a positive net monetary benefit associated with THR for all parameter values tested and in most instances the relative net benefit between the low and high values was minimal. A positive net monetary benefit means that a technology is cost effective if a maximum acceptable ICER of £20,000 is considered appropriate. The most influential parameters were the cost of revision, the utility decrement associated with revision and resource use items such as the cost of follow up appointments, the overhead cost per theatre hour, and the individual costs of prostheses components.

Additional data identified by and comments from manufacturers of relevance to the Assessment Group model assumptions.

6.34 Two manufacturers considered whether the granularity of categorisation by fixation method was appropriate. Smith and Nephew presented data from the Australian registry to highlight revision rates when prostheses were grouped by bearing surface. [redacted]. JRI commented that within the category of cementless prostheses, success rates vary by surface properties of the component onto which the bone grows.

6.35 Additional estimates to those used by the Assessment Group for the cost of revision and for the proportion of people who

experienced each cause of revision were provided in the manufacturers' submissions. The DePuy Synthes Submission reported the most common reasons for revision in 2011 as aseptic loosening (42% of revisions); pain (24% of revisions); osteolysis (13% of revisions) and dislocation/ subluxation (13% of revisions). Smith and Nephew suggested the cost of revision THR is about 38% higher than primary THR and the cost of aseptic revision in the UK to be £9048 and lower than septic revision £14441 (15% of all revisions). Stryker quoted the same study for costs of revision by indication as the Assessment Group. JRI reported the revision rates for different fixation types, and for JRI Cemented THR, by incidence per 1000-life years for revision caused by pain, dislocation, infection, aseptic loosening, lysis and periprosthetic fracture. Stryker stated that resurfacing procedures are more likely to be revised for pain, loosening, periprosthetic fracture and other reasons than non-metal-on-metal bearing surfaces. A summary of the theses additional estimates are provided in table 19 below.

Table 19 causes of revision and costs of revision

Assessment Group and Stryker			**DePuy**			**Smith and Nephew**		
Reason for revision	%	Cost	Reason for revision	%	Cost	Reason for revision	%	Cost
Aseptic Loosening	63%	£13,226	Aseptic loosening	42%	-	Aseptic revision	85%	£9048
Deep infection	25%	£24,387	- -			Septic revision	15%	£14441
Peri-Prosthetic fracture	8%	£20,216						
Dislocation	4%	£12,109	Dislocation subluxation	13%	-			
			Pain	24%	-			
			Osteolysis	13%	-			
Vanhegan (2012)			NJR 9th annual report			Jenkins et al 2013:cost effectiveness analysis		

JRI revision rate (per 1000 life year) by fixation for each indication for revision (NJR 9th annual report 2012)						
Fixation	Pain	Dislocation /subluxation	infection	Aseptic loosening	lysis	Periprosthetic fracture
All cemented	0.46	0.87	0.78	0.86	0.17	0.27
All Hybrid	0.67	1.23	0.73	0.81	0.2	0.56
All cementless	1.57	1.37	0.95	1.85	0.26	0.8
JRI cementless	0.66	1.37	0.47	1.43	0.12	0.56

6.36 Several manufacturers highlighted that the NJR data may not be sufficiently mature to capture changes in risk over time. For example one manufacturer noted that the Australian Joint Registry had shown that the risk of revision with cemented THR changed over time. Stryker noted that although cemented prostheses are associated with fewer revisions in the first 3 years after surgery after this time there is no difference in the rate of revisions between cemented and cementless prostheses. JRI commented on data from the Swedish and Australian registries and reported in the NJR 9th registry report (published 2012). It commented that all 3 registries showed that the revision rates of cemented THR were lower than cementless THR, but in a graph of the proportion of people using Swedish registry data the revision rates crossed over after 8 years (at 8 years risk of revision over time was greater with

cemented than cementless). JRI stated that the Australian Joint Registry showed that cementless THR has a higher revision rate than cemented THR in the first month but after 6 years the rate of revision for cemented THR is higher than for cementless THR.

7 Comments received during consultation of the Assessment Report.

7.1 Summary of the comments received:

- **Categories of THR:** An additional suggested category of THR of interest was cementless stem with a cemented polythene on ceramic articulation (reverse hybrid). One consultee stated that cross linked polyethylene has been shown in the Assessment Report to have lower wear rates and would benefit from separate analysis.
- **Care pathway**: Two Consultees commented on when people would be followed up. Both stated most patients are followed up at 6 weeks, 1 consultee said there is a further review at 1 year and very few patients are reviewed at day 10 as noted by the Assessment Group (see figure 2, page 40 of the assessment report). Most units follow patients at 5 yearly intervals unless there are problems (such as with MoM prostheses) but highlighted that there is variation around the country with some people being only offered 1 post-operative follow up appointment.
- **Longer term revision estimates:** One consultee stated the most common cause for early dislocation is mal-positioning, whereas wear and soft tissue problems are more common in late dislocation. Consultees noted that there are several prostheses in the NJR data that have been recalled and asked whether these prostheses types have been removed from the analysis (e.g. the DePuy ASR MoM resurfacing and total hip replacement

products that were recalled in 2010 for higher than expected revision rates). One consultee suggested that the Assessment Group estimate of 12.9% revision at 20 years may be an underestimation given revision rates in the Swedish registry with 20 year follow up were 30% in the 50-59 age category. However, they also noted a steepening of the revision curves from the Swedish registry at around 8 years suggesting that the bathtub model used by the Assessment Group may be appropriate.

- **List price prostheses costs:** One consultee suggested that only around 1 in 5 of hip prostheses is purchased through the NHS supply Chain.
- **Other costs:** Theatre time and inpatient time (which can account for a third of the cost of a joint replacement and can vary between institutions) may differ between RS and THR and between cemented and cementless THR (surgery time). Revision THR following RS can use a primary THR which is less costly than a revision prosthesis. Rapid rehabilitation techniques have halved inpatient stay in recent years but this is likely to be independent of hip prosthesis designs.
- **Other points for consideration:** The choice of implant may be dependent on factors other than patient characteristics. Surgeons and surgical units may only use one type of implant. Patient choice should be considered. One consultee noted that younger people were more likely to be offered new 'innovative' treatments, which because of the surgeons' learning curve may result in a higher revision rate, calling this effect 'innovation bias'. One consultee noted that although propensity matching by age and gender was appropriate, age is not a complete proxy for activity level and furthermore BMI and surgeon grade are also recorded in the NJR and these also could have been used by the Assessment Group in its propensity matching.

8 Equalities issues

8.1 No equalities issues were raised in the assessment report, the manufacturer's submissions or during the consultation on the Assessment Report.

9 Innovation

9.1 DePuy Synthes stated that dislocations can impact on a patient's quality of life but it is often treated without the revision of components but with a closed reduction under anaesthetic. The NJR does not report on the rate of hip dislocations which may impact on patients' quality of life. Some hip replacement systems have been designed to decrease risk of dislocation (that is, cementless cups that can accommodate larger heads). There have been developments in the materials used for bearing surfaces to reduce wear such as the use of cross linked polyethylene rather than non cross linked polyethylene. The NJR does not stratify revision rates by generation of prosthesis within a prosthesis class.

10 Authors

Mary Hughes
Technical Lead

Nicola Hay
Technical Adviser

Appendix A: Supporting evidence

Related NICE guidance

Published

- Rheumatoid arthritis: the management of rheumatoid arthritis in adults. NICE clinical guideline 79(2004).
- Osteoarthritis: the care and management of osteoarthritis in adults. NICE clinical guideline 59 (2008). Review in progress, earliest anticipated date of publication Nov 2013
- Hip disease replacement prostheses. NICE technology appraisal guidance 2(2000).
- Hip disease- metal on metal resurfacing. NICE technology appraisal guidance 44 (2002)
- Minimally invasive two-incision surgery for total hip replacement. NICE interventional procedure guidance 112(2005).
- Arthroscopic femoro-acetabular surgery for hip impingement syndrome. NICE interventional procedure guidance 408 (2011)